# Spatio-temporal Properties of Simulated Impulse Responses on a Rigid Sphere

Nara Hahn, Frank Schultz, and Sascha Spors
*Institute of Communications Engineering*
*University of Rostock, 18119 Rostock, Germany*
*Email*: nara.hahn@uni-rostock.de

## Introduction

Rigid spherical microphone arrays are widely used due to their compact size, direction-independent spatial resolution, numerical stability and compatibility with the Ambisonic format [1, 2, 3]. A reliable simulation of the microphone array signals enables the evaluation of sound field analysis techniques for varying experimental configurations and conditions. The simulations are typically carried out in the frequency domain by exploiting the analytic expression of the spatial transfer functions [4, 5]. Despite its accuracy in the frequency domain, the simulation is computationally demanding due to the involved spherical Hankel functions and the resulting signals in the time domain exhibit pre-echos and temporal aliasing.

As an alternative, this paper presents a time domain approach where the spatial impulse responses are modeled by means of infinite impulse response (IIR) filters. The polynomial expressions of the spherical Hankel functions are converted either into factorized second-order sections or partial fraction expansions. The digital filter coefficients are then obtained by mapping the parameters from the Laplace domain to the $z$ domain. The influence of different mappings (bilinear transform, matched $z$-transform, and the corrected impulse invariance method) on the spatio-temporal characteristics of the simulated impulse responses are investigated.

## Modal Responses

We consider the sound field of a plane wave driven by a Dirac impulse. The propagation direction is represented by spherical coordinates $\boldsymbol{n}_{\mathrm{pw}} = (1, \theta_{\mathrm{pw}}, \phi_{\mathrm{pw}})$ with $\theta_{\mathrm{pw}}$ and $\phi_{\mathrm{pw}}$ denoting the colatitude and azimuth angles, respectively. As illustrated in Fig. 1, the plane wave impinges on a rigid sphere whose radius is $R$. It is assumed that the plane wave would pass through the origin at $t = 0$, if the sphere was absent. The planar wavefront thus arrives at the spherical surface at $t = -\frac{R}{c}$ where $c$ denotes the speed of sound. The sound field is captured by a pressure sensor located on the sphere $\boldsymbol{x} = (R, \theta, \phi)$ with $\theta$ and $\phi$ denoting respectively the colatitude and azimuth angles. The captured sound field reads [6, Sec. 6.10.3]

$$\mathcal{H}(\omega) = \sum_{n=0}^{\infty} (2n+1) P_n(\cos\Theta) \frac{-i^{-n+1}}{(\frac{\omega}{c}R)^2 \, h_n'(\frac{\omega}{c}R)}, \qquad (1)$$

where $P_n(\cdot)$ denotes the Legendre polynomials and $h_n'(\cdot)$ the derivative of the spherical Hankel functions of the second kind. The angle between $\boldsymbol{n}_{\mathrm{pw}}$ and $\boldsymbol{x}$ is denoted by $\Theta$ (cf. Fig. 1). The angular frequency is denoted by $\omega = 2\pi f$ with $f$ being the temporal frequency. The time harmonic term $e^{i\omega t}$ is omitted for brevity. Without loss of generality, we assume that the speed of sound is $c = 343\,\mathrm{m/s}$.

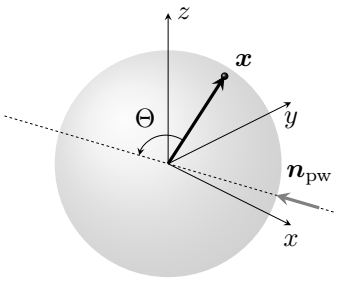



**Figure 1:** In the considered configuration, the sound field of a plane wave is captured on a rigid sphere. The receiver position is denoted by $\boldsymbol{x} = (R, \theta, \phi)$ and the plane wave propagation direction by $\boldsymbol{n}_{\mathrm{pw}} = (1, \theta_{\mathrm{pw}}, \phi_{\mathrm{pw}})$. The angle between the two vectors is denoted by $\Theta$.

