Forum Acusticum 2011 - Aalborg - Denmark

The Perceptual Effects of Dispersion Error on Room Acoustic Model Auralization

Alex Southern, Tapio Lokki and Lauri Savioja
Aalto University, Department of Media Technology, School of Science, Espoo, Finland.

Damian Murphy
Audio Lab, Department of Electronics, University of York, York, YO10 5DD, UK

Summary
Acoustic modelling can be informative in the analysis and design of room and concert hall acoustics. There are two types of acoustic models, geometric and wave-based. Geometric models assume sound propagation is ray-like and therefore suited to high frequency room impulse response (RIR) estimation. Wave-based methods are founded on a discrete numerical solution to the wave equation and are therefore a proper treatment of the physical wave motion. Wave-based methods are more appropriate for low frequency RIR estimation and are the focus of this paper. Modelling a continuous system with a discrete approximation leads to well documented limitations, namely dispersion error. It is caused as different frequencies propagate at different wave speeds depending upon their direction. The numerical characteristics of dispersion error are well understood. This paper is concerned with contributing an understanding of the apparent perceptual characteristics attributed to dispersion error. More specifically, it is a preliminary study in identifying the perceptual limits of dispersion error through the analysis of listening tests. A 3D Finite Difference Time Domain scheme is employed to generate the listening test data set. The listening test focuses on identifying the frequency region over which dispersion error becomes perceptually noticeable. The results indicate that the under the chosen model conditions dispersion becomes perceivable at 0.12-0.15 of the model sampling frequency and subsequently provide insight for further research.

PACS no. 43.55.Ka, 43.20.Mv

1. Introduction

The approximation of wave propagation is of interest in many areas of science. In acoustics, the accurate modelling of sound wave propagation within an arbitrarily shaped enclosed volume can be beneficial to acousticians and researchers in the field. Accurate modelling tools will aid the analysis and design processes in the architectural acoustics industry and this is the fundamental motivation behind the work.

Architectural acoustic analysis and design is predominantly subject to an underlying subjective judgement in the form of human perception of auditory events in an enclosed environment. That is, the successful performance of the acoustics in a given situation is ultimately determined by the listeners' perception of quality. Examples may include, the level of speech intelligibility delivered by a PA system in a public building or the auditory experience presented to an audience in a concert hall. Acoustic modelling offers the ability to inform the individual acoustician or researcher about the performance of the enclosure by simulating the characteristics of the sound wave propagation. Furthermore, the modelled data may be made audible using a suitable soundfield rendering system, such as *Spatial Impulse Response Rendering* [1], *Higher Order Ambisonics* or *Wavefield Synthesis*, e.g. [2] and references therein. The process of using soundfield rendering to lead the listener to perceive auditory cues that would exist in the acoustically modelled enclosure is referred to as *auralization* [3].

Current methods for approximating sound wave propagation for synthesizing a room impulse response (RIR) can be categorized into two main areas of study, *wave-based* and *geometric-based*. Briefly, geometric methods include the *image source method*, *beam tracing*, *ray-tracing*, *radiosity* method and *acoustic radiance transfer* all of which model sound propagation using a ray-based assumption [4]. Geometric modelling methods do not inherently model diffraction and indeed low frequency behaviour is in general not treated unless additional methods are introduced e.g. [5].

Wave-based methods approximate wave propagation by dividing the air volume into discrete sections. Sound propagation is simulated for each of the sections through a discrete numerical solution to the wave equation with appropriate boundary conditions. This approach leads to a more accurate representation of wave propagation but at considerable computational expense particularly when the full audible frequency spectrum is of interest. Wave-based methods include, *functional transformation method* [6], *finite element* and *boundary element* methods (FEM)/(BEM) [7], the *digital waveguide mesh* approach (DWM) and related [8] *finite difference time domain* (FDTD) method [9, 10]; which is the focus of this work.

Other than the computational cost wave-based methods have other limitations specific to their individual approach for approximating wave propagation. One such limitation, that is characteristic of the FDTD method, is *dispersion error* which is a consequence of the discrete numerical approximation of a continuous system. The resulting modelled wave propagation exhibits a frequency and directional dependency. While the ability of the FDTD method for acoustical modelling of low frequency behaviour is well established, the perceptual limits of dispersion error have not been considered previously. This aspect of the FDTD model is important when the synthesised RIRs are to be auralized and any audible effects of dispersion error are therefore undesirable.

