Chapter 4
Spherical Array Acoustic Impulse Response Simulation

In general, the evaluation of acoustic signal processing algorithms, such as direction of arrival (DOA) estimation (see Chap. 5) and speech enhancement (see Chap. 9) algorithms, makes use of simulated acoustic transfer functions (ATFs). By using simulated ATF it is possible to evaluate comprehensively an algorithm under many acoustic conditions, such as a range of reverberation times, room dimensions and source-array distances. Allen and Berkley's image method [2] is a widely used approach to simulate ATFs between an omnidirectional sound source and one or more microphones in a reverberant environment. In the last few decades, several extensions have been proposed [21, 29].

In recent years the use of spherical microphone arrays has become prevalent. These arrays are commonly of one of two types (discussed in Sect. 3.4.2): the open array, where microphones are suspended in free space on an 'open' sphere, and the rigid array, where microphones are mounted on a rigid baffle. As discussed in the previous chapter, the rigid sphere is often preferred as it improves the numerical stability of many processing algorithms [32] and its acoustic scattering effects are can be calculated precisely [25].

Currently, many works relating to spherical array processing consider only free-field responses; however, when a rigid array is used, the rigid baffle causes scattering of the sound waves incident upon the array that the image method does not consider. This scattering has an effect on the ATFs, especially at high frequencies and/or for microphones situated on the occluded side of the array. Furthermore the reverberation due to room boundaries such as walls, ceiling and floor must also be considered, particularly in small rooms or rooms with strongly reflective surfaces.

While measured transfer functions include both these effects, they are both time-consuming and expensive to acquire over a wide range of geometries and rooms. A method for simulating ATFs in a reverberant room while accounting for scattering is therefore essential, allowing for fast, comprehensive and repeatable testing. In this chapter, we present the SMIR (Spherical Microphone array Impulse Response) method that combines a model of the scattering in the spherical harmonic domain (SHD) with a version of the image method that accounts for reverberation in a computationally efficient way [16, 17].

The simulated ATFs include the direct path, reflections due to room reverberation, scattering of the direct path and scattering of the reverberant reflections. Reflections of the scattered sound and multiple interactions between the room boundaries and the sphere are excluded as they do not contribute significantly to the sound field, provided the distances between the room boundaries and the sphere are several times the sphere's radius [11], which is easily achieved in the case of a small scatterer [4]. Furthermore, we assume an empty rectangular shoebox room (with the exception of the rigid sphere) and specular reflections, as was assumed in the conventional image method [2]. Finally, the scattering model used assumes a perfectly rigid baffle, without absorption.

In this chapter, we first briefly summarize Allen and Berkley's image method and then present the SMIR method in the SHD. Next, we discuss some implementation aspects, namely the truncation of an infinite sum in the ATF expression and the reduction of the method's computational complexity, and then provide a pseudocode description of the method. An open-source software implementation is available online [14]. Finally, we show some example uses of the method and, where possible, compare the simulated results obtained with theoretical models.

4.1 Allen and Berkley's Image Method

The source-image or image method [2] is one of the most commonly used room acoustics simulation methods in the acoustic signal processing community. The principle of the method is to model an ATF as the sum of a direct path component and a number of discrete reflections, each of these components being represented in the ATF by a free-space Green's function. In this section, we review the free-space Green's function and the image method.

4.1.1 Green's Function

As detailed in Sect. 2.1, for a source at a position $\llcorner\mathbf{r}_s$ and a receiver at a position $\llcorner\mathbf{r}$,[1] the free-space Green's function, a solution to the inhomogeneous Helmholtz equation

---

[1] Vectors in Cartesian coordinates are denoted with a corner mark $\llcorner$ to distinguish them from vectors in spherical coordinates, which are used throughout this book and will be introduced later in the chapter.

applying the Sommerfeld radiation condition, is given by[2]

$$G(\mathbf{r}|\mathbf{r}_\mathrm{s}, k) = \frac{e^{-ik\|\mathbf{r}-\mathbf{r}_\mathrm{s}\|}}{4\pi\,\|\mathbf{r}-\mathbf{r}_\mathrm{s}\|}, \tag{4.1}$$

where $\|\cdot\|$ denotes the 2-norm and the wavenumber k is related to frequency f (in Hz), angular frequency ω (in rad $\cdot$ s^{-1}) and the speed of sound c (in m $\cdot$ s^{-1}) via the dispersion relation $k = \omega/c = 2\pi f/c$.

In the time-domain, the Green's function is given by

$$g(\mathbf{r}|\mathbf{r}_\mathrm{s}, t) = \frac{\delta(t - \frac{\|\mathbf{r}-\mathbf{r}_\mathrm{s}\|}{c})}{4\pi\,\|\mathbf{r}-\mathbf{r}_\mathrm{s}\|}, \tag{4.2}$$

where δ is the Dirac delta function and t is time. This corresponds to a pure impulse at time $t = \frac{\|\mathbf{r}-\mathbf{r}_\mathrm{s}\|}{c}$, the propagation time from $\mathbf{r}_\mathrm{s}$ to $\mathbf{r}$.

4.1.2 Image Method

Consider a rectangular room with length L_x, width L_y and height L_z. The reflection coefficients of the four walls, floor and ceiling are $\beta_{x_1}, \beta_{x_2}, \beta_{y_1}, \beta_{y_2}, \beta_{z_1}$ and β_{z_2}, where the a_1 coefficients ($a \in \{x, y, z\}$) correspond to the boundaries at $a = 0$ and the a_2 coefficients correspond to the boundaries at $a = L_a$.

If the sound source is located at $\mathbf{r}_\mathrm{s} = (x_\mathrm{s}, y_\mathrm{s}, z_\mathrm{s})$ and the receiver is located at $\mathbf{r} = (x, y, z)$, the images obtained using the walls at $x = 0$, $y = 0$ and $z = 0$ can be expressed as a vector $\mathbf{R_p}$:

$$\mathbf{R_p} = [x_\mathrm{s} - x + 2p_x x, \; y_\mathrm{s} - y + 2p_y y, \; z_\mathrm{s} - z + 2p_z z], \tag{4.3}$$

where each of the elements in $\mathbf{p} = (p_x, p_y, p_z)$ can take values 0 or 1, thus resulting in eight combinations that form a set $\mathcal{P}$. To consider all reflections we also define a vector $\mathbf{R_m}$ which we add to $\mathbf{R_p}$:

$$\mathbf{R_m} = [2m_x L_x, \; 2m_y L_y, \; 2m_z L_z], \tag{4.4}$$

where each of the elements in $\mathbf{m} = (m_x, m_y, m_z)$ can take values between $-N_m$ and N_m, and N_m is used to limit computational complexity and circular convolution errors, thus resulting in a set $\mathcal{M}$ of $(2N_m + 1)^3$ combinations. The image positions in the x and y dimensions are illustrated in Fig. 4.1.

---

[2]This expression assumes the sign convention commonly used in electrical engineering, whereby the temporal Fourier transform is defined as $\mathcal{F}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt$. For more information on this sign convention, the reader is referred to Sect. 2.3.
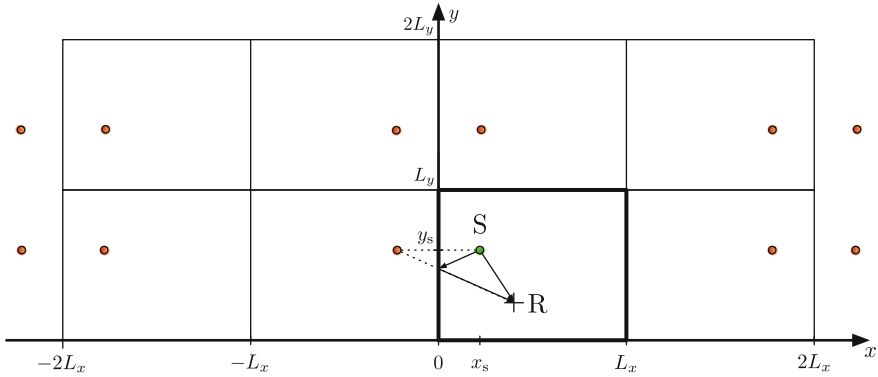
Fig. 4.1 A slice through the image space showing the positions of the images in the x and y dimensions, with a source S and receiver R. The full image space has three dimensions (x, y and z). An example of a reflected path (first-order reflection about the x-axis) is also shown

The distance between an image and the receiver is given by $\left\lVert {}_{\llcorner}\mathbf{R_p} + {}_{\llcorner}\mathbf{R_m} \right\rVert$. Using (4.1), the ATF H is then given by

$$
\begin{aligned}
H({}_{\llcorner}\mathbf{r} | {}_{\llcorner}\mathbf{r_s}, k) = \sum_{\mathbf{p} \in \mathcal{P}} \sum_{\mathbf{m} \in \mathcal{M}} & \beta_{x_1}^{|m_x + p_x|} \beta_{x_2}^{|m_x|} \beta_{y_1}^{|m_y + p_y|} \beta_{y_2}^{|m_y|} \beta_{z_1}^{|m_z + p_z|} \beta_{z_2}^{|m_z|} \\
& \times \frac{e^{-ik \left\lVert {}_{\llcorner}\mathbf{R_p} + {}_{\llcorner}\mathbf{R_m} \right\rVert}}{4\pi \left\lVert {}_{\llcorner}\mathbf{R_p} + {}_{\llcorner}\mathbf{R_m} \right\rVert}.
\end{aligned}
\tag{4.5}
$$

Using (4.2), we obtain the acoustic impulse response (AIR)

$$
\begin{aligned}
h({}_{\llcorner}\mathbf{r} | {}_{\llcorner}\mathbf{r_s}, t) = \sum_{\mathbf{p} \in \mathcal{P}} \sum_{\mathbf{m} \in \mathcal{M}} & \beta_{x_1}^{|m_x + p_x|} \beta_{x_2}^{|m_x|} \beta_{y_1}^{|m_y + p_y|} \beta_{y_2}^{|m_y|} \beta_{z_1}^{|m_z + p_z|} \beta_{z_2}^{|m_z|} \\
& \times \frac{\delta \left(t - \frac{\left\lVert {}_{\llcorner}\mathbf{R_p} + {}_{\llcorner}\mathbf{R_m} \right\rVert}{c} \right)}{4\pi \left\lVert {}_{\llcorner}\mathbf{R_p} + {}_{\llcorner}\mathbf{R_m} \right\rVert}.
\end{aligned}
\tag{4.6}
$$

4.2 SMIR Method in the Spherical Harmonic Domain

There exists a compact analytical expression for the scattering due to the rigid sphere in the SHD, therefore we first express the free-space Green's function in this domain, and then use this to form an expression for the ATF including scattering.