The sound field (1) is represented as a superposition of modal responses, with $n$ denoting the modal order. The frequency and radial dependency of the individual modes is described by the so-called radial functions $\frac{-i^{-n+1}}{(\frac{\omega}{c}R)^2 \, h_n(\frac{\omega}{c}R)}$. The spectra of exemplary radial functions are depicted in Fig. 2 (left). They exhibit the same high-frequency decay, whereas the low-frequency slope is proportional to $n$.

In practice, the expansion (1) can be evaluated only up to a finite order ($n \leq N$). The maximum order $N$ is referred to as the spatial bandwidth. Due to the band-pass filter characteristic of the radial functions (cf. Fig. 2), the series expansion (1) can be reasonably approximated by a truncated sum satisfying $N \geq \lceil \frac{\omega}{c}R \rceil$ [7, Ch. 9]. The truncation error can be of course reduced by further increasing $N$ [8][1]. In the remainder, we consider a rigid sphere of radius $R = 0.042\,\mathrm{m}$. The temporal bandwidth of interest is 20 kHz. The numerical simulations are performed with $N = 20$ which exceeds the approximate spatial bandwidth $\lceil \frac{\omega}{c}R \rceil = 16$ by some margin.

## Frequency Domain Sampling

The discrete-time modeling of the sound field is commonly performed by uniformly sampling the spectrum

[1] For the simulation of the captured signals, higher $N$ always leads to a higher accuracy. In sound field analysis using the array signals, however, there is a trade-off between truncation error and spatial aliasing. The modal bandwidth for the analysis has to be chosen based on the spatial distribution of the microphones.

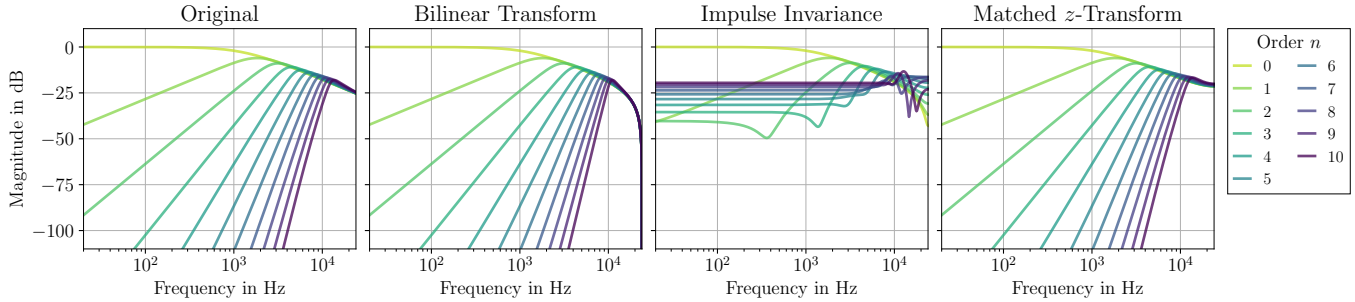



**Figure 2:** Modal spectrum of the radial functions (radius $R = 0.042\,\mathrm{m}$, modal orders $n = 0, 1, \ldots, 10$). The leftmost plot depicts the exact magnitude responses. The other three plots show the results of time-domain simulations using different Laplace-to-$z$ mappings.

of the radial functions [4, 5]. The sampled spectra are linearly combined according to (1). The discrete-time signals are then obtained by the inverse discrete Fourier transform (DFT) of the sum.

We examine the spatio-temporal properties of the impulse responses simulated by the frequency sampling method. The simulations are performed at sampling frequencies of $f_\mathrm{s} = 48\,\mathrm{kHz}$ and $192\,\mathrm{kHz}$. The DFT lengths are 4096 for $f_\mathrm{s} = 48\,\mathrm{kHz}$ and 16 384 for $f_\mathrm{s} = 192\,\mathrm{kHz}$. Prior to the inverse DFT, the spectrum is low-pass filtered by the transfer function of a second-order digital Butterworth filter with cutoff frequency of $20\,\mathrm{kHz}$. Note that the low pass filter's transfer function differs depending on the sampling frequency.