One possible application of interest of this work is for acoustically modelling concert halls with optimal accuracy over the entire audible bandwidth by combining the synthesized RIRs from different modelling methods resulting in a *hybrid RIR* [11]. The work presented here can be considered in this context if the low frequency acoustic approximation is implemented using a 3D FDTD model and the crossover frequency region can be informed by the region of perceivable difference due to the dispersion error. This work is a preliminary study to identify these perceptual limits of dispersion error.

The structure of the paper is as follows. In Section 2 a brief review of acoustic modelling and an introduction to 3D FDTD technique followed by an introduction to dispersion error. In Section 3 the main contribution of this work is set up and discussed. Section 4 concludes and describes the possibilities for some future work.

2. FDTD Acoustic Modelling

FDTD models are well established for performing acoustical modelling of enclosed geometries [9]. The 2D or 3D wave equation is discretely solved using a second order central finite difference approximation. Finite difference equations for dealing with air, surfaces, edges and corners can be formulated. The update equations for the 3D *standard rectilinear* (SRL) FDTD implementation can be defined explicitly as [10]:

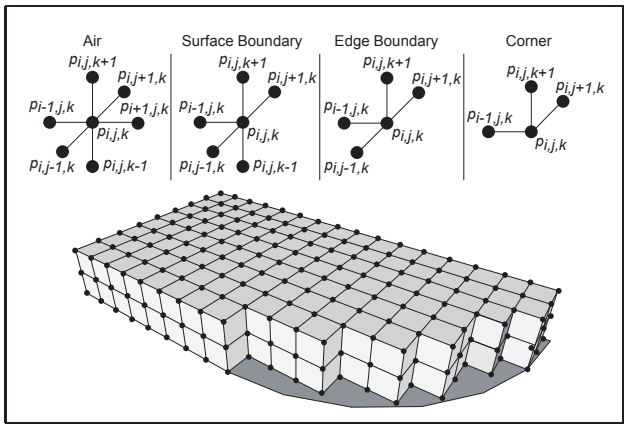

Figure 1. The four node types used in the 3D SRL FDTD scheme with which any complex volume of air may be approximated.

$$p^{n+1}_{i,j,k} = \lambda^2(p^n_{i+1,j,k} + p^n_{i-1,j,k} + p^n_{i,j+1,k} + p^n_{i,j-1,k} + p^n_{i,j,k+1} + p^n_{i,j,k-1}) + 2(1-3\lambda^2)p^n_{i,j,k} - p^{n-1}_{i,j,k} \quad (1)$$

$$p^{n+1}_{i,j,k} = [\lambda^2(2p^n_{i-1,j,k} + p^n_{i,j+1,k} + p^n_{i,j-1,k} + p^n_{i,j,k+1} + p^n_{i,j,k-1}) + 2(1-3\lambda^2)p^n_{i,j,k} + (\alpha\lambda - 1)p^{n-1}_{i,j,k}]/(1+\alpha\lambda) \quad (2)$$

$$p^{n+1}_{i,j,k} = [\lambda^2(2(p^n_{i-1,j,k} + p^n_{i,j,k+1}) + p^n_{i,j+1,k} + p^n_{i,j-1,k}) + 2(1-3\lambda^2)p^n_{i,j,k} + (2\alpha\lambda - 1)p^{n-1}_{i,j,k}]/(1+2\alpha\lambda) \quad (3)$$

$$p^{n+1}_{i,j,k} = [\lambda^2(2(p^n_{i-1,j,k} + p^n_{i,j,k+1} + p^n_{i,j+1,k}) + 2(1-3\lambda^2)p^n_{i,j,k} + (3\alpha\lambda - 1)p^{n-1}_{i,j,k}]/(1+3\alpha\lambda) \quad (4)$$

where (1), (2), (3) and (4) are the 3D air, surface boundary, edge boundary and corner update equations respectively and are depicted in Figure 1. These can be used to acoustically model any complex room geometry and curves are approximated by the grid as illustrated in Figure 1. The term $p^{n+1}_{i,j,k}$ denotes the acoustic pressure at position ijk at time step $n+1$. α is chosen between 0 and 1, and varies the boundary impedance, $\alpha = 1$ nearly anechoic and $\alpha = 0$ totally reflective. Then $\lambda = \frac{cT}{X}$ is the Courant number, X is the unit grid spacing, T the unit time step and for this work $\lambda = 1/\sqrt{D}$ and D is dimensionality so $D = 3$. The metric spacing between the nodes is related to the sampling rate of the grid, $f_s = 1/T$ (5):

