FILTER DESIGN FOR REAL-TIME AMBISONICS ENCODING DURING WAVE-BASED ACOUSTIC SIMULATIONS

Iain Henderson[1] **Archontis Politis**[2] **Stefan Bilbao**[3]

[1] Institut de Mathématiques de Toulouse, INSA Toulouse, France
[2] Audio and Speech Processing Research Group, Tampere University, Finland
[3] Acoustics and Audio Group, University of Edinburgh, Edinburgh, United Kingdom

henderso@insa-toulouse.fr, s.bilbao@ed.ac.uk

ABSTRACT

The Ambisonics format is a powerful audio tool designed for spatial encoding of the pressure field. An under-exploited feature of this format is that it can be directly extracted from virtual acoustics simulations. Finite Difference Time Domain (FDTD) simulations are particularly adapted as they simplify greatly the problem of extracting spatially-encoded signals, and enable real-time processing of the simulated pressure field. In this short contribution, we first write a time domain representation of the ambisonic channels, in terms of spatial derivatives of the acoustic field at the receiver location, and formulated as a set of ordinary differential equations. We show that in general, the natural corresponding discrete recursive integration yields a prohibitive polynomial drift in time. We then describe a real-time filtering strategy which stabilizes this numerical integration; in the discrete-time setting of FDTD simulations, this real-time filtering process features very low computational costs, avoiding the latency associated with large convolutions and frequency-domain block processing of previous approaches.

1. INTRODUCTION

Volumetric wave-based acoustic simulation through the finite difference time domain (FDTD) method was proposed in the early 1990s [1–3], and follows from the framework developed in the context of electromagnetic field simulation by Yee [4], and from much earlier work in finite difference methods for the wave equation [5].

The extraction of spatially-encoded signals from such simulations is a major ongoing interest. See, e.g., [6, 7]. In a recent paper [8], a simplified framework for time domain spatial encoding for wave-based simulation has been presented, based on earlier representations of spatial encoding in the context of soundfield recording [9, 10]. One of the major difficulties of this approach is the need for high-order time integrators, leading to potential polynomial drift in spatially-encoded output signals. The focus of this paper is on the design of such time integrators which can be used in a flexible dynamic setting (i.e., avoiding the need for frequency domain processing, and allowing for moving receiver positions over the course of a simulation).

In Section 2, a locally-defined time-domain representation of ambisonic encoding is presented. In Section 3, we describe FDTD methods used to numerically simulate the pressure field and compute spatial derivatives of the latter. In Section 4, we call attention to a pathological behaviour (temporal drift) of the simulated sound field, and describe a viable integrator design to suppress such drifts. In Section 5, we present some simulation results and concluding remarks appear in Section 6.

2. SPATIAL SOUND FIELD ENCODING

In this section, we briefly recall a definition of time-domain ambisonic encoding, as presented in [8], related to results from spatial soundfield encoding in [9]. Such a representation, which requires access to the field locally at the encoding location, is well-suited for implementation in time-domain wave-based virtual acoustics simulation.

2.1 Ambisonic Channels

The time evolution of the acoustic field is assumed satisfied by the free-field wave equation:

$$\Delta p(\mathbf{r}, t) = \frac{1}{c^2} \partial_t^2 p(\mathbf{r}, t) \tag{1}$$

where here, $p(\mathbf{r}, t)$ is the acoustic pressure as a function of spatial coordinate $\mathbf{r} \in \mathbb{R}^3$ and time $t \geq 0$. ∂_t represents partial differentiation with respect to t, and Δ is the 3D Laplacian operator. c is the wave speed.

Taking the temporal Fourier transform of (1), we obtain the free-field Helmholtz equation:

$$\Delta \hat{p}(\mathbf{r}, \omega) = -\frac{\omega^2}{c^2} \hat{p}(\mathbf{r}, \omega) \tag{2}$$

A general solution of (2) can be expressed in terms of an integral of plane waves over all directions

$$\hat{p}(\boldsymbol{r}, \omega) = \int_{S^2} \hat{a}(\boldsymbol{\gamma}, \omega) e^{i \frac{\omega}{c} \boldsymbol{\gamma} \cdot \mathbf{r}} d\Omega \tag{3}$$

where S^2 is the unit sphere of the usual 3 dimensional space, $\boldsymbol{\gamma}$ is a 3-vector of unit length, and $d\Omega$ is the surface differential element. $\hat{a}$ is the complex wave amplitude density in direction $\boldsymbol{\gamma}$. See [8] for further details.