The impulse responses simulated for $f_\mathrm{s} = 48\,\mathrm{kHz}$ are shown in Fig. 3(a). Since the sound field has rotational symmetry with respect to $\boldsymbol{n}_\mathrm{pw}$, the depicted results for $\Theta \in [0°, 360°]$ represent all the impulse responses on the sphere. The spatio-temporal structure shows the effect of the diffraction caused by the rigid sphere. The diffracted wave not only reaches the rear side of the sphere but also travels further along the circumference of the surface. This phenomenon is known as the creeping wave (*Kriechwellen* in German) [9].

The long transients of the impulse responses are apparent, which is mainly due to the time-domain aliasing caused by the sampling in the frequency domain [10, Sec. 11.3.7]. For $f_\mathrm{s} = 48\,\mathrm{kHz}$, the later parts ($t > 1.5\,\mathrm{ms}$) are dominated by temporal artifacts, and the structure of the impulse response is no longer identifiable. The smearing of the impulse response can be reduced by increasing the sampling frequency. Figure 3(b) depicts the simulation result for $f_\mathrm{s} = 192\,\mathrm{kHz}$ (oversampling by a factor of 4). The leading edge of the impulse responses are more clearly seen. The long temporal extent of the creeping waves is also revealed. Although considerably reduced, the temporal artifacts are still present for $t < -\frac{R}{c} \approx -0.1\,\mathrm{ms}$ and $t > 2.3\,\mathrm{ms}$.

Another disadvantage of the frequency domain approach is the high computational complexity. For each modal order $n$, the derivative of the spherical Hankel functions $h_n'(\cdot)$ in (1) has to be evaluated for many frequency bins. The required computations thus scale with the DFT length and the oversampling factor.

## Recursive Filter Design

This section presents the time-domain modeling of the modal impulse responses using recursive (IIR) filters. The presented approach is inspired by the radial filter design proposed in [11].

In order to derive a rational transfer function in the Laplace domain ($s = \sigma + i\omega$), we use the explicit formula of the spherical Hankel functions [12, Eq. (10.49.7)],

$$h_n(\tfrac{R}{c}s) = -i^n e^{-\frac{R}{c}s} \frac{\sum_{k=0}^{n} \beta_n(k)(\frac{R}{c}s)^k}{(\frac{R}{c}s)^{n+1}}, \tag{2}$$

where

$$\beta_n(k) = \frac{(2n-k)!}{(n-k)!\,k!\,2^{n-k}} \tag{3}$$

for $k = 0, \ldots, n$ [12, Eq. (10.49.1)]. The expression for the derivative of the spherical Hankel function can be obtained by using the recurrence relation [12, Eq. (10.51.2)],

$$h_n'(\xi) = -h_{n+1}(\xi) + \tfrac{n}{\xi} \cdot h_n(\xi), \tag{4}$$

yielding

$$h_n'(\tfrac{R}{c}s) = i^{n+1} e^{-\frac{R}{c}s} \frac{\sum_{k=0}^{n+1} \gamma_n(k)(\frac{R}{c}s)^k}{(\frac{R}{c}s)^{n+2}}, \tag{5}$$

where the coefficients $\gamma_n(k)$ are given as

$$\gamma_n(k) = \beta_{n+1}(k) - n \cdot \beta_n(k). \tag{6}$$

Plugging (5) into (1) yields

$$\mathcal{H}(s) = e^{\frac{R}{c}s} \sum_{n=0}^{\infty} (-1)^n (2n+1) P_n(\cos\Theta) \frac{(\frac{R}{c}s)^n}{\sum_{k=0}^{n+1} \gamma_n(k)(\frac{R}{c}s)^k}. \tag{7}$$

For $n \geq 1$, Each mode is characterized by $n+1$ complex poles and $n$th order zeros at $s = 0$. The exponential term $e^{\frac{R}{c}s}$ corresponds to a negative delay (time advance) of $-\frac{R}{c}$. In practice, a positive delay has to be applied to (7) to make the system causal.