$$dx = \frac{c}{\lambda f_s} \approx \frac{595.825}{f_s} \quad (5)$$

(c) European Acoustics Association, ISBN: 978-84-694-1520-7, ISSN: 221-3767

where dx is the distance in metres between two nodes and $c = 344$ is the speed of sound in m/s. The maximum room dimensions that can be practically modelled will vary depending upon the realistic time constraints, chosen sampling rate and the volume of air within the room enclosure. A higher sampling rate results in smaller internodal distance, leading to a denser grid of nodes for which at two pressure values must be stored in memory for time steps $n - 1$ and n. Naturally, higher sampling rates increase the Nyquist frequency resulting in a wider and so more usable frequency bandwidth, but at the cost of reducing the maximum air volume that can be modelled. It is also worth noting that in the 3D FDTD SRL scheme the maximum usable frequency bandwidth of the synthesised RIR is $f_b = 0.196$, which is imposed by the cutoff frequency in the horizontal plane [10]. Larger room dimensions can be modelled if objects are placed within the room, subsequently reducing the size of air volume and computation time. This is in contrast to geometric methods where more internal objects will result in more surfaces and an increase in computation time, although the size of the air volume is not itself a limiting factor.

As outlined in Section 1, a further consideration of the FDTD method arises as a consequence of representing a continuous medium with a discrete implementation, namely dispersion error.

2.1. Dispersion Error

In an ideal continuous propagation medium all waves travel with the same wavespeed in all directions. This is not the case in a FDTD acoustic model due to the imposed restriction of a sampling grid for modelling wave propagation in the air. The consequence of this approximation is that from the perspective of a given position on the sampling grid of nodes a propagating wavefront over the grid can take differing amounts of time to travel the same distance through that point. Subsequently, the wavespeed is directionally dependent and this property leads to dispersion of the wave. The amount of dispersion is also frequency dependent and larger wavelengths are less sensitive to the differences in wavespeed.

The approach taken to analyse the extent of numerical dispersion in previous work has been to obtain a ratio of the discrete numerical wavespeed over the real continuous wavespeed. This ratio is referred to in the literature as the *relative phase velocity* [10]. This relative phase velocity is most easily illustrated when the propagation medium is restricted to only two dimensions for the SRL FDTD scheme as in Figure 2(a).

In 2D it can be observed that the relative phase velocity is always 1 in the diagonal directions of the mesh. That is, there is no difference in the wavespeed of frequencies travelling in the diagonal direction. Deviation is most severe over the axial directions for larger wavenumbers (higher frequencies). Lower frequencies are around the centre of the surface plot, the relative phase velocity deviates comparatively less and is closer to the ideal value of 1 in all directions. In Figure 2(b) the relative phase velocity is shown for the 3D SRL FDTD scheme for the horizontal plane only. Dispersion occurs in all directions to some extent but is again the least severe along the horizontal diagonal of the 3D mesh. It is worth noting that in the 3D SRL scheme there is no dispersion in the diagonal directions that are formed from the 8 straight lines between the central and 8 corner points of a cube.

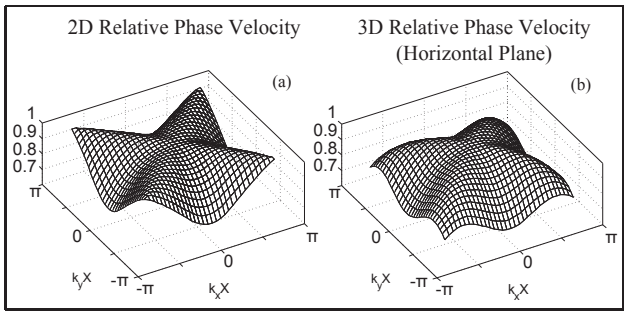

Figure 2. The effects of dispersion. The SRL relative phase velocity with respect to wavenumber in (a) for 2D and in (b) for 3D.