Additionally, recall that the set of real spherical harmonics $Y_{l,m}(\beta,\alpha) = Y_{l,m}(\boldsymbol{\gamma})$, $l = 0,\dots,\infty$ $m = -l,\dots,l$ is an orthonormal basis of the space of (square integrable) functions over the sphere S^2. Examples of spherical harmonics are presented in the appendix.

The wave amplitude density $\hat{a}$ can be decomposed over this basis as:

$$\hat{a}(\boldsymbol{\gamma},\omega) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \hat{a}_{l,m}(\omega) Y_{l,m}(\boldsymbol{\gamma}) \tag{4}$$

The functions $a_{l,m}(t)$, obtained through *inverse* time Fourier transform of $\hat{a}_{l,m}(\omega)$, are called the *ambisonic channels* of the sound field p (definitions may vary up to multiplicative constants according to the author). They encode directional information of the sound field at time t at the coordinate center $\mathbf{r} = \mathbf{0}$.

2.2 Time-domain Encoding

A notable property of the spherical harmonics of l^{th} order $Y_{l,m}(\boldsymbol{\gamma})$, when expressed in terms of Cartesian coordinates $[\gamma_x,\gamma_y,\gamma_z]^T$ over the sphere (i.e. $||\boldsymbol{\gamma}||_2 = 1$) is that they can be uniquely extended as homogeneous polynomials over the space $\mathbb{R}^3$, i.e. for all $\boldsymbol{\gamma} \in \mathbb{R}^3$. We can thus define differential operators $D_{l,m}$ by substituting $[\gamma_x,\gamma_y,\gamma_z]^T$ for $\nabla = [\partial_x,\partial_y,\partial_z]^T$ to obtain:

$$D_{l,m} = Y_{l,m}(\nabla) \tag{5}$$

which may be represented as

$$D_{l,m} = \sum_{\boldsymbol{\xi}\in\mathbb{V}_l} \sigma_{l,m}^{(\boldsymbol{\xi})} \partial_x^{\xi_x}\partial_y^{\xi_y}\partial_z^{\xi_z} \tag{6}$$

where $\mathbb{V}_l$ is the set of non negative integer-valued 3-vectors whose components sum to l and $\{\sigma_{l,m}^{(\boldsymbol{\xi})}\}$ a set of coefficients.

Using the orthonormality and homogeneity property of the spherical harmonics $Y_{l,m}$, one may show (see [8]) that

$$D_{l,m}\hat{p}|_{(\mathbf{0},\omega)} = \Big(\frac{i\omega}{c}\Big)^l \hat{a}_{l,m}(\omega) \tag{7}$$

Taking the inverse time Fourier transform of (7), we obtain

$$\frac{1}{c^l}\Big(\frac{d}{dt}\Big)^l a_{l,m}(t) = D_{l,m}p(\mathbf{0},t) \tag{8}$$

Equations (8) serve as a starting point for implementation in a discrete setting such as the FDTD framework.

3. FINITE DIFFERENCE TIME-DOMAIN METHODS

FDTD simulation methods have been thoroughly studied in the literature; Equations (1) and (8) are approximated directly in the spatiotemporal domain through the use of regular grids in discrete time. For the sake of simplicity, we will only describe here the most basic FDTD scheme.

First, assume a regular Cartesian grid, of spacing X, indexed by integer-valued 3-vectors $\mathbf{q} = [q_x,q_y,q_z]^T$. The grid function $p_{\mathbf{q}}^n$ represents an approximation of $p(\mathbf{r},t)$ at $\mathbf{r} = \mathbf{q}X$ and $t = nT$. A two-step explicit scheme for the wave equation (1) has the form

$$p_{\mathbf{q}}^{n+1} = 2p_{\mathbf{q}}^n - p_{\mathbf{q}}^{n-1} + \lambda^2 \sum_{\boldsymbol{e}_\nu\in\mathbb{B}} (p_{\mathbf{q}+\boldsymbol{e}_\nu}^n - p_{\mathbf{q}}^n) \tag{9}$$

where $\mathbb{B}$ is the set of integer valued 3-vectors of length 1, $\lambda = cT/X$ is the Courant number. According to the Courant-Friedrichs-Lewy (CFL) condition [5], for scheme (9), X and T must be chosen such that $\lambda \le 1/\sqrt{3}$ to ensure numerical stability of (9).

