Chapter 2
Physical Fundamentals of Sound Fields

The present chapter outlines the mathematical and physical tools that are employed in the subsequent chapters. It is not written in a tutorial style but serves rather as a reference.

2.1 The Wave Equation

2.1.1 General

In order for a sound field $s(\mathbf{x}, t)$ to be physically possible it has to satisfy the *scalar wave equation* in the domain (i.e., the volume) of interest. When *source-free* domains are considered, the wave equation is termed *homogeneous* and is given by (Williams 1999, Eq. (2.1), p. 15)

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

c denotes the speed of sound in air, which is assumed to be 343 m/s throughout this book. The zero on the right hand side of (2.1) indicates the absence of sources. Consistently, when (2.1) exhibits a source term on the right hand side, it is termed *inhomogeneous*. The relation between the source term and the evolving sound field is discussed in (Williams 1999, Sect. 8.6).

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

Assuming *steady-state* conditions and harmonic time dependence and applying a temporal Fourier transform as defined in Appendix B to the time domain wave

J. Ahrens, *Analytic Methods of Sound Field Synthesis*,
T-Labs Series in Telecommunication Services, DOI: 10.1007/978-3-642-25743-8_2,

equation (2.1) yields the *scalar Helmholtz equation*, which is given by (Williams 1999)

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

Equation (2.2) will play a central role in this book. k is termed wavenumber (although it is rather a coefficient than a number) and is of unit rad/m. It is related to the radian frequency ω via

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

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

$$\lambda = \frac{c}{f} = \frac{2\pi}{k}. \tag{2.4}$$

Note that exclusively sound propagation in *homogeneous* and *non-dissipative* (i.e., lossless) media is considered throughout this book. Additionally, the wave equation (2.1) and the Helmholtz equation (2.2) assume that the medium, which is air in the present case, is perfectly *linear*. A system is linear when a linear combination of two or more input signals, e.g., $x_1(t)$ and $x_2(t)$, leads to an according linear combination of the output signals $y_1(t)$ and $y_2(t)$ of the individual input signals as (Girod et al. 2001)

$$\begin{aligned} x_1(t) &\longrightarrow y_1(t) \\ x_2(t) &\longrightarrow y_2(t) \\ a x_1(t) + b x_2(t) &\longrightarrow a y_1(t) + b y_2(t). \end{aligned}$$

This assumption is only met for infinitesimal sound pressures so that (2.1) and (2.2) are essentially approximations. However, it has been shown that (2.1) and (2.2) provide useful results when sound pressures are considered that are below the threshold of pain of the human auditory system (Gumerov and Duraiswami 2004, pp. 2–3). In (Jessel 1973, p. 25, 27), the threshold of pain is settled around a sound pressure level of 120 dB, and the limits of the assumption of linearity around 160 dB.

The non-linearity can be illustrated intuitively by noting that the atmospheric pressure can be increased by, say, two times the atmospheric pressure but can obviously not be reduced by the same amount, which is clearly a non-linear behavior.

Frequently, monopole sound sources will be considered in this book the spatial transfer function of which exhibits a pole at the location of the source. Around this pole, the limits of linearity will obviously be exceeded. This exceedance will not be explicitly considered in this book since these monopoles serve as a simplification of real-world sound sources the spatial transfer functions of which do not exhibit such poles.

2.1.2 Solutions in Cartesian Coordinates

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

$$\nabla = \frac{\partial}{\partial x}\mathbf{e}_x + \frac{\partial}{\partial y}\mathbf{e}_y + \frac{\partial}{\partial z}\mathbf{e}_z, \tag{2.5}$$

whereby $\mathbf{e}_i$ denotes the unit vector in indexed direction, i.e.

$$\mathbf{e}_x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \quad \mathbf{e}_y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad \mathbf{e}_z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{2.6}$$

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

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

$$S(\mathbf{x}, \omega) = \hat{S}(\omega) e^{-i(k_x x + k_y y + k_z z)} = \hat{S}(\omega) e^{-i\mathbf{k}^T \mathbf{x}}, \tag{2.7}$$

and thus constitute plane waves.

Equation (2.7) is satisfied as long as the *dispersion relation*

$$k^2 = k_x^2 + k_y^2 + k_z^2 \tag{2.8}$$

is fulfilled. Thus, the wavenumber k represents the length of the propagation vector $\mathbf{k} = [k_x\ k_y\ k_z]^T$. Equation (2.8) can be rearranged to read

$$k_y^2 = k^2 - k_x^2 - k_z^2. \tag{2.9}$$

Note that there is no restriction on the values of k_x^2 and k_z^2 in (2.9) provided that they are real (Williams 1999, p. 21). Taking the square root of (2.9) yields

$$k_y = \begin{cases} \pm\sqrt{k^2 - k_x^2 - k_z^2} & \text{for } k^2 \geq k_x^2 + k_z^2 \\ \pm i\sqrt{k_x^2 + k_z^2 - k^2} & \text{for } k_x^2 + k_z^2 \geq k^2 \end{cases} \tag{2.10}$$

since k is non-negative.

The first case in (2.10) represents a *propagating* or *homogeneous* plane wave. The vector $\mathbf{k}$ points into the direction of propagation. Refer to Fig. 2.1a for a simulation.

The second case in (2.10) (with complex k_y) represents an *evanescent* or *inhomogeneous* wave. Inserting k_y into (2.7) yields

$$S_{\mathrm{pw}}(\mathbf{x}, \omega) = \hat{S}_{\mathrm{pw}}(\omega)\ e^{\pm\sqrt{k_x^2 + k_z^2 - k^2}\, y}\ e^{-i(k_{\mathrm{pw},x} x + k_{\mathrm{pw},z} z)} \tag{2.11}$$

Note that the first exponential in (2.11) is purely real. For $y > 0$, the positive exponent in the first exponential in (2.11) is non-physical since it blows up for $y \to +\infty$ so that the solution is restricted to the decaying term (the negative exponent) for this case (Williams 1999). Refer to Fig. 2.1b for a simulation of an evanescent wave decaying in y direction.

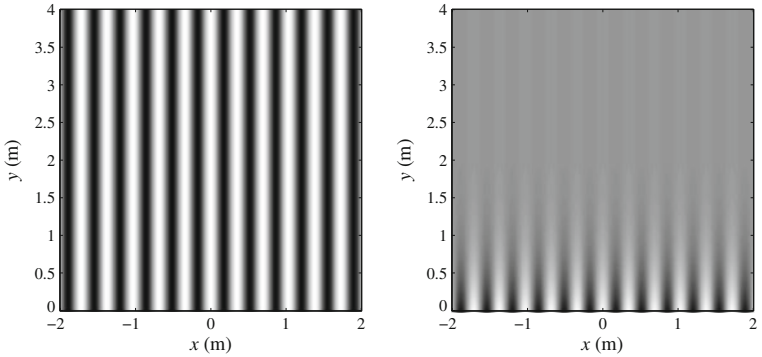

Fig. 2.1 Propagating and evanescent waves of frequency $f_{\mathrm{pw}} = 1000\,\mathrm{Hz}$. A cross-section through the horizontal plane is shown. **a** Propagating plane wave; $\mathbf{k}_{\mathrm{pw}} = [k \;\; 0 \;\; 0]^T$. **b** Evanescent wave; $\mathbf{k}_{\mathrm{pw}} = \left[\sqrt{1.01}k \;\; -i\sqrt{0.01}k \; 0\right]^T$

2.1.3 Solutions in Spherical Coordinates

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

$$\nabla = \frac{\partial}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial}{\partial \beta}\mathbf{e}_\beta + \frac{1}{r\sin\beta}\frac{\partial}{\partial \alpha}\mathbf{e}_\alpha. \tag{2.12}$$

with

$$\mathbf{e}_r = \begin{bmatrix} \cos\alpha\sin\beta \\ \sin\alpha\sin\beta \\ \cos\beta \end{bmatrix}; \quad \mathbf{e}_\beta = \begin{bmatrix} \cos\alpha\cos\beta \\ \sin\alpha\cos\beta \\ -\sin\beta \end{bmatrix}; \quad \mathbf{e}_\alpha = \begin{bmatrix} -\sin\alpha \\ \cos\alpha \\ 0 \end{bmatrix}. \tag{2.13}$$

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

$$S(\mathbf{x}, \omega) = \Pi(r)\Theta(\alpha)\Phi(\beta). \tag{2.14}$$

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

$$h_n^{(1,2)}\left(\frac{\omega}{c}r\right) = j_n\left(\frac{\omega}{c}r\right) \pm i\, y_n\left(\frac{\omega}{c}r\right). \tag{2.15}$$

Refer to Fig. 2.2 for illustrations.

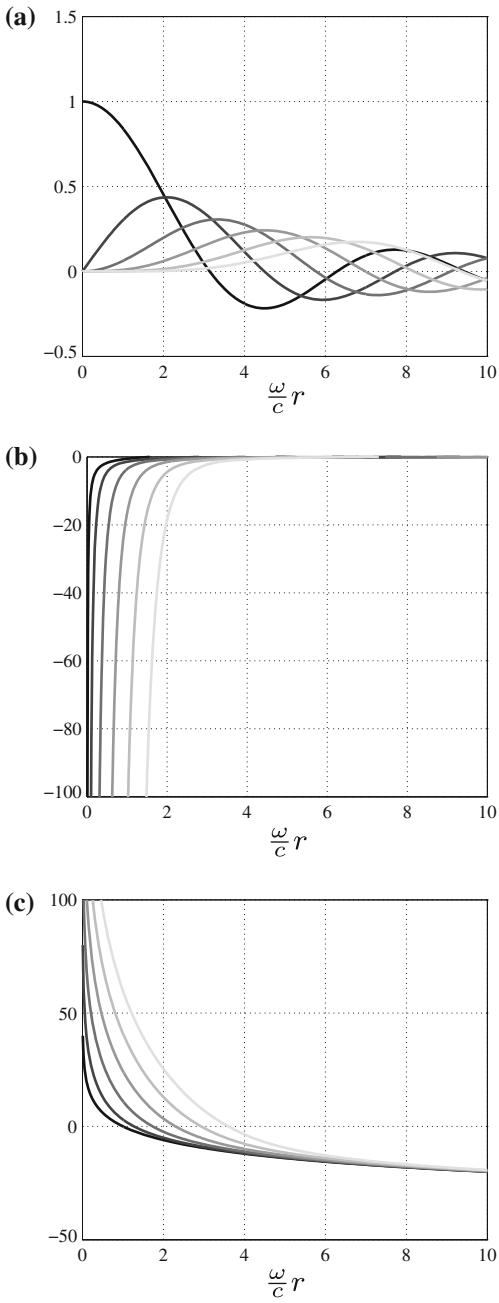

Fig. 2.2 Bessel, Neumann, and Hankel functions for $0 \leq n \leq 5$. Brighter color indicates a higher order n. **a** $j_n(\omega/c\, r) = \Re\left\{h_n^{(1,2)}(\omega/c\, r)\right\}$. **b** $y_n(\omega/c\, r) = \pm\,\Im\left\{h_n^{(1,2)}(\omega/c\, r)\right\}$. **c** $20\log_{10}\left|h_n^{(1,2)}(\omega/c\, r)\right|$

It can be shown that $h_n^{(2)}(\omega/c\, r)$ represents outgoing waves and $h_n^{(1)}(\omega/c\, r)$ represents incoming waves for the definition of the Fourier transform used in this book (refer to Appendix B). $j_n(\omega/c\, r)$ represents transitory ("passing") waves and

$$j_n\left(\frac{\omega}{c}r\right) = \frac{1}{2}\left(h_n^{(1)}\left(\frac{\omega}{c}r\right) + h_n^{(2)}\left(\frac{\omega}{c}r\right)\right) \tag{2.16}$$

holds.