4.2.1 Green's Function

We define position vectors in spherical coordinates relative to the centre of our array. Letting r be the array radius and Ω an inclination-azimuth pair, the microphone position vector is defined as $\mathbf{r} \triangleq (r, \Omega)$ where $\Omega = (\theta, \phi)$. Similarly, the source position vector is given by $\mathbf{r}_s \triangleq (r_s, \Omega_s)$ where $\Omega_s = (\theta_s, \phi_s)$. Consistent with our approach in previous chapters, it is hereafter assumed that where the addition, 2-norm or scalar product operations are applied to spherical polar vectors, they have previously been converted to Cartesian coordinates using (2.12). In addition, we assume that the source is outside the array, i.e., $r_s > r$.

The free-space Green's function (4.1) can be expressed in the SHD using the spherical harmonic expansion (SHE) in (2.22) [40]:

$$
\begin{aligned}
G(\mathbf{r}|\mathbf{r}_s, k) &= \frac{e^{-ik\|\mathbf{r}-\mathbf{r}_s\|}}{4\pi\,\|\mathbf{r} - \mathbf{r}_s\|} \\
&= -ik \sum_{l=0}^{\infty} \sum_{m=-l}^{l} j_l(kr)h_l^{(2)}(kr_s)Y_{lm}^*(\Omega_s)Y_{lm}(\Omega)
\end{aligned}
\tag{4.7}
$$

where Y_{lm} is the spherical harmonic function of order l and degree m, j_l is the spherical Bessel function of order l and $h_l^{(2)}$ is the spherical Hankel function of the second kind and of order l. This decomposition is also known as a spherical Fourier series expansion or spherical harmonic decomposition of the Green's function.

Using the spherical harmonic addition theorem (2.23), which in many cases reduces the complexity of the implementation, we can simplify the Green's function in (4.7) to

$$
G(\mathbf{r}|\mathbf{r}_s, k) = \frac{-ik}{4\pi} \sum_{l=0}^{\infty} j_l(kr)h_l^{(2)}(kr_s)(2l + 1)\mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}_s}),
\tag{4.8}
$$

where $\mathcal{P}_l$ is the Legendre polynomial of order l and $\Theta_{\mathbf{r},\mathbf{r}_s}$ is the angle between $\mathbf{r}$ and $\mathbf{r}_s$. The cosine of the angle $\Theta_{\mathbf{r},\mathbf{r}_s}$ is obtained as the dot product of the two normalized vectors $\hat{\mathbf{r}}_s = \mathbf{r}_s/r_s$ and $\hat{\mathbf{r}} = \mathbf{r}/r$:

$$
\begin{aligned}
\cos \Theta_{\mathbf{r},\mathbf{r}_s} &= \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}_s && \text{(4.9a)} \\
&= \sin\theta\cos\phi\sin\theta_s\cos\phi_s + \sin\theta\sin\phi\sin\theta_s\sin\phi_s \\
&\quad + \cos\theta\cos\theta_s && \text{(4.9b)} \\
&= \sin\theta\sin\theta_s\cos(\phi - \phi_s) + \cos\theta\cos\theta_s. && \text{(4.9c)}
\end{aligned}
$$

4.2.2 Neumann Green's Function

The free-space Green's function describes the propagation of sound in free space only. However, when a rigid sphere is present, a boundary condition must hold: the radial velocity must vanish on the surface of the sphere. The function $G_N(\mathbf{r}|\mathbf{r}_s, k)$ satisfying this boundary condition is called the *Neumann Green's* function, and describes the sound propagation between a point $\mathbf{r}_s$ and a point $\mathbf{r}$ on the rigid sphere [40]:

$$
\begin{aligned}
G_N(\mathbf{r}|\mathbf{r}_s, k) &= G(\mathbf{r}|\mathbf{r}_s, k) - \frac{-ik}{4\pi} \sum_{l=0}^{\infty} \frac{j_l'(kr)h_l^{(2)}(kr)}{h_l^{(2)'}(kr)} h_l^{(2)}(kr_s)(2l+1)\mathcal{P}_l(\cos\Theta_{\mathbf{r},\mathbf{r}_s}) \\
&= \frac{-ik}{4\pi} \sum_{l=0}^{\infty} i^{-l} b_l(k) h_l^{(2)}(kr_s)(2l+1)\mathcal{P}_l(\cos\Theta_{\mathbf{r},\mathbf{r}_s}),
\end{aligned}
\tag{4.10}
$$

where $(\cdot)'$ denotes the first derivative and the term

$$
b_l(k) = i^l \left(j_l(kr) - \frac{j_l'(kr)}{h_l^{(2)'}(kr)} h_l^{(2)}(kr) \right)
\tag{4.11}
$$

is often called the (farfield) *mode strength*. The Wronskian relation [40, Eq. 6.67]

$$
j_l(x)h_l^{(2)'}(x) - j_l'(x)h_l^{(2)}(x) = -\frac{i}{x^2}
\tag{4.12}
$$

allows us to simplify (4.11) to

$$
b_l(k) = \frac{-i^{l+1}}{h_l^{(2)'}(kr)\,(kr)^2}.
\tag{4.13}
$$

For the open sphere, substituting $b_l(k) = i^l j_l(kr)$ into (4.10) yields the free-space Green's function $G(\mathbf{r}|\mathbf{r}_s, k)$.

4.2.3 Scattering Model

The rigid sphere scattering model[3] used by the SMIR method has a long history in the literature; it was first developed by Clebsch and Rayleigh in 1871–72 [23]. It is

---

[3]Some texts [9] refer to the scattering effect as diffraction, although Morse and Ingard note that *"When the scattering object is large compared with the wavelength of the scattered sound, we usually say the sound is reflected and diffracted, rather than scattered"* [28], therefore in the case of spherical microphone arrays (particularly rigid ones which tend to be relatively small for practical reasons), scattering is possibly the more appropriate term.

presented in a number of acoustics texts [28, 36, 40], and is used in many theoretical analyses for spherical microphone arrays [26, 33].

4.2.3.1 Theoretical Behaviour

The behaviour of the scattering model is illustrated in Fig. 4.2, which plots the magnitude of the response between a source and a receiver on a rigid sphere of radius 5 cm for a source-array distance of 1 m, as a function of frequency and DOA. The response was normalized using the free-field/open sphere response, therefore the plot shows only the effect due to scattering. Due to rotational symmetry, we only looked at the one-dimensional DOA, instead of looking at both azimuth and inclination, and limited the DOA to the 0–180° range.

When the source is located on the same side of the sphere as the receiver and the direction of arrival is 0°, the rigid sphere response is greater than the open sphere response due to constructive scattering, tending towards a 6 dB magnitude gain compared to the open sphere at infinite frequency. The response on the back side of the rigid sphere is generally lower than in the open sphere case and lower than on the front side, as one would intuitively expect due to it being occluded. However, at the very back of the sphere, when the DOA is 180°, we observe a narrow *bright spot*: the waves propagating around the sphere all arrive in phase at the 180° point and as a result sum constructively.