Using the ansatz

$$p_{\mathbf{q}}^n = e^{i(\tilde{\omega}nT + \tilde{\mathbf{k}}\cdot\mathbf{q})} \tag{10}$$

in terms of numerical angular frequency $\tilde{\omega}$ and wavenumber $\tilde{\mathbf{k}}$ leads to the numerical dispersion relation which links $\tilde{\omega}$ and $\tilde{\mathbf{k}} = \tilde{k}\boldsymbol{\gamma}$, $\boldsymbol{\gamma} = [\gamma_x,\gamma_y,\gamma_z]^T$ by :

$$\tilde{\omega} = \frac{2}{T}\sin^{-1}\left(\lambda\sqrt{\sum_{\nu=x,y,z}\sin^2\left(\tilde{k}\gamma_\nu X/2\right)}\right) \tag{11}$$

Note that $\tilde{\omega} \neq c\tilde{k}$, contrary to the continuous dispersion relation. This a purely numerical effect due to discretisation error.

3.1 Spatial Finite Difference Operators

The general differential operators $D_{l,m}$ defined in (5) may be discretized using finite difference grid operators. First, for $\nu = x,y,z$, define the translation operators $s_\nu^\pm$ by $s_\nu^\pm p_{\mathbf{q}} = p_{\mathbf{q}\pm\boldsymbol{e}_\nu}$ and

$$\delta_\nu^+ = \frac{1}{X}(s_\nu^+ - 1), \quad \delta_\nu^- = \frac{1}{X}(1 - s_\nu^-) \tag{12}$$

Setting $\xi = 2M + \alpha$, $\alpha \in \{0,1\}$, one may define

$$\delta_\nu^l = (\delta_\nu^-)^M(\delta_\nu^+)^{M+\alpha} \tag{13}$$

which approximates $(\partial_\nu)^l$. Other designs are possible; we choose this one for the sake of simplicity. One may then approximate the differential operator $D_{l,m}$ by

$$\delta_{l,m} = \sum_{\boldsymbol{\xi}\in\mathbb{V}_l} \sigma_{l,m}^{(\boldsymbol{\xi})} \delta_x^{\xi_x}\delta_y^{\xi_y}\delta_z^{\xi_z} \tag{14}$$

4. INTEGRATOR DESIGN

4.1 Basic Integrator Design

As in the discretization of spatial differential operators, first define time-translation operators $s_t^\pm a^n = a^{n\pm1}$. A first basic approach in integrating (8) consists in discretizing a time derivative d/dt by either

$$\delta_t^+ = \frac{1}{T}(s_t^+ - 1) \text{ or } \delta_t^- = \frac{1}{T}(1 - s_t^-) \tag{15}$$

Setting $l = 2N + \alpha$ with $\alpha \in \{0,1\}$, we may then approximate $(d/dt)^l$ by

$$\delta_{t,l} = (\delta_t^+)^{N+\alpha}(\delta_t^-)^N \tag{16}$$

We then obtain a discrete equivalent of (8) :

$$\frac{1}{c^l}\delta_{t,l}a^n_{l,m} = \delta_{l,m}p^n_{\mathbf{0}} \tag{17}$$

Other discretization designs are possible, we choose this one for the sake of simplicity. (17) can be easily implemented recursively; as an example, (17) becomes for $a^n_{-2,2}$:

$$a^{n+1}_{-2,2} = 2a^n_{-2,2} - a^{n-1}_{-2,2} + \sqrt{15/16\pi}\lambda^2(p^n_{\boldsymbol{e}_1} + p^n_{-\boldsymbol{e}_1} - p^n_{\boldsymbol{e}_2} - p^n_{-\boldsymbol{e}_2}) \tag{18}$$

4.2 Solution Drift

The continuous ambisonic encoding equations (8) are susceptible to solution drift (polynomial, and of order $l-1$ for an lth order signal $a_{l,m}$). This drift is an inherent property of the encoding, and exhibits itself for sources in the near field. Note that standard frequency domain encoding in spatial sound recording also exhibits this effect (see, e.g., [11]). Because FDTD methods essentially capture all acoustic field information (near-field and far-field), the appearance of these drifts is to be expected in FDTD simulations.

Indeed, in general the discrete encoding equations (17) also yield such a polynomial drift in time. If not suppressed, polynomial drift rapidly attains high numerical values which render the obtained ambisonic channels $a_{l,m}(t)$ unexploitable.

As examples, drift illustrations are given in Fig.1 and Fig.2. The source considered is a broad-band Gaussian, with standard deviation 7.5×10^{-5} s. The vector $\boldsymbol{r}_s = \boldsymbol{r}_{source} - \boldsymbol{r}_{receiver}$ describes the gap between source and receiver and $d = ||\boldsymbol{r_s}||_2$ is its length in meters.