Methods for varying the dispersion error in such models have been considered previously. In the first instance, the dispersion characteristics can be altered by not restricting the wave propagation to the axial directions and using an alternative mesh topology, namely triangular in 2D and tetrahedral in 3D [12]. These cases spatially distribute the dispersion more evenly but at the consequence of increasing the implementation complexity.

Another approach is to introduce additional off-axis propagation paths to the standard rectilinear grid topology to generate new stencils. For illustrative simplicity consider this in the context of a 2D SRL FDTD scheme where we choose 4 additional propagation paths p_d diagonal to the 4 axial cases p_a, as depicted in Figure 3(a). The distance from the central node to those labelled p_a and p_d is not equal. Therefore, the problem with the approach is that the off-axis propagation paths are longer than the axial cases but are still approximated with the same unit time delay. To avoid these distance/time delay discrepancies and alleviate the extent of dispersion error an *interpolated mesh* implementation has been presented previously [13]. The same process is applicable to the 3D 26 neighbour case in Figure 3(b) where all nodes except the 6 along the axial directions must be interpolated in the equivalent manner. A comprehensive treatment of possible rectilinear FDTD stencils is in [10]. The result of employing this technique is that the dispersion error is more evenly distributed and is therefore independent of wave propagation di-

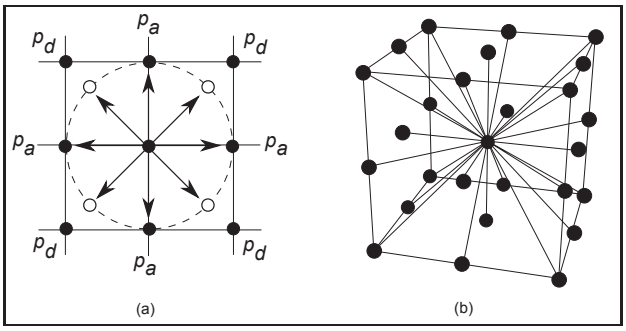

Figure 3. In (a) the 2D interpolated SRL scheme. The white points indicate the intermediate interpolated node positions. (b) A view of the 3D 26 neighbour interpolated FDTD stencil, dispersion error is less severe but at both computational and storage costs.

rection. These improvements are inevitably at a computational cost.

Further reduction of the dispersion error effects can be achieved by introducing *frequency warping* [13]. Briefly, this is achieved by frequency shifting the RIRs obtained from the model with a warped finite impulse response (FIR) filter. The filter coefficients of the FIR filter are set using the samples of the signal to be warped. The IR of the FIR filter is then the desired frequency warped signal.

3. Perceptual Study of Dispersion

An acoustic model of a typical concert hall consists of over tens of millions or hundreds of millions of nodes if the full audible frequency spectrum is to be modelled. This can therefore result in long computation times running to hours. With this in mind this work has restricted this preliminary study to the 3D SRL FDTD scheme without employing any of the techniques for reducing the effects of dispersion error which introduce further computational load. Furthermore, the frequency range of interest is restricted to low frequencies so that the work has practical application to larger concert hall modelling. Therefore with respect to (5), $f_s = 5000$Hz resulting in an internodal distance of $dx = 0.1192$m and a upper frequency limit of 980Hz.

This preliminary investigation on the perception of the inherent dispersion error is restricted to the freefield so that only the direct sound from a single source is considered. In the 3D SRL scheme dispersion does not occur along the diagonal directions connecting the centre of a cube to its corners. However, it is important to note that the dispersion characteristic in the horizontal plane of the 3D SRL scheme in Figure 2(b) is that which is under test. This is logical because the direct soundfield in a typical concert hall will usually pass through this plane. The purpose of this work is therefore to identify under what conditions the effects of the dispersion characteristic in