The polar plot of the magnitude response is shown in Fig. 4.3 and illustrates both the amplification on the front side of the sphere, and attenuation on the back side of
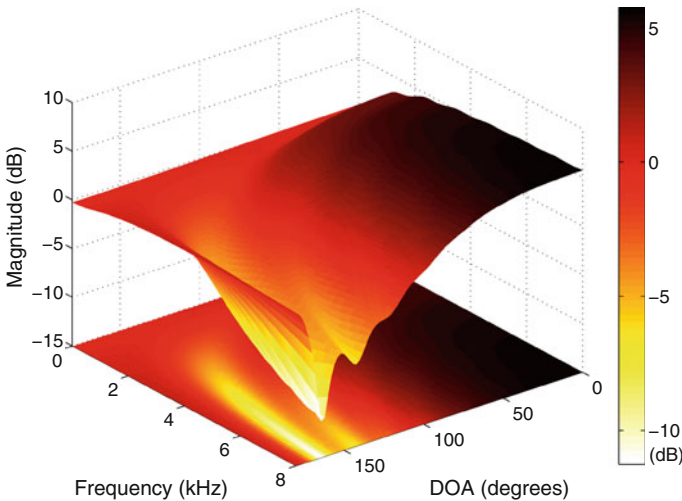
Fig. 4.2 Magnitude of the response between a source and a receiver placed on a rigid sphere of radius 5 cm at a distance of 1 m, as a function of frequency and DOA. The response was normalized with respect to the free-field response

Fig. 4.3 Polar plot of the magnitude of the response between a source and a receiver placed on a rigid sphere of radius 5 cm, at a distance of 1 m, for various frequencies
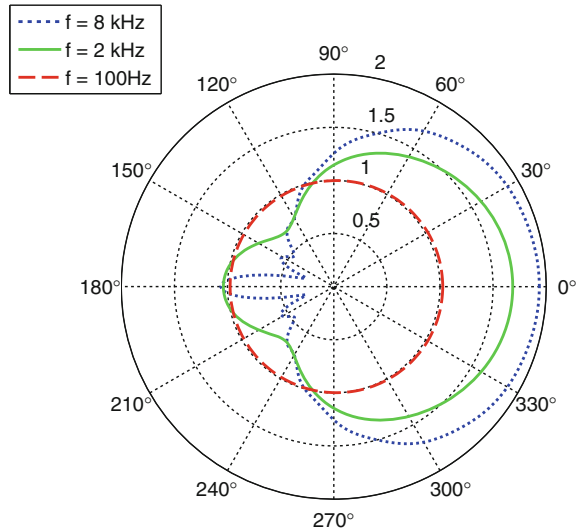
the sphere, which both increase with increasing frequency. It should be noted that although the above results are for a fixed sphere radius, as the scattering model is a function of kr, the effects of a change in radius are the same as a change in frequency; indeed the relevant factor is the radius of the sphere relative to the wavelength.

These substantial differences between the open and rigid sphere responses confirm the need for a simulation method which accounts for scattering, even for sphere radii as small as 5 cm.

4.2.3.2   Experimental Validation

In addition to being widely used in theory, this model has also been experimentally validated by Duda and Martens [9] using a single microphone inserted in a hole drilled through a 10.9 cm radius bowling ball placed in an anechoic chamber. This is a reasonable approximation to a spherical microphone array; indeed a bowling ball was used by Li and Duraiswami to construct a hemispherical microphone array [22].

Duda and Martens's experimental results broadly agree with the theoretical model. In our case we are most interested in the results in their Fig. 12a where the source-array distance is largest (20 times the array radius), as in typical spherical array usage scenarios the source is unlikely to be much closer to the array than this. The only notable difference between the theoretical and experimental results in this case is for a direction of arrival of 180°, where the high frequency response is found to be lower than expected. The authors suggest this is due to small alignment errors, which would indeed have an effect given the narrowness of the bright spot in the model (see Fig. 4.3 for $f = 8$ kHz). Given these results, we conclude that the use of this

scattering model is sufficiently accurate for simulating a small rigid array, such as the *Eigenmike* [27].

4.2.4   SMIR Method

We now present the SMIR method proposed in [16, 17], incorporating the SHE presented in Sect. 4.2.1 and the scattering model introduced in Sect. 4.2.2.

Due to the differences between the SHD Neumann Green's function in (4.10) and the spatial domain Green's function in (4.1), as well as the directionality of the array's response, two changes must be made to the image position vectors $\mathbf{R_p}$ and $\mathbf{R_m}$ in the SMIR method. Firstly, to compute the SHE in the Neumann Green's function, we require the distance between each image and the *centre* of the array [r_s in (4.10)]; this is accomplished by computing the image position vectors using the position of the centre of the array rather than the position of the receiver. Secondly, to compute the SHE we require the angle between each image and the receiver with respect to the centre of the array [$\Theta_{\mathbf{r},\mathbf{r}_s}$ in (4.10)]. In Allen and Berkley's image method, the direction of the vector $\mathbf{R_p} + \mathbf{R_m}$ is not always the same: in some cases it points from the receiver to the image and in others it points from the image to the receiver. This is not an issue for the image method as only the norm of this vector is used. Because we also require the angle of the images in the SMIR method, we modify the definition of $\mathbf{R_p}$ such that the vector $\mathbf{R_p} + \mathbf{R_m}$ always points from the centre of the array to the image.

We now incorporate these two changes into the definition of the image vectors $\mathbf{R_p}$ and $\mathbf{R_m}$. If the sound source is located at $\mathbf{r}_s = (x_s, y_s, z_s)$ and the centre of the sphere is located at $\mathbf{r}_a = (x_a, y_a, z_a)$, the images obtained using the walls at $x = 0$, $y = 0$ and $z = 0$ are expressed as a vector $\mathbf{R_p}$:

$$\mathbf{R_p} = [x_s - 2p_x x_s - x_a, \, y_s - 2p_y y_s - y_a, \, z_s - 2p_z z_s - z_a]. \tag{4.14}$$

For brevity we define $\mathbf{R_{p,m}} \triangleq \mathbf{R_p} + \mathbf{R_m}$, allowing us to express the distance between an image and the centre of the sphere as $||\mathbf{R_{p,m}}||$ and the angle between the image and the receiver as $\Theta_{\mathbf{r},\mathbf{R_{p,m}}}$, computed in the same way as (4.9), where $\mathbf{R_{p,m}}$ denotes the image positions in spherical coordinates. The image positions in the x dimension are illustrated in Fig. 4.4. Finally, the ATF $H(\mathbf{r}|\mathbf{r}_s, k)$ is the weighted sum of the individual responses $G_N(\mathbf{r}|\mathbf{R_{p,m}}, k)$ for each of the images[4]

$$H(\mathbf{r}|\mathbf{r}_s, k) = \sum_{\mathbf{p}\in\mathcal{P}} \sum_{\mathbf{m}\in\mathcal{M}} \beta_{x_1}^{|m_x - p_x|} \beta_{x_2}^{|m_x|} \beta_{y_1}^{|m_y - p_y|} \beta_{y_2}^{|m_y|} \beta_{z_1}^{|m_z - p_z|} \beta_{z_2}^{|m_z|}$$
$$\times \, G_N(\mathbf{r}|\mathbf{R_{p,m}}, k). \tag{4.15}$$

---

[4]The sign in the powers of β is different from that in Allen and Berkley's conventional image method, due to the change in the definition of $\mathbf{R_p}$ that is required for the SMIR method.
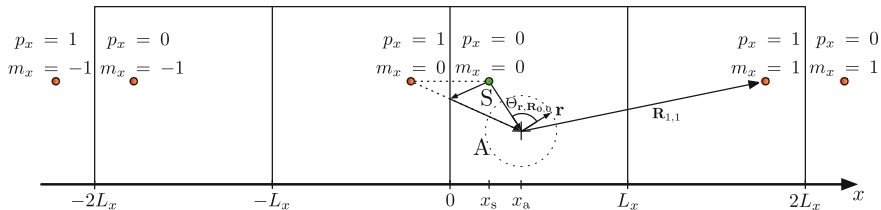
Fig. 4.4 A slice through the image space showing the positions of the images in the x dimension, with a source S and array A. The full image space has three dimensions (x, y and z). An example of a reflected path is shown for the image with $p_x = 1$ and $m_x = 0$

Since we are working in the wavenumber domain, we can allow for frequency dependent boundary reflection coefficients in (4.15), if desired. The reflection co-efficients would then be written as $\beta_{x_1}(k)$, $\beta_{x_2}(k)$ and so on. Chen and Maher [7] provide some measured reflection coefficients for a wall, window, floor and ceiling.

4.3   Implementation

4.3.1   Truncation Error

To compute the expression for the ATF in (4.15), the sum over an infinite number of orders l in the Neumann Green's function G_N must be approximated by a sum $\hat{G}_\mathrm{N}$ over a finite order L. Choosing L too small will result in a large approximation error, while choosing L too large will result in too high a computational complexity. We now investigate the approximation error in order to provide some guidelines for the choice of the order L. The results for an open sphere are provided for reference, and were computed by using a truncated SHE of the Green's function $\hat{G}$ instead of a Neumann Green's function.

For an open sphere, the error can be determined exactly because the Green's function is a decomposition of the closed-form expression in (4.1). For a rigid sphere, however, no closed-form expression exists since the scattering term can be expressed only in the SHD. We therefore estimated the error by comparing the truncated Neumann Green's function $\hat{G}_\mathrm{N}$ to a high-order Neumann Green's function. We can assume the error involved in using a high-order Neumann Green's function as a reference as opposed to the untruncated Neumann Green's function is small, due to the uniform convergence of the SHE [12]. In practice, we cannot choose very large values of L because of numerical difficulties involved in multiplying high order spherical Bessel and Hankel functions. For typical sphere radii and source-array distances, this allows us to reach L values of up to about 100 using SMIRgen, a MATLAB implementation of the SMIR method [14].