Useful recursion relations are (Gumerov and Duraiswami 2004, Eq. (2.1.86))

$$\frac{2n+1}{x} f_n(x) = f_{n-1}(x) + f_{n+1}(x) \tag{2.17}$$

and (Gumerov and Duraiswami 2004, Eq. (2.1.87))

$$(2n+1) f_n'(x) = n f_{n-1}(x) - (n+1) f_{n+1}(x), \tag{2.18}$$

whereby f_n can be any of j_n, y_n, $h_n^{(1)}$ or $h_n^{(2)}$. The prime denotes differentiation with respect to the argument.

In certain situations the large-argument approximation of the spherical Hankel functions given by (Williams 1999, Eq. (6.68), p. 197)

$$h_n^{(1,2)}\left(\frac{\omega}{c}r\right) \approx (\mp i)^{(n+1)} \frac{e^{\pm i\frac{\omega}{c}r}}{\frac{\omega}{c}r} = (\mp i)^n h_0^{(1,2)}\left(\frac{\omega}{c}r\right) \qquad \forall \frac{\omega}{c}r \to +\infty \tag{2.19}$$

will be employed in order to simplify problems. Since the argument of the spherical Hankel function is composed of a product of the angular frequency ω and the distance r in the present context, (2.19) constitutes a far-field/high-frequency approximation.

The azimuthal solutions $\Theta(\alpha)$ in (2.14) are given by the complex exponential functions $e^{im\alpha}$ with $m \in \mathbb{Z}$ and the colatitudinal solutions $\Phi(\beta)$ are given by the *associated Legendre functions* $P_n^m(\cos\beta)$. A selection of the exponential functions is illustrated in Fig. 2.3a; a selection of the associated Legendre functions is illustrated in Fig. 2.3b. The latter are purely real.

The associated Legendre functions $P_n^m(z)$ vanish for $|m|>n$ and satisfy (Gumerov and Duraiswami 2004, Eq. (2.1.46), p. 47)

$$P_n^m(-z) = (-1)^{n+m} P_n^m(z) \tag{2.20}$$

A useful recurrence relation is (Gumerov and Duraiswami 2004, Eq. (2.1.53), p. 48)

$$(1-z^2)\frac{\partial}{\partial z} P_n^m(z) = \frac{(n+1)(n+m)}{2n+1} P_{n-1}^m(z) - \frac{n(n-m+1)}{2n+1} P_{n+1}^m(z), \tag{2.21}$$

and a frequently required special value is (Gumerov and Duraiswami 2004, Eq. (2.1.43), p. 46)

$$P_n^0(1) = 1. \tag{2.22}$$

Both the exponentials $e^{im\alpha}$ and the associated Legendre functions $P_n^m(z)$ are orthogonal for a given order m.

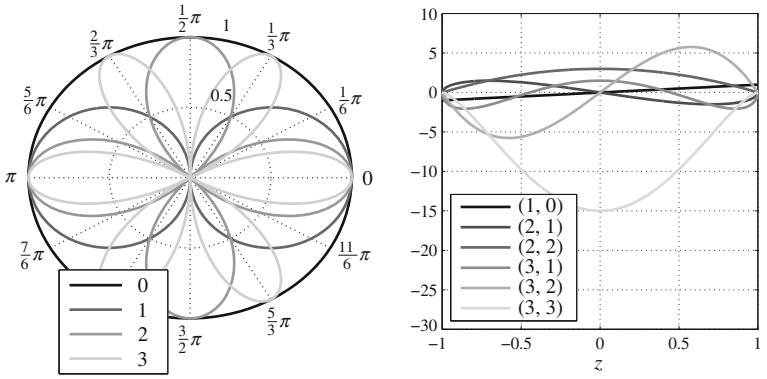

Fig. 2.3 Illustration of complex exponential functions and associated Legendre functions. **a** $\left|\Re\left\{e^{im\alpha}\right\}\right|$ for a selection of m. **b** $P_n^m(z)$ for a selection of (n, m)

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

$$Y_n^m(\beta, \alpha) = (-1)^m \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos\beta)\, e^{im\alpha}. \tag{2.23}$$

Like the associated Legendre functions, spherical harmonics $Y_n^m(\beta, \alpha)$ vanish for $|m| > n$. Refer to Fig. 2.4 for an illustration of selected spherical harmonics.

Note that other variants of the definition (2.23) exist, which differ mainly with respect to the factor $(-1)^m$, e.g. (Condon and Shortley 1935; Arfken and Weber 2005; Williams 1999). The choice of this factor is not essential but is rather made upon practical considerations.

The advantage of definition (2.23) is the fact that it inherently handles negative m and avoids the case differentiation that is required in alternative definitions. Furthermore, the complex conjugate $Y_n^m(\beta, \alpha)^*$ can be expressed by negating the degree m as (Gumerov and Duraiswami 2004)

$$Y_n^m(\beta, \alpha)^* = Y_n^{-m}(\beta, \alpha). \tag{2.24}$$

Spherical harmonics are orthonormal so that the relation

$$\int_0^{2\pi} \int_0^{\pi} Y_n^m(\beta, \alpha) Y_{n'}^{-m'}(\beta, \alpha) \sin\beta \, d\beta \, d\alpha = \delta_{nn'}\delta_{mm'} \tag{2.25}$$

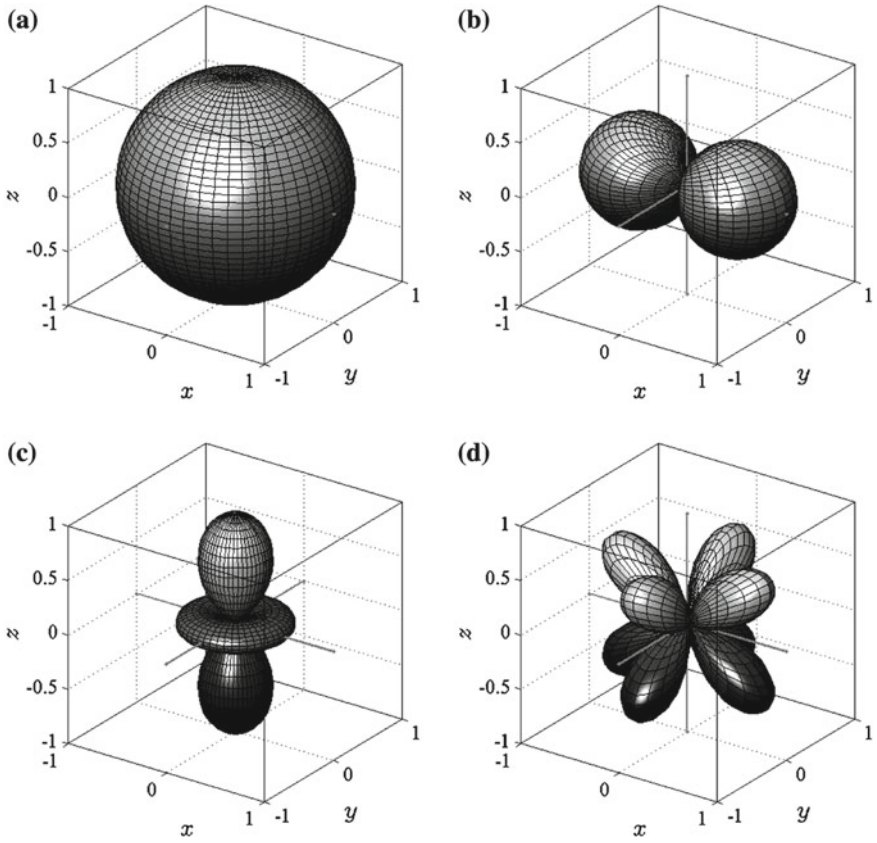

Fig. 2.4 $\left|\Re\{Y_n^m(\beta,\alpha)\}\right|$ for a selection of n and m. **a** $n=0$, $m=0$. **b** $n=1$, $m=1$. **c** $n=2$, $m=0$. **d** $n=3$, $m=-2$

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

$$\delta_{nn'} \begin{cases} 1 & \text{for } n = n' \\ 0 & \text{for } n \neq n' \end{cases} . \tag{2.26}$$

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

$$\sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_n^m(\beta,\alpha) Y_n^{-m}(\beta',\alpha') = \delta(\alpha-\alpha')\delta(\beta-\beta'). \tag{2.27}$$

Assuming

$$E_n^m(\mathbf{x}) = h_n^{(1,2)}\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha)$$
$$I_n^m(\mathbf{x}) = j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta, \alpha),$$

the relations (Gumerov and Duraiswami 2004, Eq. (3.2.7), p. 96)

$$E_n^m(-\mathbf{x}) = (-1)^n E_n^m(\mathbf{x}), \quad I_n^m(-\mathbf{x}) = (-1)^n I_n^m(\mathbf{x}). \tag{2.28}$$

hold.

The addition theorem for spherical harmonics is given by (Gumerov and Duraiswami 2004, Eq. (2.1.70), p. 53)

$$P_n^0(\cos\gamma) = \frac{4\pi}{2n+1}\sum_{m=-n}^{n} Y_n^{-m}(\beta_{\text{or}}, \alpha_{\text{or}})\, Y_n^m(\beta, \alpha), \tag{2.29}$$

with γ denoting the angle between $(\alpha_{\text{or}}, \beta_{\text{or}})$ and (α, β). Occasionally, the relation

$$Y_n^0(0, 0) = \sqrt{\frac{2n+1}{4\pi}}, \tag{2.30}$$

is exploited.

2.2 Representations of Sound Fields

2.2.1 Representation of Sound Fields as Series of Spherical Harmonics

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

$$S(\mathbf{x}, \omega) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \mathring{S}_n^m(r, \omega) Y_n^m(\beta, \alpha). \tag{2.31}$$

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

It can be shown that *interior* and *exterior* problems have to be considered separately (Williams 1999, p. 207, 217). Interior problems are problems that consider domains that are free of sound sources and obstacles, i.e., all sound sources and

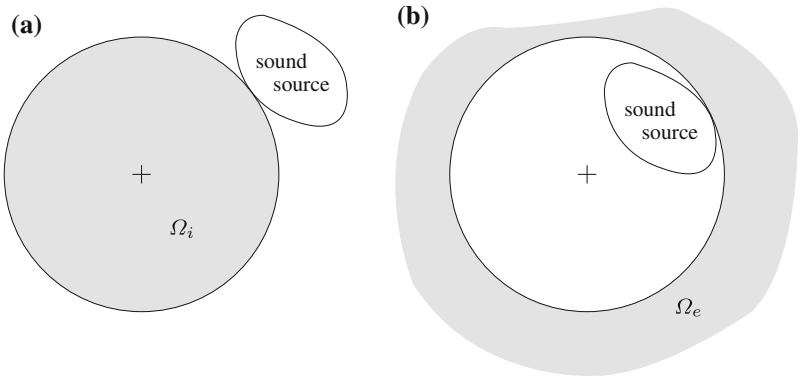

Fig. 2.5 Examples of interior and exterior problems. Shaded areas denote the domains of interest. The cross indicates the origin of the coordinate system. **a** Interior domain Ω_i. **b** Exterior domain Ω_e

obstacles are located outside the considered domain. Exterior problems on the other hand consider domains that are exterior to a distribution of sound sources and obstacles. Exterior problems do not necessarily extend to infinity. They can thus as well be interior with respect to a second sound source distribution. In the latter case, this interjacent problem is then described as a superposition of an interior and an exterior problem.