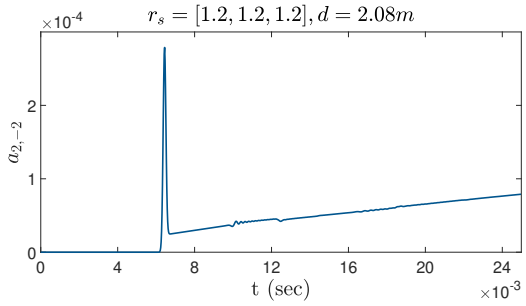

Figure 1: Linear drift in time.

Note however that these drifts are physical: they are present in the real solutions of the physical system. Moreover, simulated drifts such as in Fig.1 and Fig.2 suffer from virtually no numerical artifacts when compared to the physical drifts from equations (8). Indeed, drifts are typically extremely low frequency signals, and the finite difference operators appearing in (17) behave (almost) exactly as their differential operator counterparts from (8) for such low frequency signals. This means that the simulated drifts witnessed in Fig.1 and Fig.2 are not the consequence of numerical effects. Actually, drifts are near-field effects that

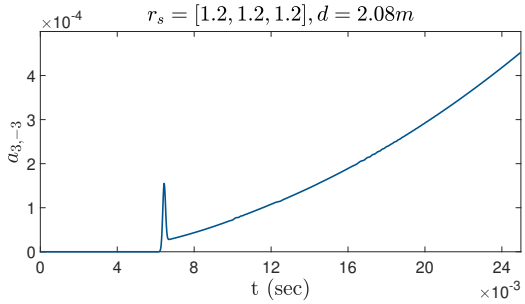

Figure 2: Quadratic drift in time.

manifest themselves as "infinite bass boosts" in the frequency domain, which have already been decribed in [12], in the context of Ambisonics.

4.3 Filtering Strategy

In order to suppress the drifts, we design high-pass Butterworth filters which will stamp out low frequencies. Digital Butterworth filters are causal, IIR filters which can easily be implemented recursively and thus used to filter simulation outputs in real time, that is during the simulation.

We first design them in the s-domain. A high-pass Butterworth filter of order b and cutoff frequency ω_c has a simple gain function :

$$G_b(\omega) = \frac{1}{\sqrt{1 + \left(\frac{\omega_c}{\omega}\right)^{2b}}} \tag{19}$$

The variable $\omega = 2\pi f$ denotes angular frequency. Butterworth filters of order b are maximally flat in the passband (no ripples) and have a roll-off of $20 \times b\ dB/decade$. We then apply a simple integrator of order l to obtain the desired analog integrator

$$I_{b,l}(\omega) = \frac{G_b(\omega)}{\omega^l} = \frac{1}{\omega^l\sqrt{1 + \left(\frac{\omega_c}{\omega}\right)^{2b}}} \tag{20}$$

A Bode plot comparing $I_{b,l}(\omega)$ with a perfect integrator is presented in Figure 3.

Note that (20) imposes that $b > l$ in order to have $I_{b,l}(\omega) \to 0$ when $\omega \to 0$, i.e. to kill the low frequencies. We also recover that $I_{b,l}(\omega) \sim 1/\omega^l$ when $\omega >> \omega_c$ (actually, true for very reasonable ω, see Fig.3) : $I_{b,l}(\omega)$ behaves as an order l integrator for high frequencies.

Using a basic bilinear transform, we may translate (20) from the analog s-domain to the digital z-domain. The resulting IIR filter can be derived analytically and directly implemented to build a recursive filter, that is through an equation of the form :

$$\tilde{a}^{n+1}_{l,m} = \sum_{i \in S_{l,m,b}} c^{(i)}_{l,m,b}\tilde{a}^{n-i}_{l,m} + \delta_{l,m}p^{n+1}_{\boldsymbol{q}} \tag{21}$$

where $\tilde{a}^n_{l,m}$ denotes the desired filtered ambisonic channels, $S_{l,m,b}$ is a known subset of the integers depending only on l, m, b and $\{c^{(i)}_{l,m,b}\}_{i \in S_{l,m,b}}$ a set of coefficients