We evaluated the truncated (Neumann) Green's function at $K = 1024$ discrete values of k (denoted by $\dot{k}$), forming a set $\mathcal{K}$ corresponding to frequencies in the range
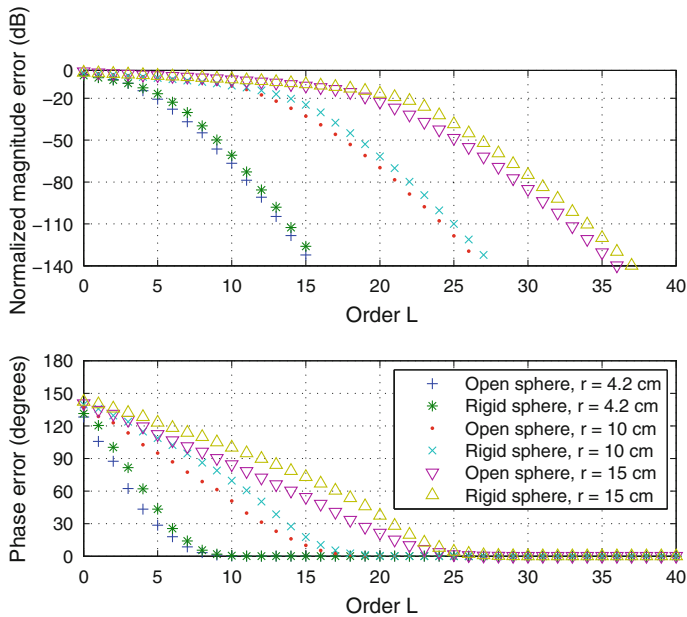
Fig. 4.5 Magnitude and phase errors involved in the truncation of the SHE in the Green's function (open sphere) and the Neumann Green's function (rigid sphere). The errors reduce rapidly beyond $L = k_{max}r$, where here $k_{max} = \frac{2\pi\,8000}{c} \approx 147\,\mathrm{m}^{-1}$

100 Hz–8 kHz,[5] and then calculated the normalized root-mean-square magnitude error ϵ_m and the root-mean-square phase error ϵ_p:

$$\epsilon_m(\mathbf{r}|\mathbf{r}_s, L) = \sqrt{\frac{1}{K}\sum_{\dot{k}\in\mathcal{K}}\frac{\left(\left|G_N(\mathbf{r}|\mathbf{r}_s,\dot{k})\right| - \left|\hat{G}_N(\mathbf{r}|\mathbf{r}_s,\dot{k},L)\right|\right)^2}{\left|G_N(\mathbf{r}|\mathbf{r}_s,\dot{k})\right|^2}}, \qquad (4.16)$$

$$\epsilon_p(\mathbf{r}|\mathbf{r}_s, L) = \sqrt{\frac{1}{K}\sum_{\dot{k}\in\mathcal{K}}\left(\angle G_N(\mathbf{r}|\mathbf{r}_s,\dot{k}) - \angle\hat{G}_N(\mathbf{r}|\mathbf{r}_s,\dot{k},L)\right)^2}. \qquad (4.17)$$

We averaged the magnitude and phase errors over 32 microphone positions uniformly distributed on the array and 50 random source positions at a fixed distance from the centre of the array.

The resulting average errors are given in Fig. 4.5, for both the open and rigid sphere cases. Three different sphere radii were used: $r = 4.2$ cm (the radius of the *Eigenmike* [24]), $r = 10$ cm and $r = 15$ cm. A source-array distance of 1 m was used; results for 1–5 m are omitted as they are essentially identical. It can be seen that

---

[5]Very low frequencies are omitted due to the fact that the spherical Hankel function $h_l(x)$ has a singularity around $x = 0$.

beyond a certain threshold, increases in L give only a very small reduction in error; this is due to the fast convergence of the spherical harmonic decomposition [12]. From Fig. 4.5, we can see that a sensible rule of thumb for choosing L is $L > \lceil 1.1 \, k_{\max} r \rceil$ where $k_{\max}$ is the largest wavenumber of interest.

4.3.2 Computational Complexity

As the ATFs are made up of a sum over all orders l which includes spherical Hankel functions h_l and Legendre polynomials $\mathcal{P}_l$, we can make use of recursion relations over l to reduce the computational complexity of these functions. For the spherical Hankel function, we make use of the following relation [40, Eq. 6.69]

$$h_m^{(2)}(x) = \frac{2m-1}{x} h_{m-1}^{(2)}(x) - h_{m-2}^{(2)}(x), \ m \geq 2 \tag{4.18}$$

where

$$h_0^{(2)}(x) = -\frac{e^{-ix}}{ix}, \ h_1^{(2)}(x) = \frac{i e^{-ix}}{x^2} - \frac{e^{-ix}}{x}. \tag{4.19}$$

For the Legendre polynomial we use a similar recursion relation [1], known as Bonnet's recursion formula

$$\mathcal{P}_m(x) = \frac{2m-1}{m} x \mathcal{P}_{m-1}(x) - \frac{m-1}{m} \mathcal{P}_{m-2}(x), \ m \geq 2 \tag{4.20}$$

where $\mathcal{P}_0(x) = 1$ and $\mathcal{P}_1(x) = x$.

While replacing the exponential in (4.1) with a SHE does lead to an increase in computational complexity when computing the ATF for a single receiver (which is unavoidable in the rigid sphere case), this can have an advantage when simulating many receiver positions. For the conventional image method, we must compute the image positions and resulting response separately for each individual receiver. However, in the SMIR method the image positions are all computed with respect to the *centre* of our array, and therefore only once for all of the microphones in the array.

An alternative to (4.15) is obtained by changing the order of the summations in the ATF and computing the sum over all images only once, instead of once per receiver:

$$\begin{aligned}
H(\mathbf{r}|\mathbf{r}_s, k) = -ik \sum_{l=0}^{\infty} i^{-l} \sum_{m=-l}^{l} Y_{lm}(\Omega) \\
\times \sum_{\mathbf{p} \in \mathcal{P}} \sum_{\mathbf{m} \in \mathcal{M}} \beta_{x_1}^{|m_x - p_x|} \beta_{x_2}^{|m_x|} \beta_{y_1}^{|m_y - p_y|} \beta_{y_2}^{|m_y|} \beta_{z_1}^{|m_z - p_z|} \beta_{z_2}^{|m_z|} \\
\times b_l(k) h_l^{(2)}(k \, ||\mathbf{R}_{\mathbf{p},\mathbf{m}}||) Y_{lm}^*(\angle \mathbf{R}_{\mathbf{p},\mathbf{m}}).
\end{aligned} \tag{4.21}$$

The expression in (4.21) requires $O\left((N_i + Q)(L + 1)^2\right)$ operations per discrete frequency, where L is the maximum spherical harmonic order, N_i is the number of images and Q is the number of microphones, while the approach in (4.15) requires $O\left(N_i Q(L + 1)\right)$ operations per discrete frequency. Since the number of images N_i is typically very large, $(N_i + Q)(L + 1)^2 \approx N_i(L + 1)^2$. Assuming the operations in the two approaches are of similar complexity, it is therefore more efficient to use the expression in (4.15) for $Q < L + 1$ and the expression in (4.21) for $Q > L + 1$. Consequently the least computationally complex approach depends on the number of microphones Q and array radius r. In the remainder of this chapter we use the expression in (4.15); this is particularly appropriate in the applications in Sect. 4.4.2 where $Q = 2$ and in Sect. 4.4.3 where $Q = 1$.

4.3.3 Algorithm Summary

A summary of the SMIR method is presented in the form of pseudocode in Fig. 4.6. The variable $nsample$ denotes the number of samples in the AIR, N_o, the maximum reflection order, and fs, the sampling frequency.

The number of computations has been reduced by processing only half of the frequency spectrum because we know the AIR is real and the corresponding ATF is conjugate symmetric. The pseudocode necessary to compute the Hankel functions and Legendre polynomials is omitted here, since their computation is straightforward using recursion relations (4.18) and (4.20).

SMIRgen, a MATLAB/C++ implementation of the method in the form of a MEX-function, is available online [14].

4.4 Examples and Applications

In this section we give a number of examples that make use of the SMIR method. Wherever possible we compare the simulated results to theoretical results obtained using approximate models. These examples are given to illustrate and partially validate the SMIR method.

4.4.1 Diffuse Sound Field Energy

In statistical room acoustics (SRA), reverberant sound fields are modelled as diffuse sound fields, allowing for a statistical analysis of reverberation instead of computing each of the individual reflections. In this subsection, we compare a theoretical prediction of sound energy on the surface of a rigid sphere, based on a diffuse model of reverberation, to simulated results obtained using the SMIR method.