When considering series of surface spherical harmonics the boundaries to interior and exterior problems are spherical and are centered around the origin of the coordinate system employed. The boundary of an interior domain is thus a sphere centered around the origin of the coordinate system that is tangent to the closest sound source of a source distribution and that does not cut through the source distribution at any point. The precise definition of the exterior domain is accordingly. Refer to Fig. 2.5 for an illustration.

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

$$S_i(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \breve{S}_{n,i}^{m}(\omega)\, j_n\left(\frac{\omega}{c} r\right) Y_n^m(\beta, \alpha), \tag{2.32a}$$

and in the exterior domain Ω_e by

$$S_e(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \breve{S}_{n,e}^{m}(\omega) h_n^{(2)}\left(\frac{\omega}{c} r\right) Y_n^m(\beta, \alpha). \tag{2.32b}$$

Equations (2.32a) and (2.32b) are also termed interior and exterior expansions respectively.

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

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

$$\breve{S}^m_{n,i}(\omega) = \frac{1}{j_n\left(\frac{\omega}{c}r\right)} \int_0^{2\pi} \int_0^{\pi} S(\mathbf{x}, \omega)\, Y_n^{-m}(\beta, \alpha) \sin\beta \, d\beta \, d\alpha \tag{2.33}$$

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

Since expansions (2.32) converge uniquely and uniformly above a certain threshold, the order of summation may be exchanged (Gumerov and Duraiswami 2004, p. 75). If the spherical harmonics $Y_n^m(\beta, \alpha)$ are then expressed by their explicit formulation (2.23), the Fourier series that is inherent to (2.32) is revealed. It is given by

$$S(\mathbf{x}, \omega) = \sum_{m=-\infty}^{\infty} \underbrace{\sum_{n=|m|}^{\infty} \breve{S}_n^m(\omega)\, j_n\left(\frac{\omega}{c}r\right) (-1)^m \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos\beta)}_{= \mathring{S}_m(r,\beta,\omega)} e^{im\alpha}. \tag{2.34}$$

exemplarily for the interior expansion. The Fourier series expansion coefficients of $S(\mathbf{x}, \omega)$ are denoted by $\mathring{S}_m(r, \beta, \omega)$. Note that the basis functions $e^{im\alpha}$ are also termed *circular harmonics*.

As mentioned in Sect. 2.1.3, the basis functions $e^{im\alpha}$ of the Fourier series are orthogonal for $m \in \mathbb{Z}$. Furthermore, they constitute a complete set and the orthogonality relation (Williams 1999)

$$\frac{1}{2\pi} \sum_{m=-\infty}^{\infty} e^{im\alpha} e^{-im\alpha'} = \delta(\alpha - \alpha') \tag{2.35}$$

holds. The inverse operation to (2.34) is given by

$$\mathring{S}_m(r, \beta, \omega) = \frac{1}{2\pi} \int_0^{2\pi} S(\mathbf{x}, \omega) e^{-im\alpha} d\alpha. \tag{2.36}$$

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

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

$$\forall r < r_s \tag{2.37a}$$

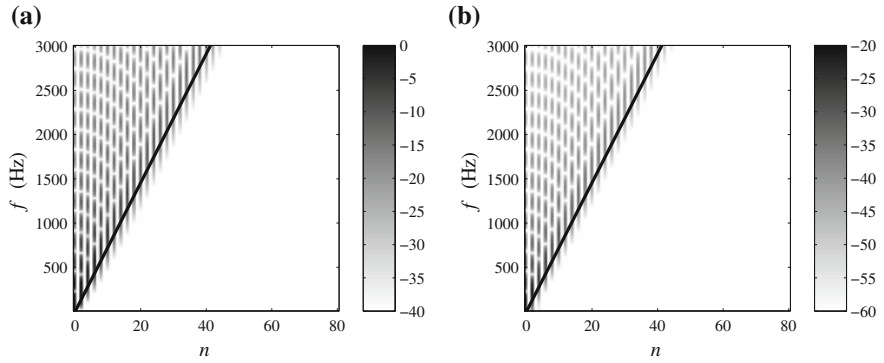

Fig. 2.6 Coefficients $20\log_{10}\left|\breve{S}_n^0(\omega)\, j_n\,(\omega/c\; r)\right|$, for $r = r_s/2\text{m}$; The black line indicates $f = (nc)/(2\pi r)$, the boundary between the primarily propagating region ($f > (nc)(2\pi r)$) and the primarily evanescent region ($f < (nc)(2\pi r)$). **a** Plane wave. **b** Point source; $r_s = 1.5\text{m}$

$$\frac{e^{-i\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_\text{s}|}}{|\mathbf{x}-\mathbf{x}_\text{s}|} = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \underbrace{(-i)\frac{\omega}{c} j_n\left(\frac{\omega}{c}r_s\right) Y_n^{-m}(\beta_s,\alpha_s)}_{=\breve{S}^m_{n,\text{sw,e}}} h_n^{(2)}\left(\frac{\omega}{c}r\right) Y_n^m(\beta,\alpha) \qquad \forall r > r_s \tag{2.37b}$$

for a spherical wave originating from $(r_\text{s}, \alpha_\text{s}, \beta_\text{s})$ and

$$e^{-i\mathbf{k}_\text{pw}^T\mathbf{x}} = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \underbrace{4\pi i^{-n} Y_n^{-m}(\phi_\text{pw},\theta_\text{pw})}_{=\breve{S}^m_{n,\text{pw}}} j_n\left(\frac{\omega}{c}r\right) Y_n^m(\beta,\alpha) \tag{2.38}$$

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

Approaching $r = r_\text{s}$ in (2.37a) and (2.37b) from the valid region of r shows that (2.37a) and (2.37b) are equal for $r = r_\text{s}$.

As derived in (Marathay and Rock 1980), when $(\omega/c)\; r < n$ or $f < (nc)/(2\pi r)$, i.e., when the argument of the spherical Bessel functions is smaller than the order, then the sound wave is primarily propagating; otherwise it is primarily evanescent. Figure 2.6 illustrates the amplitude distribution of the coefficients $\breve{S}_n^0(\omega)\, j_n\,(\omega/c\; r)$ for the interior expansion of a plane wave and a spherical wave. The according coefficients for $m \neq 0$ are qualitatively similar.

Occasionally in this book, a given sound field will be considered with respect to two different coordinate systems. The spherical harmonics expansions of the given sound field with respect to the two coordinate systems are related by a translation

operation. This translation of coordinate systems is not straightforward. Appendix E.1 summarizes one compact representation thereof. Selected alternative representations are outlined in Sects. 3.3.3 and 3.5.3. An extensive treatment can be found in (Gumerov and Duraiswami 2004). A selected rotation of the underlaying coordinate system is outlined in Appendix E.2.

2.2.2 Selected Properties of Bandlimited Spherical Harmonics Series

Consider a bandlimited series

$$S(\mathbf{x}, \omega) \approx \sum_{n=0}^{N-1} \sum_{m=-n}^{n} \mathring{S}_n^m(r, \omega) Y_n^m(\beta, \alpha). \tag{2.39}$$

Above a certain threshold $N_{\min}$, (2.39) converges uniformly for given r and ω (Kennedy et al. 2007; Gumerov and Duraiswami 2004) so that any such bandlimited series constitutes an approximation of $S(\mathbf{x}, \omega)$ the error of which decreases with increasing $N > N_{\min}$. In the case of (2.39), i.e., $\mathring{S}_n^m(r, \omega) = 0 \,\forall\, n > N - 1$, one speaks of an *N-truncated sum* (Gumerov and Duraiswami 2004, p. 75), an expansion with spatial bandwidth $N - 1$, or an $(N - 1)$-th order expansion. When simulations are presented in this book that depict quantities of infinite order, the order of the simulations is chosen such that the result becomes indistinguishable from the exact representation. Note that an $(N - 1)$-th order expansion is described by N^2 coefficients $S_n^m(r, \omega)$.

A thorough analysis of accuracy and properties of bandlimited expansions like (2.39) is cumbersome since the properties strongly depend on a number of factors including the propagation direction of the sound field $S(\mathbf{x}, \omega)$ under consideration in the domain of interest. The reader is referred to (Gumerov and Duraiswami 2004, Chap. 9) for an extensive mathematical treatment. An explicit review of this treatment is waived here since the perceptual consequences of such a spatial bandwidth limitation can not be deduced from mathematical treatments.

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

2.2.2.1 Interior Expansions

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

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

Typically, the domain inside which a bandlimited sound field description is considered to be comparable to its full-band analog is assumed to be inside the region where the argument $(\omega/c)\, r$ of the spherical Bessel function $j_n(\cdot)$ is smaller than the highest order $(N-1)$ contained in the expansion (Gumerov and Duraiswami 2004, p. 427). The radius r_{N-1} at which

$$(N-1) = \frac{\omega}{c} r_{\mathrm{N}-1} \tag{2.40}$$

represents the spatial boundary of this region. In the remainder of this book, the domain bounded by a sphere of radius r_{N-1} will be referred to as r_{N-1}-*region*. Note that r_{N-1} is indirectly proportional to the time frequency f. Expressing (2.40) in terms of the wavelength λ introduced in (2.4) yields

$$r_{\mathrm{N}-1} = \frac{N-1}{2\pi}\lambda. \tag{2.41}$$

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

Figure 2.7 depicts a monochromatic plane wave with propagation direction $(\pi/2, \pi/2)$, a 25th-order approximation of the plane wave, a 12th-order approximation of the plane wave, and the magnitude of the latter. The circles bound the corresponding r_{N-1}-region. It can be seen that the approximation with larger bandwidth describes the original sound field over a larger volume. As apparent especially in Fig. 2.7d, outside the r_{N-1}-region the amplitude of the bandlimited approximation can be higher than that of the exact representation. This circumstance constitutes a *Gibbs phenomenon* (Weisstein 2002).

This overshoot of the sound pressure can be significantly reduced by avoiding a hard truncation of the order by applying an angular fade-out (an angular *window*) towards higher orders as

$$S(\mathbf{x}, \omega) = \sum_{n=0}^{N-1} \sum_{m=-n}^{n} \breve{w}_n \breve{S}_n^m\, j_n\left(\frac{\omega}{c} r\right) Y_n^m(\beta, \alpha) \tag{2.42}$$

This procedure is also termed *angular weighting* (Ahrens and Spors 2009). Figure 2.8b illustrates the consequences of the cosine-shaped angular window shown in Fig. 2.8a when applied to the plane wave example from Figs. 2.7c, d. Note that although indicated for reference in Fig. 2.8b, the r_{12}-region as in Figs. 2.7c, d is not valid. Other types of angular windows may also be applied all of which have specific properties (Harris 1978).