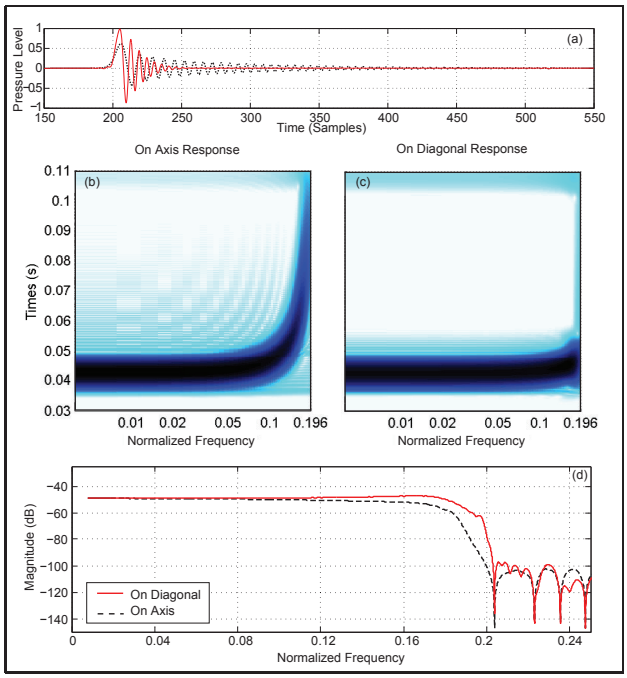

Figure 4. A comparison of the on-axis and diagonal direct sound responses low pass filtered at normalized cutoff of 0.18. (a) in the time domain. Then (b) & (c) in the time-frequency domain for on axis and diagonal respectively. (d) A magnitude response comparison. The differences are due to the dispersion in the 3D SRL FDTD scheme.

Figure 2(b) becomes perceptible by a critical listener.

To conduct a meaningful perception test, it is pertinent to first establish if there are indeed any objective differences in the modelled soundfield due to wave dispersion when considering the direct sound. Comparing the modelled IRs at varying angles of incidence to a fixed position hard source at the same distance proves informative as shown in Figure 4. The measurement distance from the source is set with respect to the number of on-axis nodes. The on-axis and diagonal measurement nodes are then chosen to have the smallest difference in Euclidean distance. In Table I, five such distances are given in terms of distance in nodes along the axis d_a and along the diagonal d_d.

Figure 4 is for d_5. Differences in the direct sound are apparent and there is clearly relatively more spreading of the signal over time in the on-axis case as expected. The spectrograms show the cause of the spreading, the higher frequencies are delayed severely on-axis compared with the diagonal case which also exhibits some dispersion as in Figure 2(b). The magnitude response is also informative, lower frequencies exhibit similar spectral shape which becomes progressively dissimilar at higher frequencies. This characteristic of frequency and directionally dependent differences in the magnitude response have not been previously presented and it is these differences that are under test in this work. The directionally dependent magnitude differ-

(c) European Acoustics Association, ISBN: 978-84-694-1520-7, ISSN: 221-3767

Table I. The on-axis/diagonal distance measurement pairs used in this work. The actual d_d and desired d'_d diagonal measurement locations are approximately the same and exhibit a negligible difference.

Ref.	d_a	$d'_d = d_a/\sqrt{2}$	d_d	$\lvert d'_d - d_d \rvert$
d_1	7	4.95	5	0.051
d_2	34	24.04	24	0.042
d_3	58	41.01	41	0.012
d_4	82	57.98	58	0.017
d_5	99	70.01	70	0.004

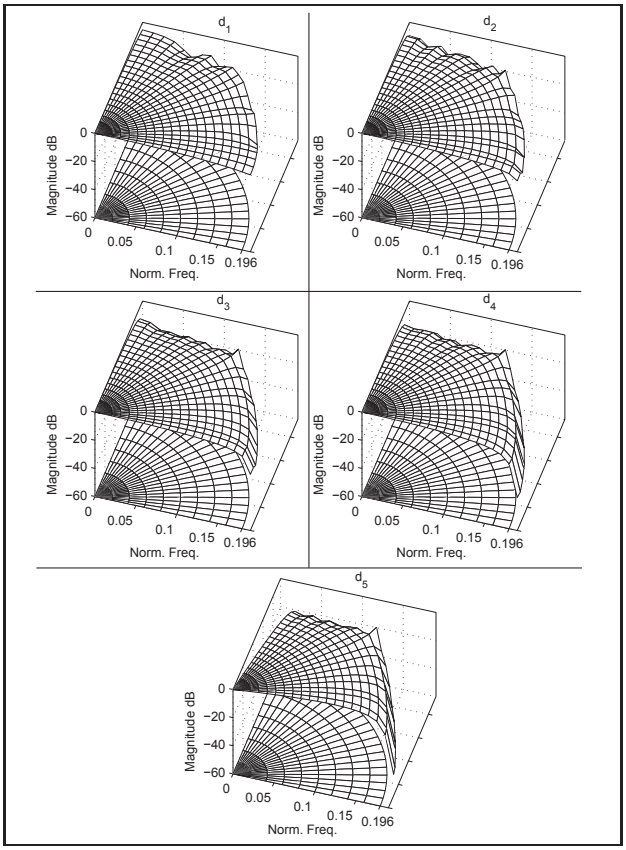

Figure 5. The directionally dependent magnitude response for a range of measurement distances.

ences are best illustrated in Figure 5 for measurement distances as obtained from Table I.