Fig. 4.6 Pseudocode for the SMIR method

1: $\mathcal{P} = \{0,1\}^3$
2: $\mathcal{M} = \{-N_m, \cdots, 0, \cdots, N_m\}^3$
3: $\mathcal{A} = \mathcal{P} \times \mathcal{M}$

4: **for** $(\mathbf{p}, \mathbf{m}) \in \mathcal{A}$ **do**
5:      **if** $|2m_x - p_x| + |2m_y - p_y| + |2m_z - p_z| \leq N_o$ **then**
6:          $\mathbf{R_{p,m}} = \begin{bmatrix} x_s - 2p_x x_s - x_a + 2m_x L_x \\ y_s - 2p_y y_s - y_a + 2m_y L_y \\ z_s - 2p_z z_s - z_a + 2m_z L_z \end{bmatrix}$
7:          $\beta(\mathbf{p}, \mathbf{m}) = \beta_{x_1}^{|m_x - p_x|} \beta_{x_2}^{|m_x|} \beta_{y_1}^{|m_y - p_y|} \beta_{y_2}^{|m_y|} \beta_{z_1}^{|m_z - p_z|} \beta_{z_2}^{|m_z|}$
8:      **else**
9:          $\mathcal{A} = \mathcal{A} \setminus \{(\mathbf{p}, \mathbf{m})\}$
10:     **end if**
11: **end for**

12: **for** $k = 1 \rightarrow nsample/2 + 1$ **do**
13:     **for** $l = 0 \rightarrow L$ **do**
14:         **if** $sphType = \text{'rigid'}$ **then**
15:             $\Delta(k,l) = \frac{-i^{l+1}}{h_l^{(2)'}(kr)(kr)^2}$
16:         **else**
17:             $\Delta(k,l) = i^l j_l(kr)$
18:         **end if**
19:         $\Gamma(k,l) = \frac{-iki^{-l}}{4\pi} \cdot \Delta(k,l)$
20:     **end for**
21: **end for**

22: **for** $(\mathbf{p}, \mathbf{m}) \in \mathcal{A}$ **do**
23:     **if** $||\mathbf{R_{p,m}}|| + r < c \cdot nsample/fs$ **then**
24:         **for** $ang = 1 \rightarrow Q$ **do**
25:             $\Theta = \hat{\mathbf{R}}_{\mathbf{p,m}} \cdot \hat{\mathbf{r}}(ang)$
26:             **for** $l = 0 \rightarrow L$ **do**
27:                 $\Upsilon(ang,l) = \mathcal{P}_l(\Theta) \cdot (2l+1)$
28:             **end for**
29:         **end for**
30:         **for** $k = 1 \rightarrow nsample/2 + 1$ **do**
31:             **for** $l = 0 \rightarrow L$ **do**
32:                 $\Lambda(k,l) = h_l(k||\mathbf{R_{p,m}}||) \cdot \Gamma(k,l)$
33:             **end for**
34:         **end for**
35:         **for** $ang = 1 \rightarrow Q$ **do**
36:             **for** $k = 1 \rightarrow nsample/2 + 1$ **do**
37:                 **for** $l = 0 \rightarrow L$ **do**
38:                     $H(\mathbf{p}, \mathbf{m}, ang, k, l) = \beta(\mathbf{p}, \mathbf{m}) \cdot \Upsilon(ang, l) \cdot \Lambda(k, l)$
39:                 **end for**
40:                 $H(\mathbf{p}, \mathbf{m}, ang, k) = \sum_l H(\mathbf{p}, \mathbf{m}, ang, k, l)$
41:             **end for**
42:         **end for**
43:     **end if**
44: **end for**
45: $H(ang, k) = \sum_{(\mathbf{p}, \mathbf{m}) \in \mathcal{A}} H(\mathbf{p}, \mathbf{m}, ang, k)$
46: $h(ang, n) = \text{IFFT}_\text{R}\{H(ang, k)\}$

A diffuse sound field is composed of plane waves incident from all directions with equal probability and amplitude [20]. Using the scattering model previously introduced, we can determine the cross-correlation between the sound pressure at arbitrary positions $\mathbf{r}$ and $\mathbf{r}'$ on the surface of a sphere, due to a unit amplitude plane wave with a random uniformly distributed direction of arrival (see the Appendix for derivation) [15]:

$$C(\mathbf{r}, \mathbf{r}', k) = \sum_{l=0}^{\infty} |b_l(k)|^2 (2l+1) \mathcal{P}_l(\cos \Theta_{\mathbf{r}, \mathbf{r}'}), \tag{4.22}$$

where $\Theta_{\mathbf{r}, \mathbf{r}'}$ is the angle between $\mathbf{r}$ and $\mathbf{r}'$. In the open sphere case, it is shown in the Appendix that this simplifies to the well-known spatial domain expression [20, 31, 39] $\mathrm{sinc}(k \left\| \mathbf{r} - \mathbf{r}' \right\|)$, where sinc denotes the unnormalized sinc function.

For the sound energy at a position $\mathbf{r}$ we substitute $\Theta_{\mathbf{r}, \mathbf{r}} = 0$ and find $C(\mathbf{r}, \mathbf{r}, k) = \sum_{l=0}^{\infty} |b_l(k)|^2 (2l+1)$. According to SRA theory [20, 39], for frequencies above the Schroeder frequency [20] the energy of the reverberant sound field H_{r} is then given by [39]

$$\begin{aligned} \mathrm{E_s} \left\{ |H_{\mathrm{r}}(\mathbf{r}, k)|^2 \right\} &= \frac{1 - \bar{\alpha}}{\pi A \bar{\alpha}} C(\mathbf{r}, \mathbf{r}, k) \\ &= \frac{1 - \bar{\alpha}}{\pi A \bar{\alpha}} \sum_{l=0}^{\infty} |b_l(k)|^2 (2l+1), \end{aligned} \tag{4.23}$$

where $\mathrm{E_s} \{\cdot\}$ denotes spatial expectation, $\bar{\alpha}$ is the average wall absorption coefficient and A is the total wall surface area.

The above theoretical expression for the average reverberant energy can be compared to simulated results obtained using the SMIR method. We computed the spatial expectation using an average over 200 source-array positions, using the approach in Radlović et al. [31]: the array and source were kept in a fixed configuration (at a distance of 2 m from each other), which was then randomly rotated and translated. Both sources and microphones were kept at least half a wavelength from the boundaries of the room, helping to ensure the diffuseness of the reverberant sound field [20]. The reverberant component H_{r} of the ATFs was computed by subtracting the direct path H_{d} from the simulated ATFs.

The room dimensions were chosen as $6.4 \times 5 \times 4$ m, as in [31, 38], such that the ratio of the dimensions was $(1.6 : 1.25 : 1)$, as recommended in [18, 31] to approximate a diffuse sound field. The reverberation time T_{60} was set to 500 ms, giving an average wall absorption coefficient of $\bar{\alpha} = 0.2656$. We simulated AIRs with a length of 4096 samples at a sampling frequency of 8 kHz. We considered frequencies from 300 Hz to 4 kHz, well above the Schroeder frequency of $2000\sqrt{\frac{0.5}{6.4 \times 5 \times 4}} = 125$ Hz, and the half-wavelength minimum distance is therefore 57 cm for a speed of sound of 343 m/s. We averaged the results over the 200 source-array positions and 32 microphone positions uniformly distributed on the array.
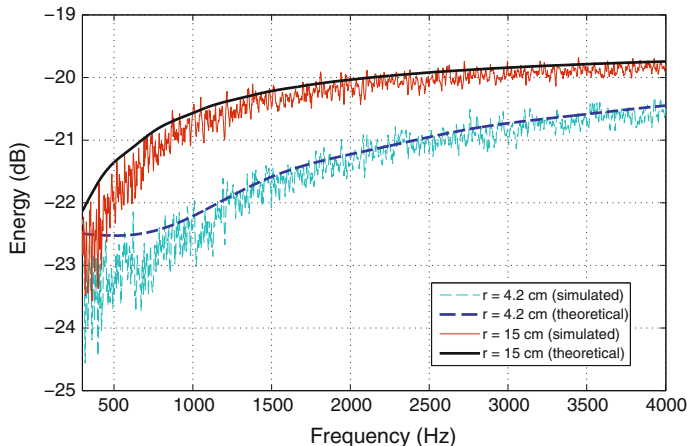
Fig. 4.7 Theoretical and simulated reverberant sound field energy on the surface of a rigid sphere, as a function of frequency for two array radii. The simulated results are averaged over 200 source-array positions, all at least half a wavelength from the room boundaries

In Fig. 4.7, we plot the theoretical and simulated energy of H_r as a function of frequency, for two array radii (4.2 and 15 cm). We note that, except at low frequencies, there is a good match between the theoretical diffuse field energy expression we derived and the results obtained using the SMIR method. At lower frequencies, the theoretical equation overestimates the energy; we hypothesize that this is due to the reverberant sound field not being fully diffuse.

4.4.2 Binaural Interaural Time and Level Differences

The topic of binaural sound and in particular head-related transfer functions (HRTFs) or head-related impulse responses (HRIRs) is of interest to researchers and engineers working on surround sound reproduction, who for example aim to reproduce spatial audio through a pair of stereo headphones. In addition, the psychoacoustic community is interested in the ability of the human brain to localize sound sources using only two ears.

Two binaural cues that contribute to sound source localization in humans are the interaural time difference (ITD) and the interaural level difference (ILD) [34]. The ITD measures the difference in arrival time of a sound at the two ears, and the ILD measures the difference in level of the sound at the two ears. In this example, we study the long-term cues assuming the source signal is spectrally white. Therefore, we can compute the cues directly using the simulated ATFs.

We used the SMIR method to simulate a simple HRTF by considering microphones placed at locations on a rigid sphere corresponding to ear positions on the

human head. Although real HRTFs vary from individual to individual, depending on many factors including the head, torso and pinnae, many of the main characteristics of the HRTF are also exhibited by a simple rigid sphere ATF [9]. The representation of HRTFs using spherical harmonics was studied in [3, 10].

Whereas HRTFs do not normally include the effects of reverberation, and as a result typically sound artificial and provide poor cues for the perception of sound source distance [37], the SMIR method also allows for the inclusion of reverberation in HRIRs. In this case, they are then referred to as binaural room impulse responses (BRIRs). BRIRs are important for the analysis of the effects of reverberation on auditory perception, for example its impact on localization accuracy. Since rotational symmetry no longer necessarily holds once the room reflections are taken into account, the measurement of BRIRs must be done for every source-head position and orientation and is therefore very time-consuming. Simulating BRIRs allows us to more easily study the effects of early and late reflections on the binaural cues.