The rational part of each mode can be factorized as

$$\frac{(\frac{R}{c}s)^n}{\sum_{k=0}^{n+1}\gamma_n(k)(\frac{R}{c}s)^k} = \frac{(\frac{R}{c}s)^n}{\prod_{l=0}^{n}(\frac{R}{c}s - p_l)} \tag{8}$$

$$= \frac{\frac{c}{R}}{s - \frac{c}{R}p_0} \cdot \prod_{l=1}^{n} \frac{s}{s - \frac{c}{R}p_l}, \tag{9}$$

where $p_l$ denotes the $l$th roots of the polynomial described by the coefficients $\gamma_n(k)$. The corresponding poles in the Laplace domain are located at $s = \frac{c}{R} \cdot p_l$. For the considered spatial bandwidth ($N = 20$) and radius ($R = 0.042\,\mathrm{m}$), the poles can be computed with sufficient numerical stability using `numpy.roots` in Python. The rational part of the modal responses can also be written as a partial fraction expansion,

$$\frac{(\frac{R}{c}s)^n}{\sum_{k=0}^{n+1}\gamma_n(k)(\frac{R}{c}s)^k} = \sum_{m=0}^{n} \frac{r_m}{s - \frac{c}{R}p_m}, \tag{10}$$

where $r_m$ denotes the residues corresponding to the poles $\frac{c}{R} \cdot p_m$. Note that only simple poles exist (i.e. $p_l \neq p_{l'}$, $\forall l \neq l'$). Equation (9) and (10) respectively constitute cascade and parallel systems.

The modal transfer functions, (9) or (10), are realized as digital IIR filters by converting the Laplace-domain representations into $z$-domain representations. In this paper, we consider the bilinear transform, the impulse invariance method, and the matched $z$-transform, which are briefly introduced in the following.

- Bilinear transform: The poles and zeros are converted by the substitution $z = \frac{1+\frac{s \cdot T_\mathrm{s}}{2}}{1-\frac{s \cdot T_\mathrm{s}}{2}}$ [13, Sec. 8.3.3] with $T_\mathrm{s} = \frac{1}{f_\mathrm{s}}$ denoting the sampling interval. The frequency axis ($i\omega$) in the Laplace domain is mapped onto the circumference of the unit circle ($e^{i\omega T_\mathrm{s}}$) in the $z$-plane. The resulting discrete-time system is free of aliasing but the spectrum is warped due to the nonlinear mapping.
- Matched $z$-transform: The poles and zeros are mapped by $z = e^{s \cdot T_\mathrm{s}}$ [13, Sec. 8.3.4]. Since the conversion does not change the pole and zero frequencies, the spectrum is linearly mapped for $|f| < \frac{f_\mathrm{s}}{2}$. The spectral components beyond the Nyquist limit, however, causes aliasing.
- Impulse invariance method: The filter is designed in such a way that the discrete-time impulse response is a sampled version of the continuous-time counterpart [13, Sec. 8.3.2]. Although the time-domain structure is well preserved, the spectrum inevitably exhibits aliasing distortions. In this paper, we use the corrected impulse invariance method [14, 15].

Note that the bilinear transform of a cascaded continuous-time systems is the cascade of their individual bilinear transforms [16, Pr. 7.25]. This is also the case for the matched $z$-transform. The two methods are thus applied to the factorized expression (9). The (corrected) impulse invariance method applied to a parallel continuous-time systems is equivalent to the parallel combination of the impulse invariant discrete-time systems [16, Pr. 7.25]. Therefore, the impulse invariance method is applied to the partial fraction expansion (10).