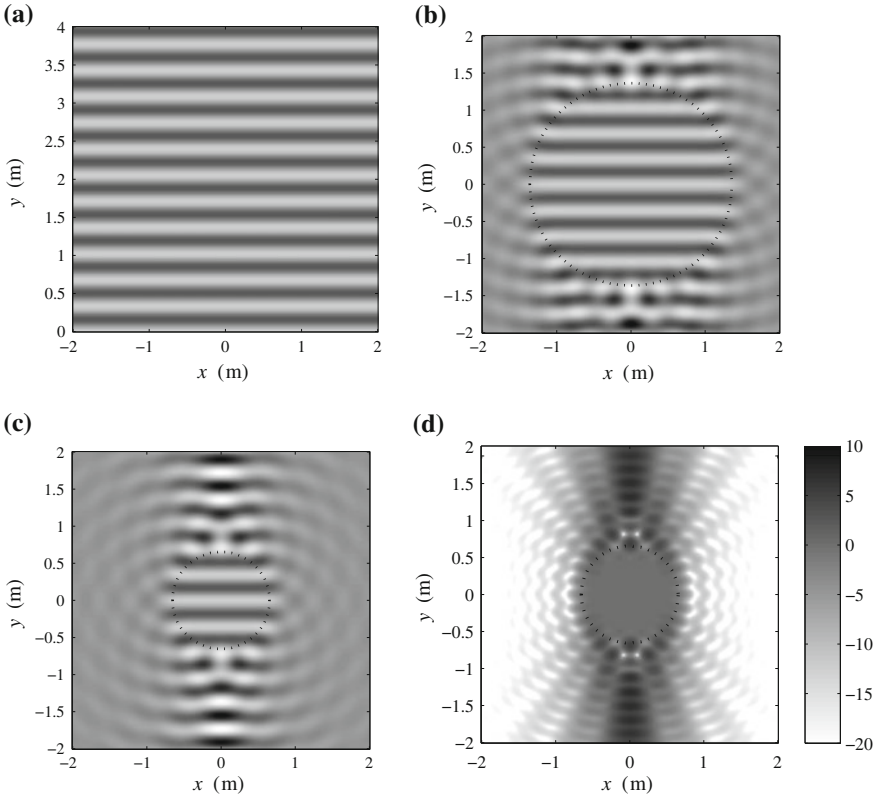

Fig. 2.7 Cross-section through the horizontal plane of a monochromatic plane wave sound field $S(\mathbf{x}, \omega)$ with propagation direction $(\pi/2, \pi/2)$ of frequency $f = 1000\,\text{Hz}$ (Fig. 2.7a) and bandlimited approximations thereof with different bandwidths (Figs. 2.7 b–d). The dotted circles bound the $r_{\text{N}-1}$-region. **a** $\Re\{S(\mathbf{x}, \omega)\},\ N = \infty$. **b** $\Re\{S(\mathbf{x}, \omega)\},\ N = 26$. **c** $\Re\{S(\mathbf{x}, \omega)\},\ N = 13$. **d** $20 \log_{10} |S(\mathbf{x}, \omega)|,\ N = 13$. Values are clipped as indicated by the colorbar

From Figs. 2.7c, d it is evident that the approximation can exhibit very low amplitude in those locations that are outside of the r_{N-1}-region and that are not along the channel of propagation of the sound field which crosses the r_{N-1}-region.

Consider now a sound field carrying a signal that is broadband with respect to the time frequency. When the spatial bandwidth of the sound field is constant over the entire time-frequency range, then the sound field has a larger "extent" at low frequencies than at high frequencies. At positions closer to the expansion center more energy is apparent at higher time frequencies than at farther positions. This circumstance is illustrated in Fig. 2.9, which shows the amplitude of a 5-th order plane wave $(N = 6)$, which propagates inside the horizontal plane. Figure 2.9 is essentially a broadband extension of Fig. 2.7d. It can be seen that the spatial extent of the sound field under consideration can shrink to only a few centimeters at some kHz.

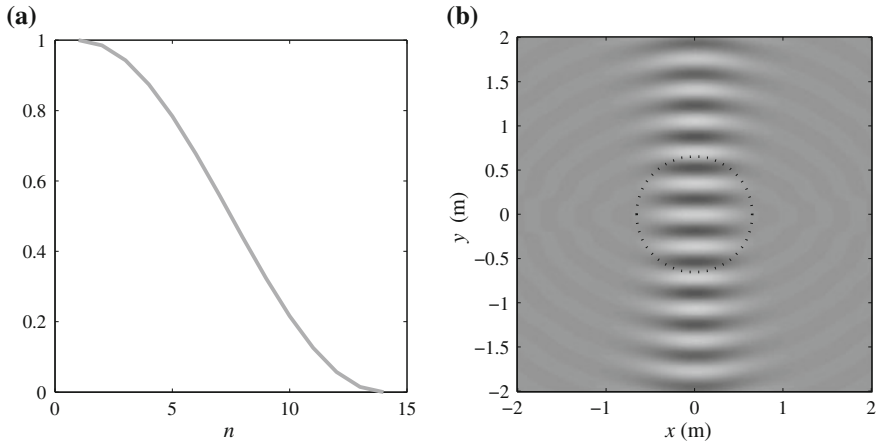

Fig. 2.8 Angular weighting for reduction of the Gibbs phenomenon apparent in Figs. 2.7 c, d. **a** Cosine-shaped angular window $\breve{w}_n$ applied to the expansion in Fig. 2.8b. **b** Cross-section through the horizontal plane of a 12-th order approximation of a monochromatic plane wave sound field $S(\mathbf{x}, \omega)$ with propagation direction $(\pi/2, \pi/2)$ of frequency $f = 1000\,\text{Hz}$ with angular weighting as shown in Fig. 2.8a. The dotted circle bounds the r_{12}-region

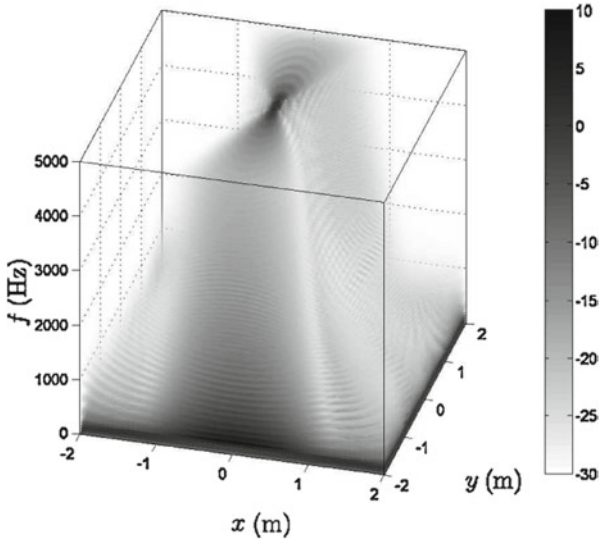

Fig. 2.9 $20\log_{10}|S(\mathrm{x}, \omega)|$ of a 5-th order plane wave $(N = 6)$ with propagation direction $(\pi/2, \pi/2)$; a cross-section through the horizontal plane is shown. The magnitude is indicated both via brightness as well as via transparency. Values below the lower limit indicated by the errorbar are fully transparent; opacity increases proportionally to the magnitude and reaches full opacity for values above the upper limit indicated by the errorbar

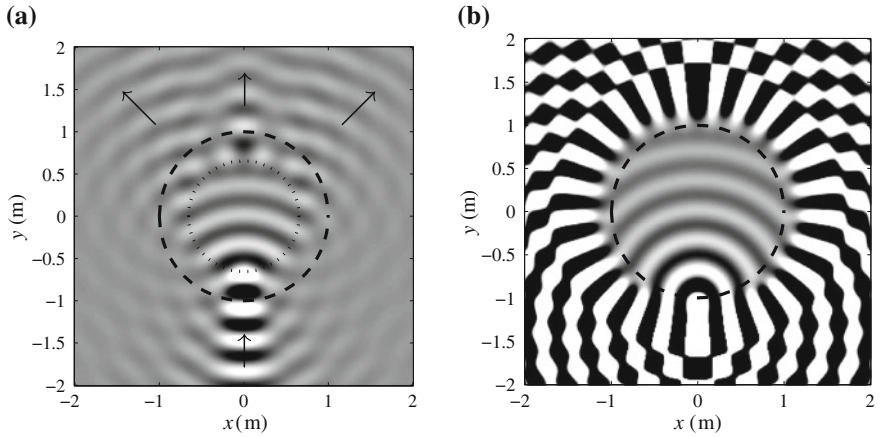

Fig. 2.10 Bandlimited interior expansion of the monopole source located at $r_\mathrm{s} = 1\,\mathrm{m}$; $(\alpha_\mathrm{s}, \beta_\mathrm{s}) = (-\pi/2, \pi/2)$ emitting a monochromatic signal of $f = 1000\,\mathrm{Hz}$. A cross-section through the horizontal plane is shown. **a** $N = 13$; the dotted line bounds the r_{12}-region; the dashed line bounds the domain of validity of the interior expansion; the arrows indicate the local propagation direction. **b** $N = 26$; the dashed line bounds the domain of validity of the interior expansion; the r_{25}-region is larger than the domain of validity and is therefore not indicated

Using the assumption of an $r_{\mathrm{N}-1}$-region, it can be shown than a 16-th order is required in order for an plane wave to be accurately described over a volume of the size of human head at $f = 10\,\mathrm{kHz}$ (assuming a head radius of 8.5 cm). An $(N-1)$-th order sound field is described by N^2 coefficients $\breve{S}_n^m(\omega)$, which are 136 coefficients in the present case.

The situation can be different for other types of sound fields. Spherical waves the origin of which is far away from the expansion center behave similarly as shown in Fig. 2.9 whereas the concentration of energy is less pronounced for spherical waves the origin of which is closer to the expansion center.

Figure 2.10a shows an example of a spherical wave where the r_{N-1}-region is smaller than the domain of validity of the interior expansion. Recall the the domain of validity is determined by the distance of the monopole source (that evokes the spherical wave) to the center of the expansion. A remarkable property of the sound field depicted in Fig. 2.10a is its local propagation direction, which is indicated by the arrows.

The situation changes when the bandlimit N is chosen such that the r_{N-1}-region is larger than the domain of validity of the interior expansion. Refer to Fig. 2.10b for an illustration. In the domain of validity, the sound field is accurately described. Outside of this domain, the sound field can not be interpreted. Sound fields like the one depicted in Fig. 2.10b are further analyzed and manipulated in Sect. 5.6.2.

Another important aspect are the time-domain properties of spatially bandlimited sound fields. Figure 2.11 depicts a spatially bandlimited plane wave that carries a time-domain impulse and is thus broadband with respect to time frequency.