We begin by looking at ITDs in an anechoic environment, in order to illustrate the effect of the head in isolation. We compare simulated results to approximate theoretical results provided by a ray-tracing formula attributed to Woodworth and Schlosberg that looks at the distance travelled from the source to an observation point on the sphere, either in free-space if the observation point is on the near side of the sphere, or via a point of tangency if the observation point is on the far side [9].

The simulated results were obtained by using the SMIR method to generate HRIRs at a sampling frequency of 32 kHz, with a sphere radius of 8.75 cm and microphones placed at $(\theta, \phi) = (90°, 100°)$ (corresponding to the left ear) and $(\theta, \phi) = (90°, 260°)$ (corresponding to the right ear). The HRIRs were then band-pass filtered between 2.8 and 3.2 kHz.[6] The DOA was varied by rotating the source around the sphere at a fixed distance of 1 m and inclination of 90°. The simulated ITD was computed by determining the time delay that maximized the interaural cross-correlation between the two simulated and band-pass filtered HRIRs. The cross-correlation was interpolated using a second-order polynomial in order to obtain sub-sample delays.

In Fig. 4.8 we plot the ITDs as a function of direction of arrival, where 0° corresponds to the median plane on the front side of the sphere and 180° corresponds to the median plane on the back side of the sphere. As expected, as the DOA increases from 0° to 80° and the source gets closer to the ipsilateral ear, the ITD increases monotonically until it reaches its maximum at 80°, at which point the source is furthest from the contralateral ear. The ITD then decreases from 80° to 180° as the

---

[6]While the ray-tracing formula is frequency-independent, it has been shown [6] that ITDs actually exhibit some frequency dependence, and that because the ray-tracing concept applies to short wavelengths, this model yields only the high frequency time delay. Kuhn provides a more comprehensive discussion of this model and the frequency-dependence of ITDs [19]. It should be noted the simulation results in Fig. 4.8 are in broad agreement with Kuhn's measured results at 3.0 kHz.
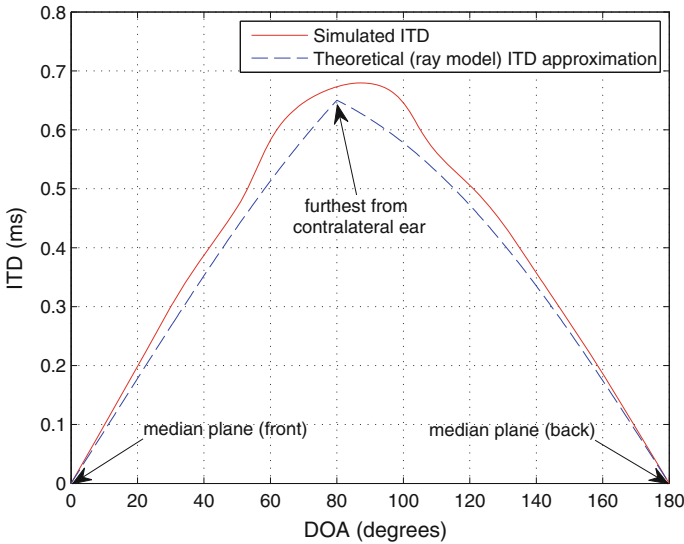
Fig. 4.8 Comparison of ITDs as a function of source DOA, in simulation and using the theoretical ray model approximation. The simulated ITDs are based on HRIRs computed using the SMIR method in an anechoic environment

source nears the median plane and gets closer to the contralateral ear. The response from 180° to 360° is not shown due to the symmetry about 180°. As we expect, the simulated results are reasonably close to the theoretical ray-tracing results [9], with a difference of less than 70 µs.

Using the SMIR method, we analyzed the ILDs in a reverberant environment under three scenarios: the sphere was either placed in the centre of the room with a DOA of 0° (where the source is equidistant from the two ears), or at a distance of approximately 0.5 m from one of the walls with DOAs of 0° and 100° (where the source is aligned with the left ear). In all three cases the source was placed at a distance of 1 m from the centre of the sphere. We chose a room size of $9 \times 5 \times 3$ m with a reverberation time T_{60} of 500 ms, and simulated BRIRs with a length of 4096 samples at a sampling frequency of 8 kHz.

In Figs. 4.9, 4.10 and 4.11 we plot the ILDs for the three above cases, as well as the ILDs we would obtain in an anechoic environment, which are entirely due to scattering. The ILDs were computed by taking the difference in magnitude between the left ear response and the right ear response. A negative ILD therefore indicates that the magnitude of the ipsilateral ear response is lower than that of the contralateral
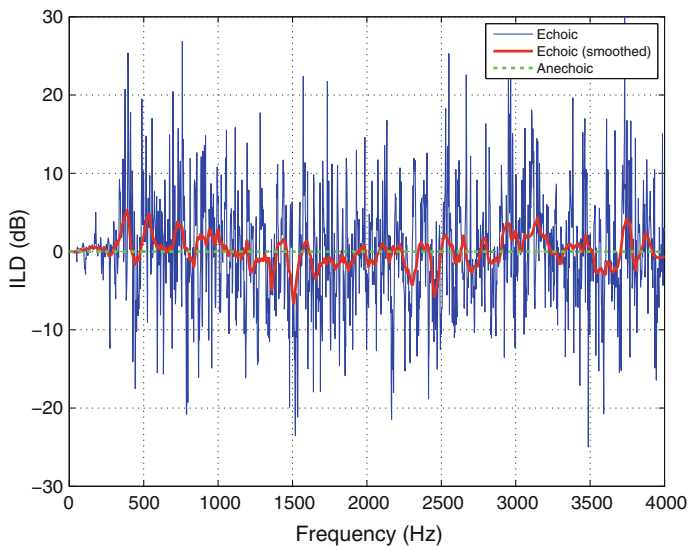
Fig. 4.9 Comparison of ILDs in echoic and anechoic environments, with the sphere placed in the centre of the room and a DOA of $0°$. The ILDs are based on HRTFs (anechoic) and BRIRs (echoic) computed using the SMIR method
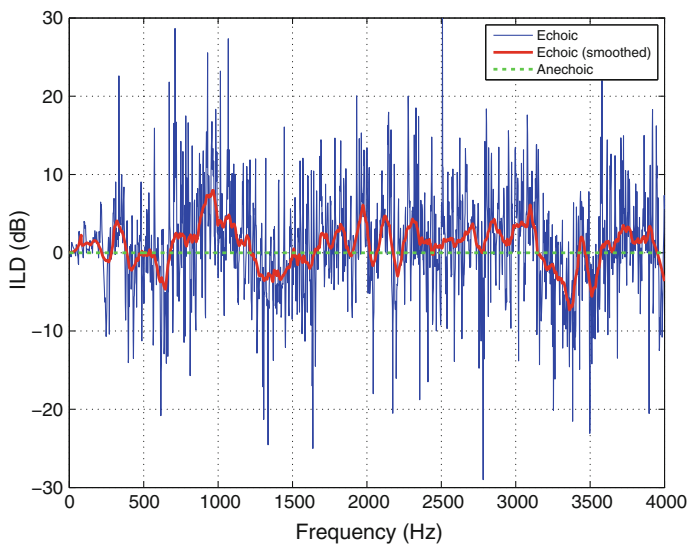
Fig. 4.10 Comparison of ILDs in echoic and anechoic environments, with the sphere placed near a room wall and a DOA of $0°$
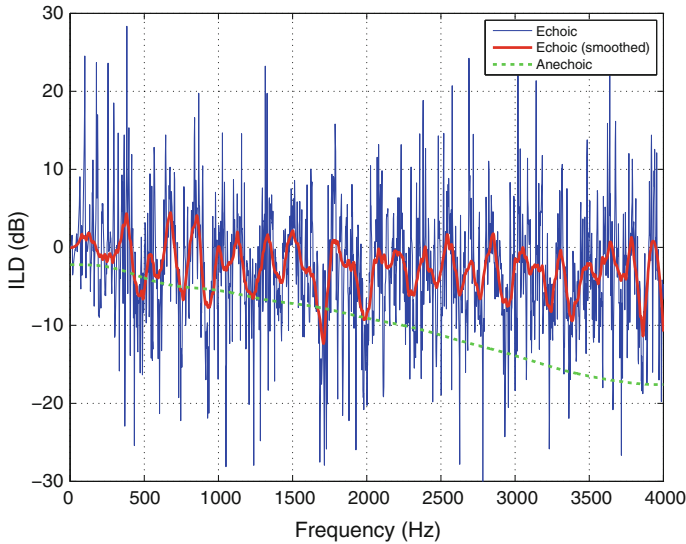
Fig. 4.11 Comparison of ILDs in echoic and anechoic environments, with the sphere placed near a room wall and a DOA of 100°

ear response. The smoothed echoic ILDs were obtained using a Savitzky-Golay smoothing filter [35].

The main effect of reverberation we can observe is the introduction of random frequency-to-frequency variations; these are particularly obvious when most of the reverberant energy is diffuse, for example, when the sphere is placed in the centre of the room (Fig. 4.9). Room reflections also increase the overall reverberant energy, particularly in the contralateral ear which receives less direct path energy, thus reducing the ILDs. This is especially noticeable when the contralateral ear is placed near a wall: the contralateral ear receives more energy than in the anechoic case and the ILD is therefore closer to zero (Fig. 4.11).

Placement of the sphere near a wall additionally introduces systematic distortions in the ILDs associated with the prominent early reflection from this wall. This is visible in Fig. 4.11 and most noticeably in Fig. 4.10.

