

# The Impact of Reverberation Techniques on Immersion in Spatial Audio for Virtual Reality

---

Simon Durbridge

November 22, 2016

## 1 ABSTRACT

## 2 INTRODUCTION

Increases in available computing power, and great strides in research and development have brought a new surge of interest to virtual reality (VR) applications. With the introduction of improved VR systems, such as Google Jump, Oculus, Vive and facebook360, using both high end computers, and mobile phone technology to provide immersive visualisation; Further steps are being taken to provide users with immersive sound environments [6]. These environments may be created in an attempt to emulate real places, or to characterise fictional places.

A significant part of how humans identify with their surroundings, includes the perception of the reverberant characteristics of the surrounding area [8]. Cues such as the timing and strength of early reflections help humans to localize sound sources, as well as how far from the nearest boundaries the perceivers are. Some VR application development platforms provide a simplified model for how reverberation behaves in an audio system (negating early reflections), and may be an oversimplification when attempting to create an immersive audio experience that is tending towards the suspension of disbelief<sup>1</sup>.

The aim of this report is to introduce a testing method, to allow for the evaluation of different reverberation methods with respect to immersive VR audio. Initially, some theory behind reverberation and perception will be discussed. Following this, a brief description of current spatial audio techniques will be given. Finally, a testing framework will be proposed, in which subjects will evaluate different VR environments and reverb algorithms.

---

<sup>1</sup>This report is not intended to explore or discuss the artistic concept of the suspension of disbelief. For a discussion on this, please refer to CITE

### 3 REVERBERATION & PERCEPTION

#### 3.1 REVERBERATION

Reverberation (in a simplistic description), is the steady-state of diffusely scattered sound energy in a space due to the reflection of that energy from boundaries to a high order. Specifically, the amplitude of these reflections are such as to balance in amplitude with the steady state of the acoustic system, at or beyond the critical distance from a source (a classic analogy is given by Everest) [2].

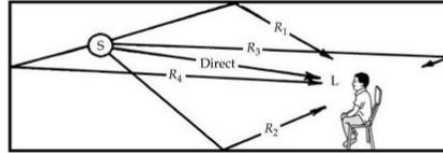


Figure 3.1: A conceptualisation of direct and reflected sound [2]

That is in contrast with strong reflections that may occur before excitation of the reverberant field (early reflections), or may occur later and above the steady state amplitude (echos) [2]. Early and strong reflections are of significant interest in this study, due to the cues humans receive from perception of them. A reverberant sound field is often quantified by the decay time from steady state, to an amplitude of the steady state level  $-60\text{dB}$ . This is commonly described as the  $RT_{60}$ , and was proposed by WC Sabine in 1900. A more comprehensive description of reverberation and overview of the associated parameters is given by Rossing [7]. Further to this, the ratio of direct sound from a source to the amplitude of the reverberant field at any place in the sound field is also of interest [1]. A third parameter of interest is the clarity score, that is determined by the ratio between direct sound and an integration of the sound received directly from the source for some arbitrary time (commonly 80ms).

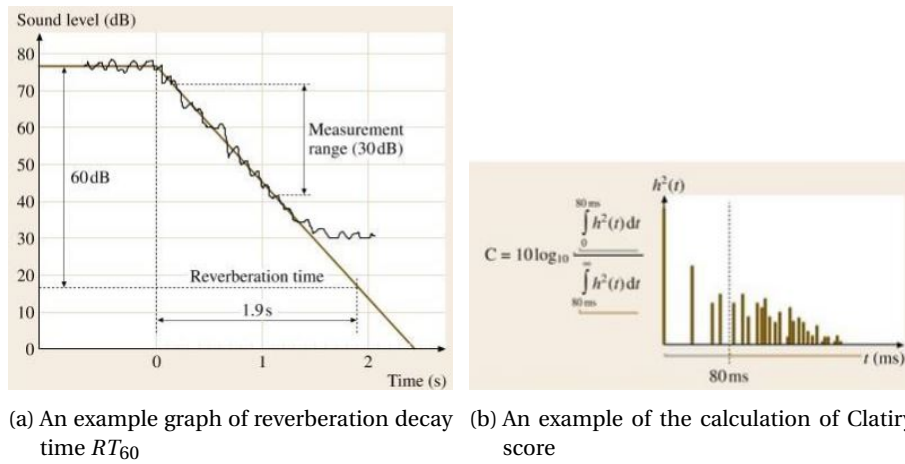


Figure 3.2: Graphics pertaining to quantification of reverberant sound [7]

### 3.2 PERCEPTION OF REFLECTIONS

There is a plethora of research and understanding of the objective quantification of reverberation [7], less so about the subjective effects [4] beyond the seminal studies of early reflections and level threshold shifts [3]. It may be argued that many facets of human auditory perception relate in some way to the perception of reverberation, the concepts of interest in this paper are those linked with localisation and perception of the auditory scene. Within these concepts are Haas/Precedence effect, and auditory masking.