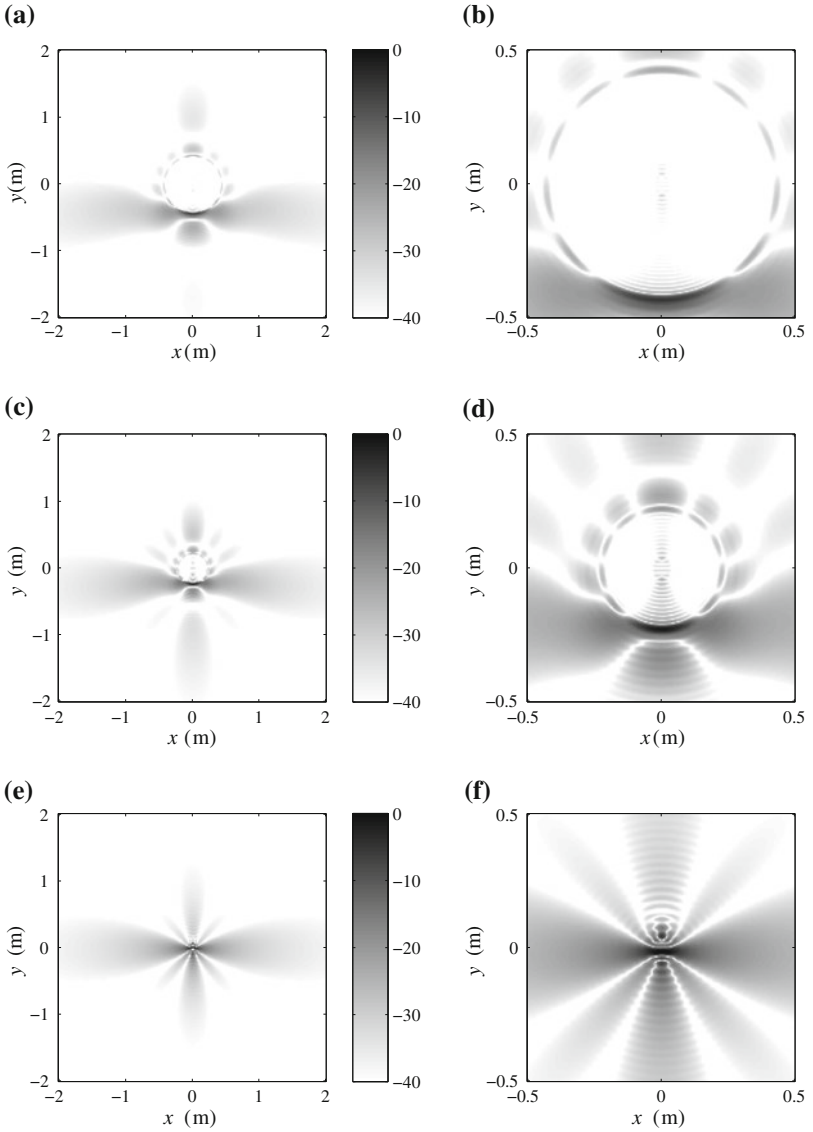

Fig. 2.11 $20\log_{10}|s(\mathbf{x},t)|$ of a 5-th order plane wave ($N=6$) with propagation direction $(\pi/2,\pi/2)$, which carries a time-domain impulse; a cross-section through the horizontal plane is shown on different scales (*left* column: large scale, *right* column: small scale). **a** large scale, $t=-1.8$ ms. **b** small scale, $t=-1.8$ ms. **c** large scale, $t=-0.9$ ms. **d** small scale, $t=-0.9$ ms. **e** large scale, $t=0$ ms. **f** small scale, $t=0$ ms

The absolute value of the sound pressure is shown in dB, i.e.,

$$20\log_{10}|\Re\{s_{\mathrm{S}}(\mathbf{x},t\}|\,.$$

The simulation was obtained via a numerical inverse Fourier transform of (2.32a). Close to center, the wave fronts are indeed as desired, whereby the accuracy of the bandlimited sound field description rises for source for positions closer to the expansion center. A distances far from the expansion center, the wave fronts tend to be smeared with respect to time, which is a consequence of the lack of energy at high frequencies at these locations. Note that an r_{N-1}-region (Sect. 2.2.2.1) can not be indicated here since it is frequency dependent.

Inspecting Fig. 2.11 thoroughly reveals that not all energy is propagating in the same direction like the plane wave. E.g., in Fig. 2.11b, a circular wave front is apparent, which propagates towards the origin of the coordinate system (i.e., towards the expansion center). This converging circular wave front is apparent at any location though with decreasing amplitude with respect to the distance from the center. Increasing the spatial bandwidth or applying angular weighting as in Fig. 2.8b also reduces the amplitude.

Obviously, for locations close to the center the converging wave front arrives only a few microseconds earlier than the plane wave. This interval increases for farther locations. For the situation depicted in Fig. 2.11, the amplitude of the additional wave front is around 20 dB below the maximum amplitude of the desired plane wave.

Note finally that the spatial structure of the plane wave depicted in Fig. 2.11 is essentially symmetric with respect to time, i.e., once the plane wave front has passed the origin, it is followed by a circular diverging wave.

2.2.2.2 Exterior Expansions

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

An elementary type of sound source is a point source (or *monopole*) the spatial transfer function of which constitutes a spherical wave and is given by (Williams 1999; Gumerov and Duraiswami 2004)

$$\frac{e^{-i\frac{\omega}{c}r}}{r} = -i\frac{\omega}{c}h_0^{(2)}\left(\frac{\omega}{c}r\right) = -\sqrt{4\pi}i\frac{\omega}{c}h_0^{(2)}\left(\frac{\omega}{c}r\right)Y_0^0(\beta,\alpha), \tag{2.43}$$

and thus employs only 0-th order.

For illustration of the properties of sound sources with more complex radiation properties consider a sound source whose spatial transfer function is given by (Ahrens and Spors 2010)

$$S_{n,e}^m(\omega) = \begin{cases} (-1)^{(m+n)}i^{-n}\frac{(N-1)!N!}{(N+n)!(N-n-1)!}Y_n^{-m}(\beta_{\text{or}},\alpha_{\text{or}}) & \forall\, n \le N-1 \\ 0 & \text{elsewhere} \end{cases}. \tag{2.44}$$

$(\alpha_{\text{or}}, \beta_{\text{or}})$ denotes the main radiation direction of the source, i.e., its nominal orientation. Equation (2.44) represents a purely real spatial transfer function.

Refer to Fig. 2.12, which depicts the sound field radiated by sources whose spatial transfer functions are given by (2.44) for $(\alpha_{\text{or}}, \beta_{\text{or}}) = (0, \pi/2)$ and $N = 4$ and $N = 21$

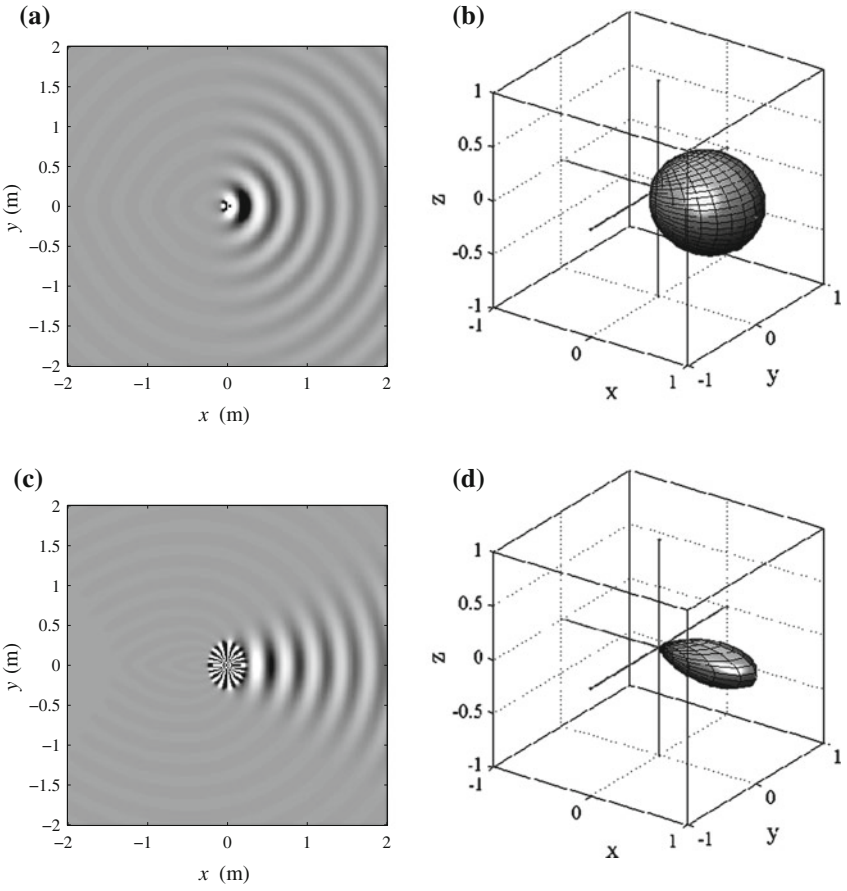

Fig. 2.12 Sound fields the horizontal plane and far-field signature functions (see Sect. 2.2.5) of monochromatic sound sources with a spatial transfer function given by (2.44). **a** $N=4, f=1000$ Hz. **b** Normalized far-field signature function, $N=4$. **c** $N=21, f=1000$ Hz. **d** Normalized far-field signature function, $N=21$

respectively. The far-field directivities of the two sound sources are also depicted (refer to Sect. 2.2.5 for a treatment of far-field radiation). It can be seen especially in Fig. 2.12c that the emitted sound field exhibits very high values in the vicinity of the sound source. This circumstance is also represented in Fig. 2.2c by the fact that the higher the order n of a spherical Hankel function the larger is its magnitude especially for low arguments.

The high pressure values apparent in Figs. 2.12a, c are caused by the evanescent components of the sound field. Note that the sound pressure in Fig. 2.12c clips over a larger area than in Fig. 2.12a. Considerable evanescent field components indicate the vicinity of vibrating surface (i.e a sound source) (Williams 1999). Larger bandwidths

thus suggest a larger spatial extent of a source. Note however that this is not a general rule.

Finally, it can be seen from Fig. 2.12 that the directivity of the source with bandwidth $N = 21$ exhibits a stronger focus in the main radiation direction. A strong frequency dependency like with the properties of interior expansion treated above is not present here.

2.2.3 Multipoles

Radiating solutions to the Helmholtz equation can be represented by *multipole expansions*, i.e., by combinations of monopoles located at infinitesimal distance from each other (Gumerov and Duraiswami 2004, p. 71). Lower order multipoles are also referred to as *monopoles*, *dipoles*, *quadrupoles*, *octopoles*, etc.

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

2.2.4 The Signature Function

Interior sound fields $S_i(\mathbf{x}, \omega)$ can be represented by a continuum of propagating plane waves with respect to the surface of a notional unit sphere as (Colton and Kress 1998; Gumerov and Duraiswami 2004)

$$S_i(\mathbf{x}, \omega) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} \bar{S}_i(\phi, \theta, \omega) e^{-i\mathbf{k}^T\mathbf{x}} \sin\phi \, d\phi \, d\theta. \tag{2.45}$$

The coefficients $\bar{S}_i(\phi, \theta, \omega)$ of the decomposition are termed *signature function* (Gumerov and Duraiswami 2004, p. 82).

Note that, although $S(\mathbf{x}, \omega)$ is represented by a continuum of *propagating* plane waves, it is an exact representation of $S(\mathbf{x}, \omega)$, i.e., it covers evanescent components. For completeness, various relations between the signature function $\bar{S}_i(\phi, \theta, \omega)$ and other representations of $S_i(\mathbf{x}, \omega)$ are derived in Appendix E.4. The most important relation in the context of this book is

$$\bar{S}_i(\phi, \theta, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n \breve{S}_{n,i}^m(\omega) Y_n^m(\phi, \theta). \tag{2.46}$$

2.2.5 Far-Field Radiation

As outlined in Sect. 2.2.1, the spatial transfer function of any stationary sound source of finite spatial extent can be represented in the exterior domain by a series of spherical harmonics $Y_n^m(\beta, \alpha)$ and appropriate coefficients as stated by (2.32b).