All these effects have also been observed experimentally with a manikin by Shinn-Cunningham et al. [37]. The SMIR method is therefore an inexpensive way of predicting the effects of head movement and environmental changes (such as reverberation time) on HRTFs or BRIRs, without as much need for physical and acoustic measurements to be performed.

4.4.3   Mouth Simulator

The principle of reciprocity can often be advantageously used in room acoustics measurements. The principle states that ATFs are symmetric in the coordinates of the sound source and the observation point: *"If we put the sound source at* **r**, *we observe at point* **r**$_0$ *the same sound pressure as we did before at* **r**, *when the sound source was at* **r**$_0$*"* [20]. We can apply this principle to ATF simulations, and use the SMIR method to generate the ATF between one or more sources on a sphere and a single omnidirectional microphone placed away from the sphere.

A specific application of this is a mouth simulator: we model the head as a rigid sphere (as in Sect. 4.4.2) of radius r_h, and the mouth as an omnidirectional point source placed on this rigid sphere. This is straightforwardly implemented in the SMIR method by replacing the source position with the microphone position $\mathbf{r}_{mic}$, the microphone position with the mouth position $\mathbf{r}_{mouth} = (r_h, \Omega_{mouth})$, and the array position with the head position:

$$H(\mathbf{r}_{mic}|\mathbf{r}_{mouth}, k) = H(\mathbf{r} = \mathbf{r}_{mouth}|\mathbf{r}_s = \mathbf{r}_{mic}, k).$$

As a result we can simulate the ATF between a mouth on a head, and a single microphone in free space. Repeated use of the algorithm allows for multiple receivers.

Although more accurate modelling of the head and mouth is possible using finite element or boundary element methods [5, 30] for example, the SMIR method is valuable for application to this problem due its comparative simplicity and the fact that, if desired, it can also take into account room reverberation. The SMIR method can, for example, be used as a mouth simulator in the evaluation of a speech enhancement algorithm [13], instead of the omnidirectional source model that is commonly used. While the diameter of the mouth plays an important role in determining the filter characteristic of the vocal tract [8], we assume for the purposes of the scattering model that the mouth is a point source.

As an illustration of this application, Fig. 4.12 shows the energy of the ATF between the mouth and a microphone as a function of microphone position at frequencies of 100 Hz and 3 kHz in an anechoic environment. The mouth was positioned on a sphere of radius 8.75 cm. Only two dimensions, x and y, are shown for brevity since the z dimension is identical to x and y. We observe that at 100 Hz there is no scattering and the radiation pattern is omnidirectional so that the sphere has little effect. At 3 kHz the effect of scattering starts to become more significant, and the energy at the back of the sphere is reduced while the energy at the front is increased. Finally the bright spot discussed in Sect. 4.2.3 is particularly apparent at the very back of the sphere in the bottom plot.
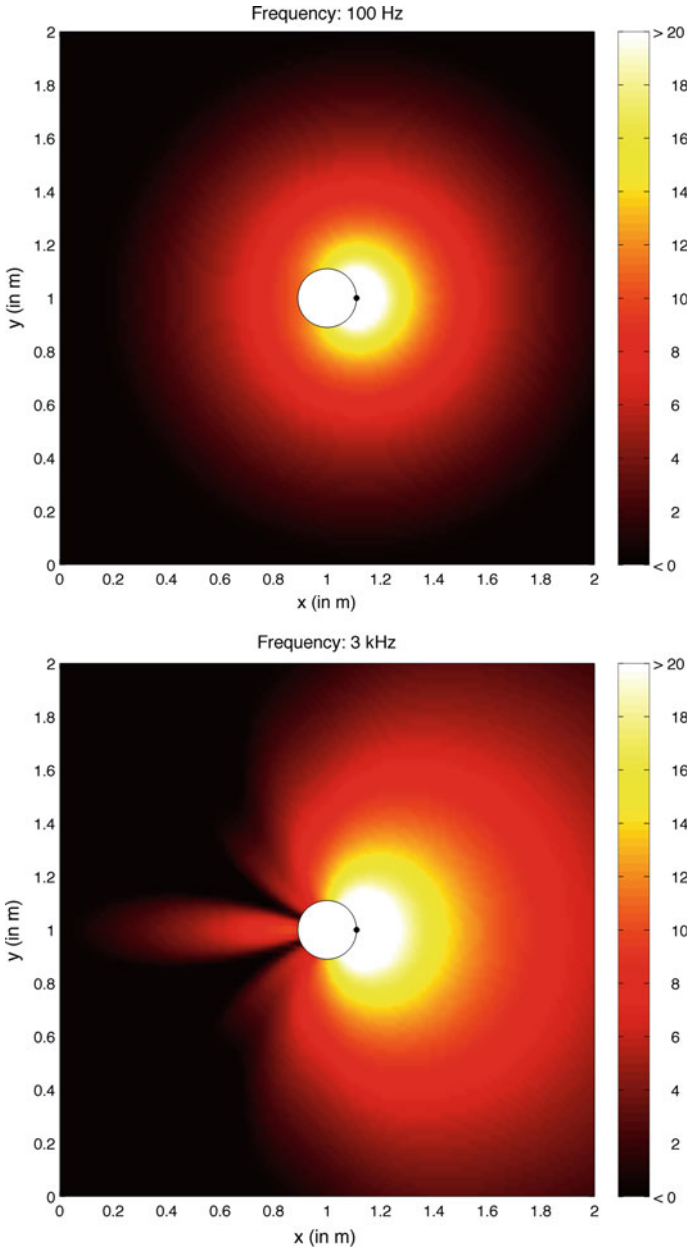
Fig. 4.12 Sound energy radiation pattern (in dB) at 100 Hz (*top*) and 3 kHz (*bottom*). The mouth position is denoted by a *black dot*

4.5 Chapter Summary and Conclusions

Spherical microphone arrays on a rigid baffle are of great interest, due to their numerical robustness and precisely calculable scattering effects. In order to analyze, work with and develop acoustic signal processing algorithms that make use of a spherical microphone array, a simulator is needed that can take into account the effects of the acoustic environment of the array as well as the scattering effects of the rigid spherical baffle. Accordingly, in this chapter the SMIR method was presented for the simulation of AIRs or ATFs for a rigid spherical microphone array in a reverberant environment.

We presented a scattering model used to model the rigid sphere, justifying its use with references to the literature, and provided an overview of the model's behaviour. We showed that the error with respect to the theoretical model can be controlled at the expense of increased computational complexity. Finally we provided a number of examples showing additional applications of this method.

Appendix: Spatial Correlation in a Diffuse Sound Field

The sound pressure at a position $\mathbf{r} = (r, \Omega)$ due to a unit amplitude plane wave incident from direction Ω_s is given by [40]

$$P(\mathbf{r}, \Omega_s, k) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} 4\pi\varphi(\Omega_s) b_l(k) Y_{lm}^*(\Omega_s) Y_{lm}(\Omega), \qquad (4.24)$$

where $\varphi(\Omega_s)$ is a random phase term and $|\varphi(\Omega_s)| = 1$. Assuming a diffuse sound field, the spatial cross-correlation between the sound pressure at two positions $\mathbf{r} = (r, \Omega)$ and $\mathbf{r}' = (r, \Omega')$ is given by:

$$\begin{aligned}
C(\mathbf{r}, \mathbf{r}', k) &= \frac{1}{4\pi} \int_{\Omega_s \in \mathcal{S}^2} P(\mathbf{r}, \Omega_s, k) P^*(\mathbf{r}', \Omega_s, k) d\Omega_s \\
&= \frac{1}{4\pi} \int_{\Omega_s \in \mathcal{S}^2} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} 4\pi b_l(k) Y_{lm}^*(\Omega_s) Y_{lm}(\Omega) \\
&\quad \times \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} 4\pi b_{l'}^*(kr) Y_{l'm'}(\Omega_s) Y_{l'm'}^*(\Omega') d\Omega_s.
\end{aligned}$$

Using the orthonormality property of the spherical harmonics in (2.18) and the addition theorem in (2.23), we eliminate the cross terms followed by the sum over m and obtain

$$C(\mathbf{r}, \mathbf{r}', k) = \frac{1}{4\pi} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (4\pi)^2 |b_l(k)|^2 Y_{lm}(\Omega) Y_{lm}^*(\Omega')$$

$$= \frac{1}{4\pi} \sum_{l=0}^{\infty} (4\pi)^2 |b_l(k)|^2 \frac{2l+1}{4\pi} \mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}'})$$

$$= \sum_{l=0}^{\infty} |b_l(k)|^2 (2l+1) \mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}'}), \tag{4.25}$$

where $\Theta_{\mathbf{r},\mathbf{r}'}$ is the angle between $\mathbf{r}$ and $\mathbf{r}'$.