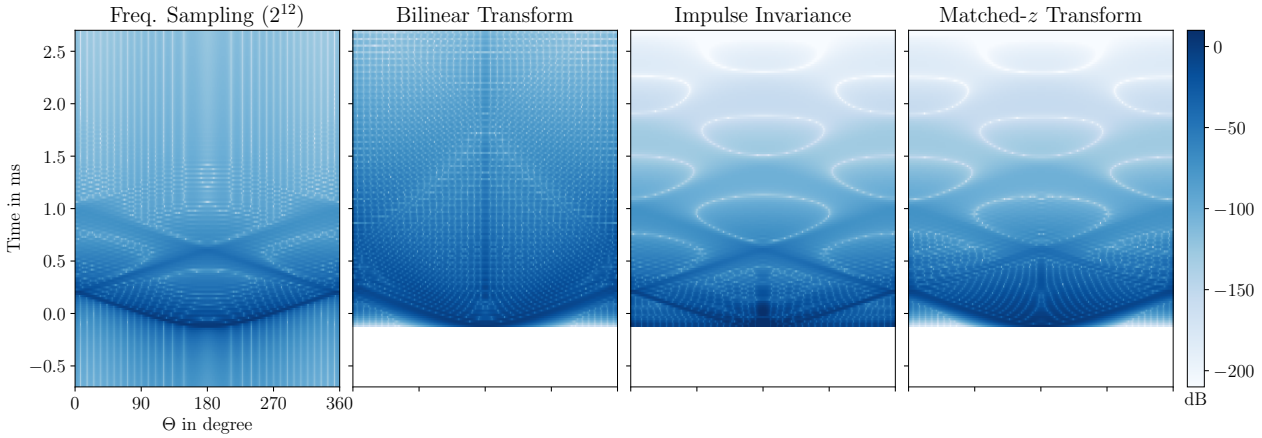
The modal impulse responses are simulated in the time domain by using the above methods. The frequency responses of the digital radial filters are depicted in Fig. 2 ($f_\mathrm{s} = 48\,\mathrm{kHz}$). In the case of the bilinear transform, the spectral deviation increases in the neighborhood of $\frac{f_\mathrm{s}}{2}$. Due to the mentioned frequency warping, the spectrum decays increasingly steeper at higher frequencies. The results for the matched $z$-transform exhibit comparably moderate errors. The impulse invariance method results in strong spectral impairments in comparison with other methods. The higher-order zeros at $\omega$ are not properly modeled, leading to high errors at low frequencies. It can be seen that the spectral deviation increases for higher modal orders $n$.

The spatial impulse responses are computed by combining the radial filters according to the expansion (7) with a finite modal bandwidth ($\infty \rightarrow N$). For a direct comparison with the frequency sampling method, the same low-pass filter (second-order Butterworth with cutoff frequency at 20 kHz) is applied in the time domain. Figure 3(a) shows the spatial impulse responses for $f_\mathrm{s} = 48\,\mathrm{kHz}$. The bilinear transform introduces significant group delay distortion thereby affecting the structure of the impulse responses. The diffraction and the creeping wave are barely identifiable. The matched $z$-transform and the impulse invariance method show similar results in terms of spatio-temporal structure. In particular, the late part of the impulse responses is almost identical, as the IIR filters have the same set of poles.

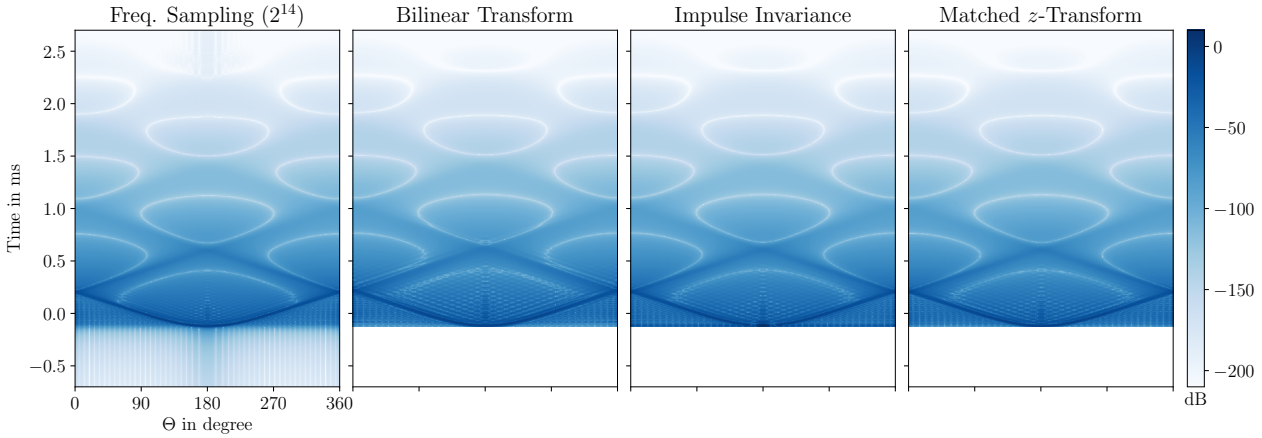
The results for oversampling ($f_\mathrm{s} = 192\,\mathrm{kHz}$) are shown in Fig. 3(b). Increasing the sampling rate reduces frequency warping and aliasing artifacts, which not only improves the spectral accuracy (not shown here) but also the temporal characteristics. The difference between the time-domain simulations is subtle. The result for the frequency-domain sampling is also in a close agreement except for the pre-ringing.