When (2.32b) is evaluated in the far-field, i.e., for $(\omega/c)\, r \to +\infty$, then the large-argument approximation of the spherical Hankel functions (2.19) (Williams 1999) can be applied resulting in (Gumerov and Duraiswami 2004, Eq. (2.3.39), p. 81)

$$\begin{aligned} S_e(\mathbf{x}, \omega) &\approx \frac{i}{\frac{\omega}{c}} \frac{e^{-i\frac{\omega}{c}r}}{r} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n \, \breve{S}_{n,e}^m(\omega)\, Y_n^m(\beta, \alpha) \\ &= h_0^{(2)}\left(\frac{\omega}{c} r\right) \underbrace{\sum_{n=0}^{\infty} \sum_{m=-n}^{n} i^n \, \breve{S}_{n,e}^m(\omega)\, Y_n^m(\beta, \alpha)}_{= \bar{S}_e(\beta,\alpha,\omega)} . \end{aligned} \tag{2.47}$$

Thus, at sufficient distance any stationary sound source of finite spatial extent radiates like a point source (i.e., $\sim \frac{1}{r} \exp(-i(\omega/c)\ r)$, see (2.43) whereby the angular dependency of the transfer function is given by the *far-field signature function* $\bar{S}_e(\beta, \alpha, \omega)$ (Gumerov and Duraiswami 2004, p. 296). The latter is given by an appropriate summation of the coefficients $\breve{S}_{n,e}^m(\omega)$ respectively. Note the similarity between $\bar{S}_e(\beta, \alpha, \omega)$ and $\bar{S}_i(\phi, \theta, \omega)$ given by (2.46).

The distance that is sufficient in order for (2.47) to be valid is reached when the distance from the observation point to the sound source is much larger than the spatial extent of the sound source. For small sources like the human voice, the region of validity is reached at low distances. For extended sources like a car or a train, the required distance is significantly larger.

Note that it is actually not rigorous to apply the large-argument approximation on (2.32b) since the former does not hold uniformly in n. Rigorous treatments can be found in (Colton and Kress 1998; Gumerov and Duraiswami 2004), which also lead to (2.47). The detailed derivation of (2.47) is not performed here since it is not relevant for the remainder of this book.

Finally, (2.47) proves that the signature function corresponds to what is commonly referred to as *directivity* or *directivity function* (Williams 1999, p. 39; Blackstock 2000, p. 463) and constitutes the two-dimensional equivalent of polar diagrams. Examples of far-field signature functions are depicted in Fig. 2.12.

The fact that the far-field representation (2.47) avoids the necessity of evaluating spherical Hankel functions significantly reduces the computational complexity. Additionally, one can benefit from all advantages that the two representations comprise since they can be used interchangeably. The discrete property of the coefficients $\breve{S}_{n,e}^m(\omega)$ makes this representation suitable for storage and transmission of the radiation properties for a source under consideration (Wefers 2008). The more

intuitive representation of $\bar{S}_e(\beta, \alpha, \omega)$ is helpful in the modeling of desired radiation properties.

2.2.6 The Wavenumber Domain

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

$$\tilde{S}(k_x, y, z, \omega) = \int_{-\infty}^{\infty} S(\mathbf{x}, \omega) e^{ik_x x} dx \tag{2.48}$$

exemplarily for the x-dimension. The inverse operation to (2.48) is given by (B.4) in Appendix B. The spatial Fourier domain is also referred to as *wavenumber domain* or *k-space* (Williams 1999).

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

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

Recalling the dispersion relation (2.8)–(2.10) from Sect. 2.1.2, a sharp segregation of propagating and evanescent components of the sound field under consideration is straightforward. The region $\omega/c < k_{x/y/z}$ is purely evanescent; the region $\omega/c \geq k_{x/y/z}$ is purely propagating (Williams 1999, p. 30). Figure 2.13 illustrates this circumstance on the example of the k_x-f-spectrum of a monopole source residing in the coordinate origin given by (C.10). Obviously, the evanescent components are more prominent for closer observation points.

Another convenient property of the wavenumber domain is the fact that the propagation direction of the described sound field can be directly deduced (Williams 1999, Sect. 2.8). As apparent from (C.5), which is stated here again for convenience,

$$\tilde{S}(k_x, y, z, \omega) = 2\pi\delta(k_x - k_{\mathrm{pw},x}) e^{-ik_{\mathrm{pw},y} y} e^{-ik_{\mathrm{pw},z} z} \cdot 2\pi\delta(\omega - \omega_{\mathrm{pw}}), \tag{2.49}$$

a monochromatic plane wave is represented by a Dirac delta function in k_x-space. This is illustrated schematically in Fig.2.14a.

The triangular area between the two gray lines that indicate $\omega/c = |k_x|$ in Fig. 2.14 is where propagating components are located. Evanescent components are located

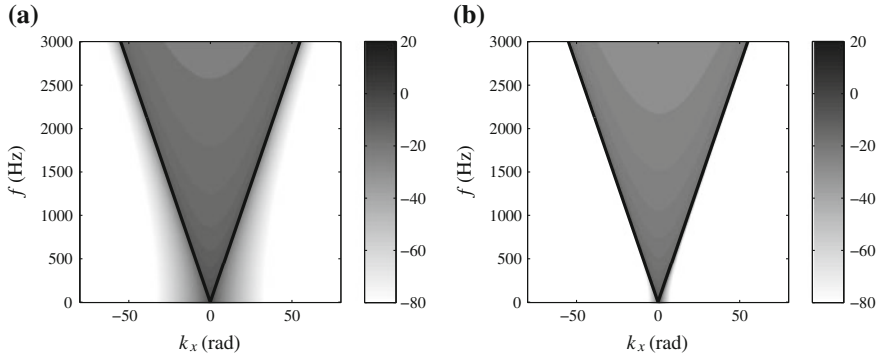

Fig. 2.13 $20\log_{10}\left|\tilde{S}(k_x, y, z, \omega)\right|$ of a monopole source residing in the coordinate origin for different observation points. The black lines indicate $\omega/c = |k_x|$. **a** $y = 0.2\,\mathrm{m}$; $z = 0\,\mathrm{m}$. **b** $y = 1\,\mathrm{m}$; $z = 0\,\mathrm{m}$

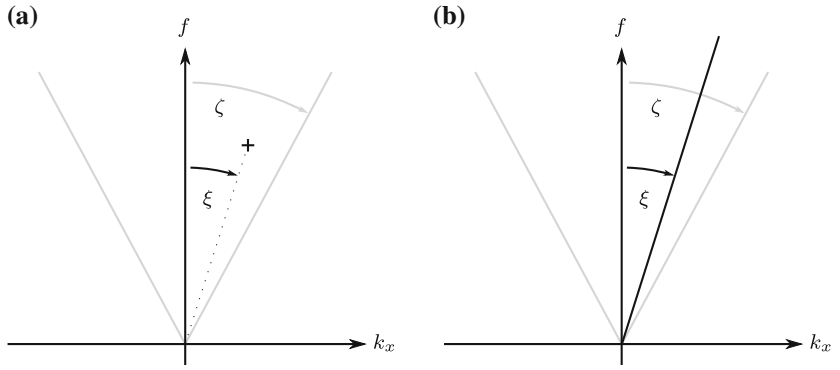

Fig. 2.14 Schematic of the k_x-f-spectrum of a plane wave with propagation direction $0 < \theta_{\mathrm{pw}} < \pi/2$. The gray lines indicate $\omega/c = |k_x|$. **a** Monochromatic plane wave; the mark indicates the location of the energy. **b** Broadband plane wave; the black line indicates the location of the energy

outside of this area. This can be deduced from the dispersion relation (2.8). Note that when a sound field is considered in wavenumber domain with respect to two dimensions, the this triangular area becomes the inner of a cone.

The angle ξ between the f-axis ($k_x = 0$) and the straight line through the origin and the Dirac delta function determines the propagation direction of the plane wave under consideration. Assuming a horizontal propagation direction of the plane wave, i.e., $\left(\phi_{\mathrm{pw}} = \pi/2\right)$, the described plane wave propagates approximately in direction of the y-axis for $\xi \approx 0$, i.e., $\theta_{\mathrm{pw}} \approx \pi/2$. For $\xi > 0$, then $0 < \theta_{\mathrm{pw}} < \pi/2$; and for $\xi < 0$, then $\pi/2 < \theta_{\mathrm{pw}} < \pi$. The relation between the propagation direction of the described plane wave represented by θ_{pw} can be deduced from the according ξ using trigonometric considerations using (A.3). If $|\xi|$ is larger than the absolute value $|\zeta|$ of the angle between the f-axis and the straight line along $\omega/c = k_x$, then an

evanescent wave is apparent. If $|\xi|$ is slightly smaller than $|\zeta|$, then the plane wave propagates approximately parallel to the x-axis.

A plane wave that is broadband with respect to time frequency is represented by a straight line through the origin at given angle ξ to the f-axis, as illustrated in Fig. 2.14b.

Considering the discussion above, the fact that a monopole source radiates in all directions is indeed represented in Fig. 2.13. Close to the source, the evanescent components apparent in the k_x-f-representation are stronger (Fig. 2.13a) compared to far distances (Fig. 2.13).

Note that the forward and inverse spatial Fourier transforms as used in this book (2.48) and (B.4) use signs in the exponent that are reversed with respect to the forward and inverse temporal Fourier transforms defined in (B.1) and (B.2) respectively. The motivation to do so is related to the propagation direction of plane waves as explained in the following.

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

$$S(\mathbf{x}, \omega) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \tilde{S}(\mathbf{k}, \omega) e^{-i\mathbf{k}^T\mathbf{x}} dk_x dk_y dk_z. \tag{2.50}$$

The exponential function in (2.50) can be interpreted as a plane wave propagating in direction $\mathbf{k}$ (refer to (2.7) and Appendix C.1). Thus, the spatial Fourier domain constitutes a plane wave representation of a sound field with respect to a three-dimensional space. The wave vector $\mathbf{k} = k \cdot [\cos\theta \sin\phi \;\; \sin\theta \sin\phi \;\; \cos\phi]^T$ then points into the direction of propagation of the plane wave component under consideration. The propagation direction is also represented by the colatitude ϕ and the azimuth θ. Using signs in the exponent of the spatial Fourier transform similar to the temporal one as e.g., in (Rabenstein et al. 2006) results in the angles ϕ and θ describing the direction the plane wave is "coming from", which is considered less elegant.

In order to illustrate the physical meaning of the wavenumber k_i, the analogies of the spatial Fourier transform (2.48) and the temporal Fourier transform defined in (B.1) are outlined below exemplarily for the x and k_x-dimensions respectively. The frequency variable in the time Fourier transform is the *radian time frequency* ω, which is related to the *time frequency* f via $\omega = 2\pi f$. In practice, the time-frequency scale (not the radian frequency scale) is used in order to refer to specific values.

The frequency variable in the spatial Fourier transform is the *wavenumber in x-direction* k_x. k_x can thus be interpreted as the *spatial radian frequency* and is of unit rad/m. Via the relation $k_x = 2\pi f_x$, a *space frequency* f_x can be established. Note that $\lambda_x = (2\pi)/k_x = 1/f_x$ is termed *trace wavelength* in x direction and k_x is termed *trace wavenumber* in x direction (Williams 1999).