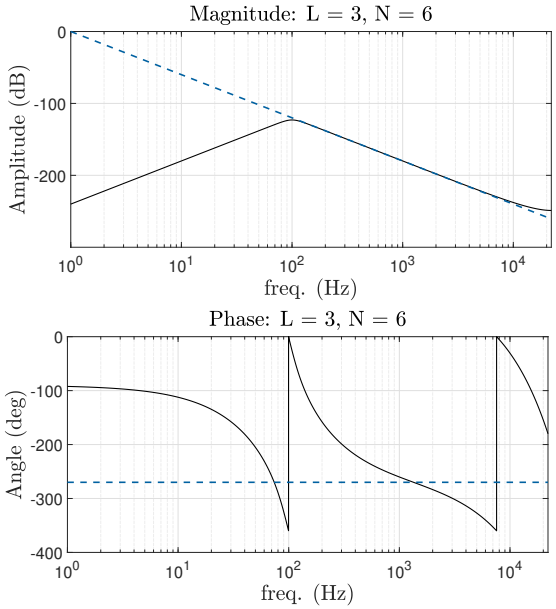

Figure 3: Bode plots : solid line corresponds to $I_{b,l}(\omega)$ for $b = 6, l = 3$, $f_c = 100$ Hz; dashed line corresponds to the Bode plot of a perfect integrator of order $l = 3$.

only depending on l, m, b, which can be derived analytically. Equation (21) yields (very) low computational costs and is easily implementable.

Numerical simulation results show that ω_c in (20) may have to depend on l to yield acceptable results. This is briefly discussed in section (5).

5. SIMULATION RESULTS

In this section, numerical simulation results are presented, with sample rate $1/T = 44100$ Hz, sound speed $c = 344$ m/s and Courant number $\lambda = 1/\sqrt{3}$. Equation (9) was implemented in a free-field setting (transparent boundary conditions), as well as filtered discrete encoding equations (21). Source input is, as in section section (4.2), a broadband Gaussian, with standard deviation 7.5×10^{-5} s.

5.1 Filter Outputs

Filter outputs $\tilde{a}$ are showed here for ambisonic orders $l = 3$ in Fig.4 and $l = 4$ in Fig.5. For well-chosen Butterworth orders b and cutoffs f_c, the drifts are rapidly removed. Fig.3 (phase diagram) shows that, compared to a perfect integrator, a delay appears as b grows for frequencies higher than the cutoff : this is noticeable in Fig.5. This delay should be taken into account when decoding the ambisonic channels, or compensated post-process.

Still, this shows that this filtering strategy is effective and renders possible live Ambisonics outputs from FDTD simulations.

5.2 Discussion on the choice of the cutoff frequencies

A first strategy for choosing the cutoff frequencies f_c would be to set $f_c = 20$ Hz, for all orders. Indeed, the human hearing ranges from 20 Hz to 20 kHz, so virtually no

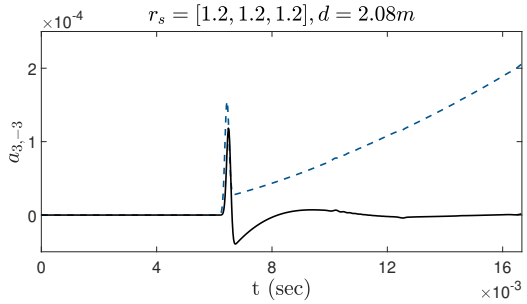

Figure 4: Filtered quadratic drift : $l = 3, b = 6, f_c = 125$ Hz.

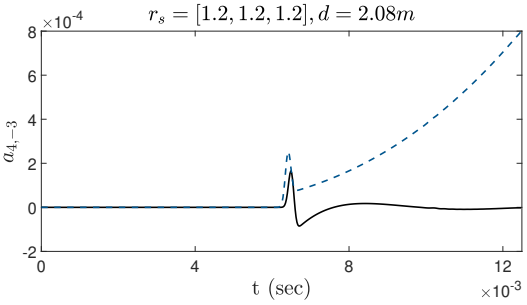

Figure 5: Filtered cubic drift: $l = 4, b = 8, f_c = 150$ Hz

information would be lost. However, simulations show that the chosen cutoff f_c may have to increase as l grows. Fig.6 shows outputs for 7^{th} order ambisonic channel $a_{7,-1}$, with a cutoff $f_c = 25$ Hz. The left hand side figure is a zoom on the right hand side one. One may clearly see the original gaussian bump arriving (left figure). This small bump triggers oscillations of very large amplitude when compared to that of the gaussian impulse (right figure).