In the open sphere case, we can derive a simplified expression for $C(\mathbf{r}, \mathbf{r}', k)$. Firstly, we note that the expression in (4.25) is real, and therefore, for a reason which will soon become clear, $C(\mathbf{r}, \mathbf{r}', k)$ can advantageously be expressed as

$$C(\mathbf{r}, \mathbf{r}', k) = -\Im \left\{ -i \sum_{l=0}^{\infty} |b_l(k)|^2 (2l+1) \mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}'}) \right\}, \tag{4.26}$$

where $\Im$ denotes the imaginary part of a complex number. By substituting the open sphere mode strength $b_l(k) = i^l j_l(kr)$ into (4.26), we obtain

$$C(\mathbf{r}, \mathbf{r}', k) = -\Im \left\{ -i \sum_{l=0}^{\infty} j_l^2(kr)(2l+1) \mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}'}) \right\}. \tag{4.27}$$

Using $\Re\{h_l^{(2)}(kr)\} = j_l(kr)$, where $\Re$ denotes the real part of a complex number, we can now write (4.27) as

$$C(\mathbf{r}, \mathbf{r}', k) = -\Im \left\{ -i \sum_{l=0}^{\infty} j_l(kr) \left[h_l^{(2)}(kr) - i \Im\{h_l^{(2)}(kr)\} \right] (2l+1) \mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}'}) \right\}$$

$$= -\Im \left\{ -i \sum_{l=0}^{\infty} j_l(kr) h_l^{(2)}(kr)(2l+1) \mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}'}) \right\}$$

$$+ \Im \left\{ \underbrace{\sum_{l=0}^{\infty} j_l(kr) \Im\{h_l^{(2)}(kr)\}(2l+1) \mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}'})}_{\star} \right\}. \tag{4.28}$$

As the expression marked with a $\star$ is real, its imaginary part is zero and (4.28) can be simplified to

$$C(\mathbf{r}, \mathbf{r}', k) = -\Im \left\{ -i \sum_{l=0}^{\infty} j_l(kr) h_l^{(2)}(kr)(2l+1) \mathcal{P}_l(\cos \Theta_{\mathbf{r},\mathbf{r}'}) \right\}. \tag{4.29}$$

Finally, using (4.7) and (4.8), we obtain the well-known spatial domain result for two omnidirectional receivers in a diffuse sound field [20, 31, 39]:

$$C(\mathbf{r}, \mathbf{r}', k) = -\Im \left\{ \frac{e^{-ik\|\mathbf{r}-\mathbf{r}'\|}}{k\|\mathbf{r}-\mathbf{r}'\|} \right\}$$
$$= \frac{\sin(k\|\mathbf{r}-\mathbf{r}'\|)}{k\|\mathbf{r}-\mathbf{r}'\|}. \tag{4.30}$$

References

1. Abramowitz, M., Stegun, I.A. (eds.): Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover Publications, New York (1972)
2. Allen, J.B., Berkley, D.A.: Image method for efficiently simulating small-room acoustics. J. Acoust. Soc. Am. **65**(4), 943–950 (1979)
3. Avni, A., Rafaely, B.: Sound localization in a sound field represented by spherical harmonics. In: Proceedings 2nd International Symposium on Ambisonics and Spherical Acoustics, pp. 1–5. Paris, France (2010)
4. Betlehem, T., Poletti, M.A.: Sound field reproduction around a scatterer in reverberation. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 89–92 (2009). doi:10.1109/ICASSP.2009.4959527
5. Botteldoore, D.: Finite-difference time-domain simulation of low-frequency room acoustic problems. J. Acoust. Soc. Am. **98**(6), 3302–3308 (1995)
6. Brown, C., Duda, R.: A structural model for binaural sound synthesis. IEEE Trans. Speech Audio Process. **6**(5), 476–488 (1998). doi:10.1109/89.709673
7. Chen, Z., Maher, R.C.: Addressing the discrepancy between measured and modeled impulse responses for small rooms. In: Proceedings of Audio Engineering Society Convention (2007)
8. Deller, J.R., Proakis, J.G., Hansen, J.H.L.: Discrete-Time Processing of Speech Signals. MacMillan, New York (1993)
9. Duda, R.O., Martens, W.L.: Range dependence of the response of a spherical head model. J. Acoust. Soc. Am. **104**(5), 3048–3058 (1998). doi:10.1121/1.423886
10. Evans, M.J., Angus, J.A.S., Tew, A.I.: Analyzing head-related transfer function measurements using surface spherical harmonics. J. Acoust. Soc. Am. **104**(4), 2400–2411 (1998)
11. Gumerov, N., Duraiswami, R.: Modeling the effect of a nearby boundary on the HRTF. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 5, pp. 3337–3340 (2001). doi:10.1109/ICASSP.2001.940373
12. Gumerov, N.A., Duraiswami, R.: Fast Multipole Methods for the Helmholtz Equation in Three Dimensions. Elsevier, Oxford (2005)
13. Habets, E.A.P., Benesty, J.: A two-stage beamforming approach for noise reduction and dereverberation. IEEE Trans. Audio, Speech, Lang. Process. **21**(5), 945–958 (2013)
14. Jarrett, D.P.: Spherical microphone array impulse response (SMIR) generator. http://www.ee.ic.ac.uk/sap/smirgen/
15. Jarrett, D.P., Habets, E.A.P., Thomas, M.R.P., Gaubitch, N.D., Naylor, P.A.: Dereverberation performance of rigid and open spherical microphone arrays: theory and simulation. In: Proceedings Joint Workshop on Hands-Free Speech Communication and Microphone Arrays (HSCMA), pp. 145–150. Edinburgh (2011)
16. Jarrett, D.P., Habets, E.A.P., Thomas, M.R.P., Naylor, P.A.: Simulating room impulse responses for spherical microphone arrays. In: Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), pp. 129–132. Prague, Czech Republic (2011)

17. Jarrett, D.P., Habets, E.A.P., Thomas, M.R.P., Naylor, P.A.: Rigid sphere room impulse response simulation: algorithm and applications. J. Acoust. Soc. Am. **132**(3), 1462–1472 (2012)
18. Knudsen, V., Harris, C.: Acoustical Designing in Architecture. Wiley, New York (1950)
19. Kuhn, G.F.: Model for the interaural time differences in the azimuthal plane. J. Acoust. Soc. Am. **62**(1), 157–167 (1977). doi:10.1121/1.381498
20. Kuttruff, H.: Room Acoustics, 4th edn. Taylor and Francis, London (2000)
21. Lehmann, E., Johansson, A.: Diffuse reverberation model for efficient image-source simulation of room impulse responses. IEEE Trans. Audio Speech Lang. Process. **18**(6), 1429–1439 (2010). doi:10.1109/TASL.2009.2035038
22. Li, Z., Duraiswami, R.: Hemispherical microphone arrays for sound capture and beamforming. In: Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 106–109 (2005)
23. Logan, N.A.: Survey of some early studies of the scattering of plane waves by a sphere. Proc. IEEE **53**(8), 773–785 (1965). doi:10.1109/PROC.1965.4055
24. Meyer, J., Agnello, T.: Spherical microphone array for spatial sound recording. In: Proc. Audio Engineering Society Convention, pp. 1–9. New York (2003)
25. Meyer, J., Elko, G.: A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 2, pp. 1781–1784 (2002)
26. Meyer, J., Elko, G.W.: Position independent close-talking microphone. Signal Process. **86**(6), 1254–1259 (2006). doi:10.1016/j.sigpro.2005.05.036
27. mh acoustics LLC: The Eigenmike microphone array. http://www.mhacoustics.com/mh_acoustics/Eigenmike_microphone_array.html
28. Morse, P.M., Ingard, K.U.: Theoretical Acoustics. International Series in Pure and Applied Physics. McGraw Hill, New York (1968)
29. Peterson, P.M.: Simulating the response of multiple microphones to a single acoustic source in a reverberant room. J. Acoust. Soc. Am. **80**(5), 1527–1529 (1986)
30. Pietrzyk, A.: Computer modeling of the sound field in small rooms. In: Proceedings of AES International Conference on Audio, Acoustics and Small Spaces, vol. 2, pp. 24–31. Copenhagen (1998)
31. Radlović, B.D., Williamson, R., Kennedy, R.: Equalization in an acoustic reverberant environment: robustness results. IEEE Trans. Speech Audio Process. **8**(3), 311–319 (2000)
32. Rafaely, B.: Plane-wave decomposition of the pressure on a sphere by spherical convolution. J. Acoust. Soc. Am. **116**(4), 2149–2157 (2004)
33. Rafaely, B.: Analysis and design of spherical microphone arrays. IEEE Trans. Speech Audio Process. **13**(1), 135–143 (2005). doi:10.1109/TSA.2004.839244
34. Sandel, T.T., Teas, D.C., Feddersen, W.E., Jeffress, L.A.: Localization of sound from single and paired sources. J. Acoust. Soc. Am. **27**(5), 842–852 (1955). doi:10.1121/1.1908052
35. Savitzky, A., Golay, M.J.E.: Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. **36**(8), 1627–1639 (1964). doi:10.1021/ac60214a047
36. Sengupta, D.L.: The sphere. In: J.J. Bowman, T.B.A. Senior, P.L.E. Uslenghi (eds.) Electromagnetic and Acoustic Scattering by Simple Shapes, pp. 353–415. North-Holland (1969) (chap. 10)
37. Shinn-Cunningham, B.G., Kopco, N., Martin, T.J.: Localizing nearby sound sources in a classroom: binaural room impulse responses. J. Acoust. Soc. Am. **117**(5), 3100–3115 (2005). doi:10.1121/1.1872572
38. Talantzis, F., Ward, D.B.: Robustness of multichannel equalization in an acoustic reverberant environment. J. Acoust. Soc. Am. **114**(2), 833–841 (2003)
39. Ward, D.B.: On the performance of acoustic crosstalk cancellation in a reverberant environment. J. Acoust. Soc. Am. **110**, 1195–1198 (2001)
40. Williams, E.G.: Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography, 1st edn. Academic Press, London (1999)