2.2.7 The Angular Spectrum Representation

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

$$S(\mathbf{x}, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \tilde{S}(k_x, y, k_z, \omega)\, e^{-i(k_x x + k_z z)} dk_x dk_z. \tag{2.51}$$

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

$$\frac{\partial^2}{\partial y^2} \tilde{S}(k_x, y, k_z, \omega) + k_y^2 \tilde{S}(k_x, y, k_z, \omega) = 0, \tag{2.52}$$

whereby

$$k_y = \sqrt{k^2 - k_x^2 - k_z^2}, \quad \forall\, k_x^2 + k_z^2 \leq k^2 \tag{2.53a}$$

$$k_y = i\sqrt{k_x^2 + k_z^2 - k^2}, \quad \forall\, k_x^2 + k_z^2 > k^2. \tag{2.53b}$$

A propagating sound field is described when (2.53a) is satisfied and an evanescent sound field is described when (2.53b) is satisfied.

There are two solutions to (2.52) which are given by

$$\tilde{S}_1(k_x, y, k_z, \omega) = \check{S}_1(k_x, k_z, \omega)\, e^{ik_y y} \tag{2.54a}$$

$$\tilde{S}_2(k_x, y, k_z, \omega) = \check{S}_2(k_x, k_z, \omega)\, e^{-ik_y y}. \tag{2.54b}$$

Introducing (2.54) into (2.51) yields two expressions for $S(\mathbf{x}, \omega)$ which are given by

$$S(\mathbf{x}, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \check{S}_1(k_x, k_z, \omega)\, e^{-i(k_x x + k_y y + k_z z)} dk_x dk_z \tag{2.55a}$$

$$S(\mathbf{x}, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \check{S}_2(k_x, k_z, \omega)\, e^{-i(k_x x - k_y y + k_z z)} dk_x dk_z. \tag{2.55b}$$

$\check{S}_1(k_x, k_z, \omega)$ and $\check{S}_2(k_x, k_z, \omega)$ are termed the *angular spectrum representation* or *plane wave spectrum* of $S(\mathbf{x}, \omega)$ in a source-free half-space (Nieto-Vesperinas 2006).

The integral (2.55a) is convergent for $y \geq 0$ and represents $S(\mathbf{x}, \omega)$ in the case that all sound sources are located at $y < 0$. Equation (2.55b) is convergent for $y \leq 0$ and represents $S(\mathbf{x}, \omega)$ in the case that all sound sources are located at $y > 0$.

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

Note that the signature function presented in Sect. 2.2.4 represents the decomposition of a sound field into plane wave with respect to the unit sphere.

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

Equation (2.55) takes the form of a two-dimensional inverse Fourier transform and can thus be inverted by the forward transform as indicated in Appendix B. Setting then $y = 0$ yields

$$\check{S}\,(k_x, k_z, \omega) = \iint_{-\infty}^{\infty} S(x, 0, z, \omega) e^{i(k_x x + k_z z)} dx dz, \tag{2.56}$$

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

Introducing (2.56) into (2.55a)yields

$$S(\mathbf{x}, \omega) = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} S(\mathbf{x}_0, \omega) \underbrace{\iint_{-\infty}^{\infty} e^{-i\left(k_x (x - x_0) + k_y y + k_z (z - z_0)\right)} dk_x dk_z}_{= \mathscr{P}(\mathbf{x} - \mathbf{x}_0, \omega)} dx_0 dz_0, \tag{2.57}$$

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

Equation (2.57) describes the relationship between the sound field $S(\mathbf{x}, \omega)$ at an arbitrary point $\mathbf{x}$ in the half-space $y \geq 0$ and its boundary value $S(x, 0, z, \omega)$ at the reference plane. Extensive literature exists regarding the theoretical limits on, applicability of, and analytical solutions to, the angular spectrum decomposition. Refer to the standard literature on *Fourier optics* such as (Nieto-Vesperinas 2006) for references.

2.2.8 Spatial Spectra and Spatial Bandlimitation

The term *spectrum*, in the present context, refers to the coefficients of a decomposition of a given quantity under consideration into given basis functions. Note that other meanings also exist (Weisstein 2002).

A very common spectrum is the *time-frequency spectrum* of a signal, which refers to the coefficients $S_0(\omega)$ of the decomposition of a signal $s_0(t)$ into sine or cosine waves respectively (Girod et al. 2001). The time-domain signal $s_0(t)$ can be synthesized from its spectrum $S_0(\omega)$ via an according transform, in this case the inverse Fourier transform (B.2). Note that the exponential in (B.2) represents cosine waves in complex notation. The spectrum $S_0(\omega)$ of the signal $s_0(t)$ can be obtained via the forward Fourier transform (B.1).

Another type of spectra considered in this book is the *space-frequency spectrum* or *spatial spectrum*, which refers to the coefficients of a decomposition of a quantity under consideration in elementary spatial basis functions. Thus, the spherical harmonics expansion coefficients $S_n^m(r, \omega)$ constitute one representation of the spatial spectrum of a sound field $S(\mathbf{x}, \omega)$, i.e., they represent a decomposition of $S(\mathbf{x}, \omega)$ into surface spherical harmonics. The corresponding forward transform is similar to (2.33) and the inverse transform is given by (2.31).

Also, $\bar{S}(\phi, \theta, \omega)$, $\tilde{S}(\mathbf{k}, \omega)$, and $\check{S}(k_x, k_z, \omega)$ constitute other representations of the spatial spectrum of $S(\mathbf{x}, \omega)$, in this case a decomposition of $S(\mathbf{x}, \omega)$ into different sets of plane waves. The according transforms for decomposition and recomposition are outlined in the respective sections.

From above considerations, the meaning of the term *spatial bandlimitation* becomes clear: Spatial bandlimitation constitutes a recomposition of a given signal using only a subset of the corresponding spectral coefficients and basis functions. One example of a spatial bandlimitation is given by (2.39). Bandlimitations for the other representations of the spatial spectrum can be obtained, e.g., by modifying the boundaries of the integral of the inverse transforms accordingly.

The reader might have an intuitive understanding of the effect of bandlimiting a time-domain signal in terms of the way the timbre of the signal changes. The effect on the time-domain representation of the signal, i.e., its wave form is less intuitive.

A spatial bandlimitation applied to a sound field affects the spatial structure of the latter. The way this spatial structure changes depends heavily on the fact with respect to which representation of the spatial spectrum the signal is bandlimited. A bandlimitation in terms of the spherical harmonics expansion as in (2.39) can concentrate the energy of the sound field under consideration in specific regions when an interior problem is considered as depicted in Fig. 2.7. With exterior problems, a spatial bandwidth limitation has a different effect.

A spatial bandwidth limitation with respect to one of the plane wave representations $\bar{S}(\phi, \theta, \omega)$, $\tilde{S}(\mathbf{k}, \omega)$, and $\check{S}(k_x, k_z, \omega)$ affects in diverse ways the properties of the sound field under consideration with respect to the way it propagates. Spatial bandwidth limitations will be essential to the discussion presented in Sect. 4.

2.3 Boundary Conditions

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

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

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

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

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

2.3.1 Dirichlet Boundary Condition

Dirchlet boundary conditions concern the sound pressure. The *homogeneous Dirchlet boundary condition* is given by

$$S(\mathbf{x}, \omega) = 0 \qquad \forall\, \mathbf{x} \in \partial\Omega \tag{2.58}$$

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

The *inhomogeneous Dirchlet boundary condition*

$$S(\mathbf{x}, \omega) = F_D(\mathbf{x}, \omega) \quad \forall\ \mathbf{x} \in \partial\Omega \tag{2.59}$$

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

2.3.2 Neumann Boundary Condition

The *homogeneous Neumann boundary condition* is given by

$$\left.\frac{\partial S(\mathbf{x}, \omega)}{\partial \mathbf{n}(\mathbf{x})}\right|_{\partial\Omega} = 0 \tag{2.60}$$

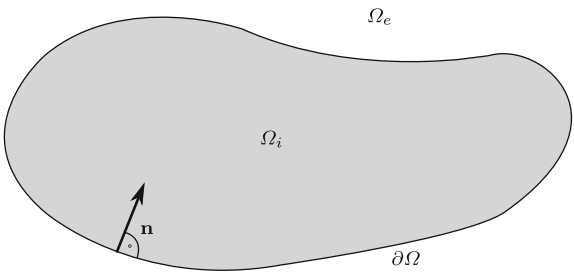

Fig. 2.15 Illustration of interior domain Ω_i which is enclosed by boundary $\partial\Omega$. Ω_e is the domain exterior with respect to $\partial\Omega$. $\mathbf{n}$ denotes the inward pointing surface normal on $\partial\Omega$

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

$$\frac{\partial}{\partial\mathbf{n}(\mathbf{x})}S(\mathbf{x},\omega) = \langle\nabla S(\mathbf{x},\omega),\mathbf{n}\rangle, \tag{2.61}$$

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

$$\langle\nabla,\mathbf{n}(\mathbf{x})\rangle = \cos\alpha_\mathrm{n}\sin\beta_\mathrm{n}\frac{\partial}{\partial x} + \sin\alpha_\mathrm{n}\sin\beta_\mathrm{n}\frac{\partial}{\partial y} + \cos\beta_\mathrm{n}\frac{\partial}{\partial z}. \tag{2.62}$$

Equation (2.60) states that the gradient of the sound pressure in direction of the normal $\mathbf{n}(\mathbf{x})$ on the boundary pointing into the domain of interest vanishes at the boundary $\partial\Omega$. Note that the directional gradient of a pressure field is directly proportional to the particle velocity (Williams 1999). A vanishing directional gradient of the sound pressure means also a vanishing particle velocity and thus a rigid boundary. Refer to Fig. 2.15 for an illustration of the interior example.

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

$$\left.\frac{\partial S(\mathbf{x},\omega)}{\partial\mathbf{n}(\mathbf{x})}\right|_{\partial\Omega} = F_N(\mathbf{x},\omega)\Big|_{\partial\Omega} \tag{2.63}$$

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

2.3.3 Sommerfeld Radiation Condition

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

$$\lim_{r\to+\infty} r\left(\frac{\partial}{\partial r}S(\mathbf{x},\omega)+i\frac{\omega}{c}S(\mathbf{x},\omega)\right)=0 \tag{2.64}$$

for the definitions of the Fourier transform used in this book. It is employed in exterior problems and provides a boundary condition at infinity. A sound field $S(\mathbf{x},\omega)$ satisfying (2.64) is composed of outgoing waves only. In simple words, the Sommerfeld radiation condition takes care that no energy contributions to the sound field under consideration stem from infinity.

2.4 Green's Functions

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

$$\nabla^2 G(\mathbf{x},\mathbf{x}_0,\omega)+k^2 G(\mathbf{x},\mathbf{x}_0,\omega)=-\delta(\mathbf{x}-\mathbf{x}_0) \tag{2.65}$$

are termed *Green's functions* (Williams 1999, p. 265). $\delta(\mathbf{x}-\mathbf{x}_0)$ denotes a three-dimensional Dirac delta function at position $\mathbf{x}_0$, which represents excitation of space at $\mathbf{x}_0$. Green's functions thus describe the response of the domain of interest to a spatial Dirac excitation and thus the way sound propagates. When considered in time domain (i.e., $g(\mathbf{x},\mathbf{x}_0,t)$), they can be interpreted as the spatial impulse response of the domain.