What is not clear is if these differences can be heard by a listener when this IR is appropriately processed and convolved with a typical real source signal.

3.1. Perception Test Description

For the direct sound only, the IR pairs for the on-axis and diagonal measurement distances d_1 to d_5 are synthesized with a 3D SRL FDTD scheme resulting in 10 IRs, or 5 IR pairs. Each IR is then low-pass FIR filtered (LPF) with five cutoff frequencies f_1 to f_5, producing a new set of 50 IRs, or 25 IR pairs. The LPF (MATLAB, 80 tap `fir1`) acts as an anti-aliasing filter and more importantly varies the total objective difference between the magnitude responses

Table II. Correct discriminations in ABX for TRB & VC.

TRB	f_1	f_2	f_3	f_4	f_5
d_1	3	4	⑧	6	⑧
d_2	6	6	5	⑧	⑨
d_3	3	5	4	⑦	⑧
d_4	4	2	5	⑨	⑧
d_5	⑧	3	⑦	⑨	⑨

VC	f_1	f_2	f_3	f_4	f_5
d_1	2	2	3	6	⑦
d_2	4	5	4	6	⑧
d_3	5	⑦	2	⑦	⑧
d_4	4	2	6	3	⑦
d_5	3	5	2	⑦	⑨

of each measurement pair. A set of normalized cutoff frequencies were selected by the authors based on informal listening tests as 0.06, 0.09, 0.12, 0.15, 0.18 being f_1 to f_5 respectively. Each IR was then resampled to 48kHz and convolved with two different anechoic recordings also at 48kHz, resulting in a total of 100 test signals consisting of 50 pairs, 25 of which were for a trombone (TRB) phrase and 25 for a violin cello (VC) phrase, as these are low frequency sources in concert halls.

A null hypothesis is chosen so that there is assumed to be no perceived difference between the 25 pairs in both the TRB and VC instances. The results of a double blind ABX listening test then attempts to disprove this hypothesis for all 25 distance and cutoff frequency combinations for TRB and VC. A total of nine listeners were presented with all 50 pairs labelled A, B and X, with X randomly distributed as either A or B, and each listener then indicated which sample was the odd one out.

3.2. Analysis of Results

The results of the listening tests are given in Table II for the TRB and VC pairs. The circled entries are statistically significant according to a 95% confidence level which is sufficient for such perception tests. The circled results disprove the null hypothesis for those distance and cutoff frequency combinations and they generally agree with the expected trend. The higher cutoff frequencies resulted in the listeners consistently agreeing that a difference was perceptible in the on-axis and diagonal measurement locations, this confidence decreases with the cutoff frequency over the 0.12-0.15 region. f_5 was consistently discriminated despite the measurement distance and this further suggests the notion of a perceptual limit that is below the numerical cutoff limit f_b in the 3D SRL FDTD scheme. It is harder to make clear conclusions about any dependence on measurement distances this can be understood from Figure 6. The pre-LPF on-axis and diagonal IRs for d_1 to d_5 are shown and the attenuation due to distance travelled has been compen-

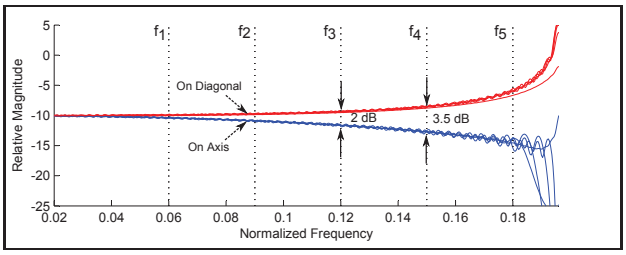

Figure 6. Comparison of the pre-LPF on-axis and diagonal IRs after compensation for attenuation due to distance. The vertical lines represent f_1 to f_5, the crossover region is $f_3 = 0.12$ to $f_4 = 0.15$.

sated to make direct comparison. There are objective differences in the magnitude response, but over the distance of 99 nodes the ABX test suggests they are not perceived reliably, though this may depend on the spectrum of the source stimulus.