## Conclusion

This paper investigates different frequency and time domain approaches for the simulation of spherical microphone array signals. The frequency domain simulations, based on the inverse DFT of uniformly sampled spectrum, are shown to exhibit temporal artifacts. We present a time-domain simulation where the spatial impulse responses are realized efficiently by recursive filters. The digital filter coefficients are derived based on the analytical representation of the modal responses of the sound field. It is demonstrated that the resulting discrete-time impulse responses exhibit well-defined transients. These advantages, however, come at the expense of impaired spectral accuracy which depends on the transformation of the system function from the Laplace domain to the $z$ domain. It is revealed that, for the considered configuration, the matched $z$-transform is superior in terms




**(a)** $f_\mathrm{s} = 48\,\mathrm{kHz}$




**(b)** $f_\mathrm{s} = 192\,\mathrm{kHz}$

**Figure 3:** Simulated impulse responses on a rigid sphere (radius $R = 0.042\,\mathrm{m}$, spatial bandwidth $N = 20$). All impulse responses are low-pass filtered by a second-order Butterworth filter with a cutoff frequency of 20 kHz. For the frequency sampling method (first column), the DFT lengths are indicated in the parentheses.

of spectral accuracy in comparison to the bilinear transform and the corrected impulse invariance method. We also show that the simulation accuracy can further be improved by mild oversampling.

## References


[1] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *IEEE Int. Conf. Acoust. Speech Signal Processing (ICASSP)*, vol. 2, 2002, pp. 1781–1784.

[2] B. Rafaely, *Fundamentals of Spherical Array Processing*. Cham, Switzerland: Springer, 2015.

[3] F. Zotter and M. Frank, *Ambisonics*. Cham, Switzerland: Springer, 2019.

[4] B. Bernschütz, C. Pörschmann, S. Spors, and S. Weinzierl, "SOFiA sound field analysis toolbox," in *Proc. Int. Conf. Spatial Audio (ICSA)*, Detmold, Germany, 2011, pp. 7–15.

[5] D. Jarrett, E. Habets, M. Thomas, and P. Naylor, "Rigid sphere room impulse response simulation: Algorithm and applications," *J. Acoust. Soc. Am.*, vol. 132, no. 3, pp. 1462–1472, 2012.

[6] E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Cambridge, UK: Academic Press.

[7] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Oxford, UK: Elsevier, 2005.

[8] R. Kennedy, P. Sadeghi, T. Abhayapala, and H. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Trans. Signal Proces.*, vol. 55, no. 6, pp. 2542–2556, 2007.

[9] W. Franz, "Über die Greenschen Funktionen des Zylinders und der Kugel," *Zeitschrift für Naturforschung A*, vol. 9, no. 9, pp. 705–716, 1954.

[10] B. Girod, R. Rabenstein, and A. Stenger, *Signals and Systems*. Chichester, UK: Wiley, 2001.

[11] H. Pomberger, "Angular and radial directivity control for spherical loudspeaker arrays," Master's thesis, University of Music and Dramatic Arts, Graz, Austria, 2008.

[12] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions Handback*. New York, NY, USA: Cambridge Univ. Press, 2010.

[13] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[14] L. B. Jackson, "A correction to impulse invariance," *IEEE Signal Process. Lett.*, vol. 7, no. 10, pp. 273–275, 2000.

[15] W. F. Mecklenbräuker, "Remarks on and correction to the impulse invariant method for the design of IIR digital filters," *Signal Process.*, vol. 80, no. 8, pp. 1687–1690, 2000.

[16] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice Hall, 1999.