The free-field Green's function $G_0(\cdot)$ depends only on the distance between $\mathbf{x}$ and $\mathbf{x}_0$ and is stated here as (Williams 1999, Eq. (8.41), p. 265)

$$G_0(\mathbf{x}-\mathbf{x}_0,\omega)=\frac{1}{4\pi}\frac{e^{-i\frac{\omega}{c}|\mathbf{x}-\mathbf{x}_0|}}{|\mathbf{x}-\mathbf{x}_0|}. \tag{2.66}$$

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

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

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

$$\frac{\partial G_0(\mathbf{x},\omega)}{\partial x}=\frac{1}{4\pi}\left(i\frac{\omega}{c}-\frac{x}{r}\right)\frac{e^{-i\frac{\omega}{c}r}}{r^2}. \tag{2.67}$$

Equation (2.67) can be interpreted as the spatial transfer function of a dipole source whose main axis is along the x-axis (Williams 1999). The far-field signature function of $\partial G_0(\mathbf{x},\omega)/\partial x$ is similar to Fig. 2.4b.

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

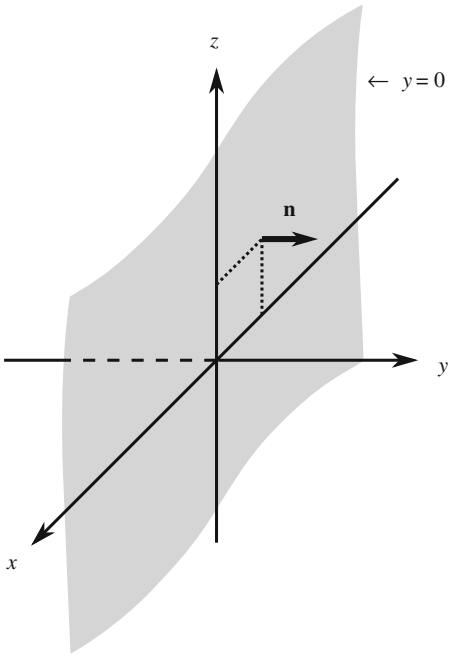

Fig. 2.16 Illustration of Rayleigh's first integral formula. For convenience it is assumed that the boundary $\partial\Omega$ of the target half-space is situated along the x-z-plane. It is indicated by the grey shading and has infinite extent. The target half-space contains the positive y-axis

2.5 The Rayleigh Integrals

The *Rayleigh I Integral*, also referred to as *Rayleigh's First Integral Formula*, may be formulated in time-frequency domain and under free-field conditions as (Williams 1999, Eq. (2.75), p. 36)

$$P(\mathbf{x}, \omega) = -\int_{\partial\Omega} 2 \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}, \omega)\Big|_{\mathbf{x}=\mathbf{x}_0} \cdot G(\mathbf{x} - \mathbf{x}_0, \omega)\, dA(\mathbf{x}_0), \tag{2.68}$$

$\mathbf{x}_0$ denotes a position on the plane $\partial\Omega$; $S(\mathbf{x}, \omega)$ denotes an arbitrary sound field that is source-free in one of the half-spaces bounded by $\partial\Omega$. The latter is referred to as *target half-space* in this book. Refer to Fig. 2.16 for an illustration. Due to the close relationship between (2.68) and the angular spectrum representation presented in Sect. 2.2.7, the properties of both representations with respect to convergence are similar (Nieto-Vesperinas 2006).

$\frac{\partial}{\partial \mathbf{n}}$ denotes the gradient in direction of $\mathbf{n}$, the unit length normal vector on the plane $\partial\Omega$ pointing into the target half-space. And finally, $P(\mathbf{x}, \omega)$ can be interpreted as the sound pressure evoked by a continuous monopole distribution that is located along $\partial\Omega$. $P(\mathbf{x}, \omega)$ is perfectly symmetric with respect to $\partial\Omega$ and is identical to $S(\mathbf{x}, \omega)$ for all positions inside the target half-space.

In words, the Rayleigh I Integral (2.68) states that the sound field $S(\mathbf{x}, \omega)$ inside a given source-free half-space (the target half-space) is uniquely determined by the

gradient of $S(\mathbf{x}, \omega)$ taken in direction of the normal on the boundary of the target half-space pointing into the target half-space and evaluated at that boundary.

Other similar integrals named after Rayleigh have been established (Williams 1999). They are termed *Rayleigh II*, *Rayleigh III*, etc. or *Rayleigh's Second*, *Third*, etc. Integral respectively.

2.6 The Kirchhoff-Helmholtz Integral

The *Kirchhoff-Helmholtz Integral* (or *Kirchhoff Integral* or *Helmholtz Integral*) is one of the essential theorems in acoustics. For interior problems it is given by (Williams 1999)

$$a(\mathbf{x})P(\mathbf{x},\omega) = -\oint_{\partial\Omega} \left(G(\mathbf{x},\mathbf{x}_0,\omega)\,\frac{\partial}{\partial \mathbf{n}(\mathbf{x}_0)} S(\mathbf{x},\omega)\Big|_{\mathbf{x}=\mathbf{x}_0} - S(\mathbf{x}_0,\omega)\,\frac{\partial}{\partial \mathbf{n}(\mathbf{x}_0)} G(\mathbf{x},\mathbf{x}_0,\omega) \right) dA(\mathbf{x}_0), \tag{2.69}$$

with

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

$\partial\Omega$ denotes a surface enclosing the source-free volume Ω_i, $A(\mathbf{x}_0)$ an infinitesimal surface element of $\partial\Omega$, $\mathbf{x}_0$ a point on $\partial\Omega$; Ω_e denotes the domain outside $\partial\Omega$, $G(\mathbf{x}, \mathbf{x}_0, \omega)$ a Green's function fulfilling the given boundary conditions, and $\partial/\partial\mathbf{n}(\mathbf{x}_0)$ the gradient in direction of the inward pointing surface normal $\mathbf{n}(\mathbf{x}_0)$. Refer to Fig. 2.15. An according formulation of (2.69) for exterior problems exists (Williams 1999).

The Kirchhoff-Helmholtz Integral (2.69) represents solutions to the homogeneous Helmholtz equation (2.2) with inhomogeneous boundary conditions. The sound field $P(\mathbf{x}, \omega)$ described by (2.69) equals $S(\mathbf{x}, \omega)$ $\forall \mathbf{x} \in \Omega_i$ provided that $S(\mathbf{x}, \omega)$ is source-free in Ω_i.

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

Under free-field conditions, i.e., when the boundary $\partial\Omega$ is *acoustically transparent*, then $G(\mathbf{x}, \mathbf{x}_0, \omega)$ is given by the free-field Green's function (2.66).

The Kirchhoff-Helmholtz Integral actually provides a direct formulation for sound field synthesis. As mentioned in Sect. 2.4, under free-field conditions the Green's function $G(\mathbf{x} - \mathbf{x}_0, \omega)$ employed in the Kirchhoff-Helmholtz Integral can be interpreted as the spatial transfer function of a monopole sound source and its directional gradient $\partial/(\partial\mathbf{n})G(\cdot)$ as the spatial transfer function of a dipole sound source whose main axis lies parallel to $\mathbf{n}$ (Williams 1999). Reinterpreted in terms of sound field synthesis, by means of an enclosing acoustically transparent continuous layer of *secondary monopole sources* and an according layer of *secondary dipole sources*, any source-free sound field can be synthesized inside this enclosing boundary.

However, this approach to sound field synthesis requires two layers of secondary sources, which is considered inconvenient. Typically, it is desired to avoid the dipole layer since it is very difficult to implement in practice. The fact that the sound field synthesized via the Kirchhoff-Helmholtz Integral is zero outside the secondary source distribution and thus that the acoustical properties of the listening room are negligible is only a theoretical benefit (Fazi and Nelson 2007) because the exterior sound field of practical implementations will not vanish as will be shown in Chap. 4.

It is thus rather desired to employ a monopole-only formulation. In Chap. 3, it will be shown that methods exist that may be employed in order to solve the problem of sound field synthesis and that avoid the necessity of secondary dipole sources.

References

Abramowitz, M., & Stegun, I.A. (eds) (1999). *Handbook of Mathematical Functions.* New York: Dover Publications Inc.

Ahrens, J., & Spors, S. (2009, June). Spatial encoding and decoding of focused virtual sound sources. In: *Ambisonics Symposium*.

Ahrens, J., & Spors, S. (2010, March). An analytical approach to 3D sound field reproduction employing spherical distributions of non- omnidirectional loudspeakers. *IEEE International Symposium, on Communications, Control and Signal Processing (ISCCSP)* (pp. 1–5).

Arfken, G., & Weber, H. (2005). *Mathematical Methods for Physicists.* San Diego: Elsevier Academic Press.

Blackstock, D. T. (2000). *Fundamentals of Physical Acoustics.* Wiley and Sons, Inc: New York.

Colton, D., & Kress, R. (1998). *Inverse Acoustic and Electromagnetic Scattering Theory.* Berlin: Springer.

Condon, E. U., & Shortley, G. H. (1935). *The Theory of Atomic Spectra.* Cambridge: Cambridge University Press.

Fazi, F., & Nelson, P. (2007). A theoretical study of sound field reconstruction techniques. In: *19th International Congress on Acoustics*. (Sept.).

Girod, B, Rabenstein, R, Stenger, A (2001). *Signals and Systems.* New York: Wiley.

Gumerov, N. A., & Duraiswami, R. (2004). *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions.* Amsterdam: Elsevier.

Harris, F. J. (1978). On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE, 66*, 51–83.

Jessel, M. (1973). *Acoustique Théorique: Propagation et Holophonie [Theoretical acoustics: Propagation and holophony].* New York: Wiley.

Kennedy, R. A., Sadeghi, P., Abhayapala, T. D., & Jones, H. M. (2007). Intrinsic limits of dimensionality and richness in random multipath fields. *IEEE Transactions on Signal Processing, 55*(6), 2542–2556.

Marathay, A. S., & Rock, D. F. (1980). Evanescent wave contribution to the diffracted amplitude for spherical geometry. *Pramana, 14*(4), 315–320.

Morse, P. M., & Feshbach, H. (1953). *Methods of Theoretical Physics.* Feshbach Publishing, LLC: Minneapolis.

Nieto-Vesperinas, M. (2006). *Scattering and Diffraction in Physical Optics.* Singapore: World Scientific Publishing.

Rabenstein, R., Steffen, P., & Spors, S. (1980). Representation of twodimensional wave fields by multidimensional signals. *EURASIP Signal Processing Magazine, 14*(4), 315–320.

Wefers, F. (2008, March). OpenDAFF: Ein freies quell-offenes Software-Paket für richtungsabhängige Audiodaten [OpenDAFF: An open-source software package for direction-dependent audio data]. *Proceedings of 34th DAGA* (pp. 1059–1060). text in German.

Weisstein, E. W. (2002). *CRC Concise Encyclopedia of Mathematics.* London: Chapman and Hall/CRC.

Williams, EG (1999). *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography.* London: Academic.

http://www.springer.com/978-3-642-25742-1

Analytic Methods of Sound Field Synthesis

Ahrens, J.

2012, XVI, 300 p., Hardcover
ISBN: 978-3-642-25742-1