4. Conclusions and Future Work

This work has conducted a preliminary study of the perception of numerical wave dispersion that occurs as an inherent consequence of the 3D FDTD method. Objective differences in the magnitude response of the on-axis and diagonal impulse responses to a source are present in addition to the well documented delay in the high frequencies. This work has shown that these magnitude response differences generally lead to perceptible differences in the direct sound measurements prior to any artefacts associated with frequency dependent delay.

Considering only the direct sound has proven informative about the approximate regions of interest, which is useful for when the extension to the reverberant field is made. In the reverberant field it is likely that these magnitude differences will themselves be masked by the stochastic nature of room reflections, but instead the delayed high frequencies will become more apparent as the wave propagation has taken place over a greater distance.

The test model was restricted to $f_s = 5000$Hz and this is a realistic choice if the FDTD model is employed as part of a hybrid acoustic model where higher frequencies are modelled geometrically. When considering the modelled direct sound through the horizontal plane of the 3D SRL FDTD scheme for $f_s = 5000$Hz, the normalized frequency region at which the dispersion becomes perceivable is around 0.12 - 0.15. However it is important to note that a change in f_s does not necessarily mean that the normalized frequency region will remain the same. This is because the spectral characteristics of the dispersion will be positioned differently on the frequency scale. As the human auditory system has a non-linear sensitivity to changes in frequency it is reasonable to assume that this region will not shift in a linear fashion, though the non-linearity itself may turn out to be negligible.

Acknowledgement

This research benefited from funding from the Academy of Finland, project no. [218238] and the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. [203636].

References

[1] V. Pulkki and J. Merimaa. Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests. *J. Audio Eng. Soc*, 54(1/2):3–20, 2006.

[2] J. Daniel, R. Nicol, and S. Moreau. Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging. In *AES Convention 114, Paper 5788, Amsterdam, Netherlands*, 2003.

[3] M. Kleiner, B.I. Dalenbäck, and P. Svensson. Auralisation - An Overview. *J. Audio Eng. Soc.*, 41(11):861–875, 1993.

[4] S. Siltanen, T. Lokki, S. Kiminki, and L. Savioja. The room acoustic rendering equation. *J. Acoust. Soc. of Am.*, 122(3):1624–1635, 2007.

[5] U.P. Svensson, R.I. Fred, and J. Vanderkooy. An analytic secondary source model of edge diffraction impulse responses. *J. Acoust. Soc. of Am.*, 106(5):2331–2344, 1999.

[6] S. Petrausch and R. Rabenstein. Efficient 3D simulation of wave propagation with the functional transformation method. In *In Proc. of 18th Symposium of Simulation Technique, ASIM, Erlangen, Germany*, Sept. 2005.

[7] Andrzej Pietrzyk. Computer modeling of the sound field in small rooms. In *AES Conference: 15th International Conference: Audio, Acoustics and Small Spaces*, Oct 1998.

[8] D. T. Murphy, A. Kelloniemi, J. Mullen, and S. Shelley. Acoustic modeling using the digital waveguide mesh. *IEEE Signal Processing Magazine*, 24(2):55–66, 2007.

[9] D. Botteldooren. Finite difference time domain simulation of low frequency room acoustic problems. *J. Acoust. Soc. of Am.*, 98(6):3302–3308, 1995.

[10] K. Kowalczyk and M. van Walstijn. Room acoustics simulation using 3-D compact explicit FDTD schemes. *Audio, Speech, and Language Processing, IEEE Transactions on*, 19(1):34 –46, 2011.

[11] A. Southern, S. Siltanen, and L. Savioja. Spatial room impulse responses with a hybrid modelling method. In *AES Convention 130, Paper 8385, London, UK*, May 2011.

[12] S.A. Van Duyne and J.O. III. Smith. The tetrahedral digital waveguide mesh. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 1995.*, pages 234 –237, Oct 1995.

[13] L. Savioja and V. Välimäki. Interpolated rectangular 3-D digital waveguide mesh algorithms with frequency warping, 2003.