One solution would be trying to remove all near-field effects by choosing a high enough cutoff. Nonetheless, near-field effects are desirable since they are key in describing the proximity between source and receiver, and thus should be exploited. Therefore an optimal balance should be sought between removing drifts and depicting source proximity.

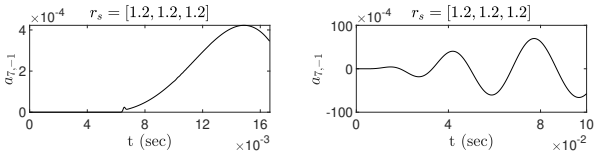

Figure 6: Inappropriate cutoff $f_c = 25$Hz, for ambisonic channel $a_{7,-1}$.

The current strategy is an empirical cutoff frequency selection. This is a viable strategy, though obviously suboptimal and heavy to execute.

6. CONCLUDING REMARKS

The principal issue with extracting ambisonic channels from wave-based simulations is the (physical) drift naturally present in the encoded channels. Here, we pre-

sented a simple filtering strategy which efficiently deals with this drift, thus enabling real-time ambisonic extraction from wave-based simulations. We presented applications up to Ambisonics of order 4, which falls into the domain of High Order Ambisonics (HOA). From perceptual considerations, these are known to render very high spatial quality of reproduction when used with well designed ambisonic decoders. Yet, this strategy can theoretically be used up to any order, but necessitates more precise investigations on the dependency between the cutoff frequencies and the ambisonic orders.

7. APPENDIX

Table 1: $Y_{l,m}(\boldsymbol{\gamma})$ for $l = 0, 1, 2$.

$m \backslash l$	0	1	2
-2	·	·	$\sqrt{15/4\pi}\gamma_x\gamma_y$
-1	·	$\sqrt{3/4\pi}\gamma_y$	$\sqrt{15/4\pi}\gamma_y\gamma_z$
0	$\sqrt{1/4\pi}$	$\sqrt{3/4\pi}\gamma_z$	$\sqrt{5/16\pi}\left(2\gamma_z^2 - \gamma_x^2 - \gamma_y^2\right)$
1	·	$\sqrt{3/4\pi}\gamma_x$	$\sqrt{15/4\pi}\gamma_x\gamma_z$
2	·	·	$\sqrt{15/16\pi}\left(\gamma_x^2 - \gamma_y^2\right)$

8. REFERENCES

[1] D. Botteldooren, "Acoustical finite-difference time-domain simulation in a quasi-Cartesian grid," *J. Acoust. Soc. Am.*, vol. 95, no. 5, pp. 2313–2319, 1994.

[2] L. Savioja, T. Rinne, and T. Takala, "Simulation of room acoustics with a 3-D finite-difference mesh," in *Proc. Int. Comp. Music Conf.*, (Århus, Denmark), pp. 463–466, Sept. 1994.

[3] O. Chiba, T. Kashiwa, H. Shimoda, S. Kagami, and I. Fukai, "Analysis of sound fields in three dimensional space by the time-dependent finite-difference method based on the leap frog algorithm," *J. Acoust. Soc. Jpn.(J)*, vol. 49, pp. 551–562, 1993.

[4] K. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagation*, vol. 14, pp. 302–307, 1966.

[5] R. Courant, K. Friedrichs, and H. Lewy, "Über die partiellen Differenzengleichungen de mathematischen Physik," *Mathematische Annalen*, vol. 100, pp. 32–74, 1928.

[6] A. Southern and D. Murphy, "Methods for 2nd order spherical harmonic spatial encoding in digital waveguide mesh virtual acoustic simulations," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, (New Paltz, New York, USA), pp. 203–206, October 2007.

[7] A. Southern, D. Murphy, and L. Savioja, "Spatial encoding of finite difference time domain acoustic models for auralization," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 20, no. 9, pp. 2420–2432, 2012.

[8] S. Bilbao, A. Politis, and B. Hamilton, "Local time-domain spherical harmonic spatial encoding for wave-based acoustic simulation," *IEEE Signal Processing Letters*, vol. 26, pp. 617–621, 2019.

[9] G. Dickins and R. Kennedy, "Towards optimal soundfield representation," in *106th Convention of the Audio Engineering Society*, (Munich, Germany), May 1999. Paper 4925.

[10] G. Dickins, "Soundfield representation, reconstruction and perception," M.Sc. thesis, The Australian National University, 2003.

[11] B. Rafaely, *Fundamentals of Spherical Array Processing*. Springer, 2015.

[12] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format," in *116th Convention of the Audio Engineering Society*, (Berlin, Germany), May 2003.