The Haas effect is described as the first wave-front perceived by a listener, determines the source localisation in an auditory scene. The following reflections received within a 0.7 to 35ms window are integrated with the initial sound, such that the sound is given the impression of increased loudness, an increase in source width and tonal shifts of the perceived sound [2]. Beyond 40ms, following strong reflections are heard as echos. Begault [1] suggests that the precedence effect is the intrinsic mechanism of the human auditory system, that allows for the localization of sound sources on a reverberant sound-field. Begault also suggests that there is a link between inter-aural time difference (ITD) perception for lateralization, and changes in apparent sound source 'width', suggesting a 5us to 1.5ms ITD 'window' within which a sound sources lateral location is determined. Another aspect of note, is that if a sound source is occluded and the reflected sound is louder than the direct sound, the precedence effect is diminished [9].

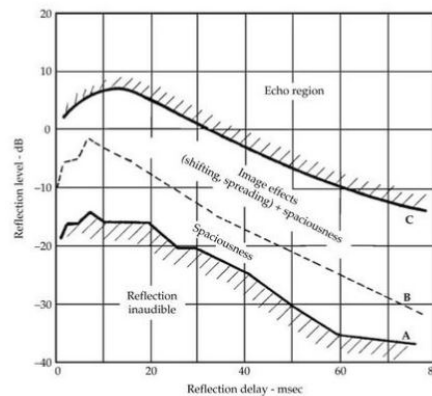


Figure 3.3: A plot of the window regions of precedence between reflection 'level' and delay [2]

Sigfried Linkwitz [5] also discussed the idea that human perception of sound in enclosed spaces is powerful enough that given an appropriate sound source, a person can subconsciously differentiate between reflected sounds and the stereo sound-stage created by two sound sources. The link is with auditory scene analysis. The argument is that more realistic reverberation modelling may provide a more realistic scene for our brains to analyse.

Many of the texts referenced in this section of the report do not discuss auditory masking in great detail<sup>2</sup>, though it may be of considerable importance when considering perception of noise-like sound. Auditory masking can be conceptualised as the time-envelope dependent blocking of new sounds being perceived, due to some other sound having already excited the inner ear. Specifically, sounds

<sup>2</sup>Masking is noted by Begault, but not discussed

having already excited a portion of the basilar membrane are thus stopping other sounds from being perceived as separate, that would otherwise excite the same portion of the basilar membrane [2]. The basilar membrane is often regarded as to be discretized into critical bands or equivalent rectangular bandwidths (ERB), and these bandwidths are different depending on sound level as well as centre frequency. General works by Moore [?] should be reviewed for a more in-depth discussion of masking. The relevance of masking in this context, is that continuous noise is often used as a test signal in masking experiments, and steady-state reverberation may be similar to noise and thus cause masking of reflections and general reverberation. As such, creation of realistic reverberation may be important for masking in a natural way as opposed to having a spectrally dense reverb that may cause excessive masking.

Echolocation

Timing & Intensity

Concert Hall Reverberation

Apparent Source Width

Auditory Spaciousness

### 3.3 REVERBERATION ALGORITHMS

Derived Reverberation

Geometric Modelling

Wave Modelling

## 4 AMBISONIC SPATIAL AUDIO FOR VIRTUAL REALITY

### 4.0.1 AUDIO SPATIAL PERCEPTION

Spatial Cues  
ITD

ILD

HRTFS

Direct to Reverberant Ratio

Precedence

Auditory Scene Analysis

#### 4.0.2 AMBISONICS

the concept of ambisonics

## 5 EXPERIMENT PROTOCOL

Users will be asked to evaluate 5 different instances of VR environments, and determine which audio is from the real environment.

Karjalainen *et al* [4] undertook a study into the perception of late reverberation, using a perceptual model to reinforce listening test data.

## 6 LISTS

### 6.1 EXAMPLE OF LIST (3\*ITEMIZE)

- First item in a list
  - First item in a list
    - \* First item in a list
    - \* Second item in a list
  - Second item in a list
- Second item in a list

## 6.2 EXAMPLE OF LIST (ENUMERATE)

1. First item in a list
2. Second item in a list
3. Third item in a list

$$\begin{aligned}(x+y)^3 &= (x+y)^2(x+y) \\ &= (x^2+2xy+y^2)(x+y) \\ &= (x^3+2x^2y+xy^2)+(x^2y+2xy^2+y^3) \\ &= x^3+3x^2y+3xy^2+y^3\end{aligned}\tag{6.1}$$

$$A = \begin{bmatrix} A_{11} & A_{21} \\ A_{21} & A_{22} \end{bmatrix}\tag{6.2}$$

$$RT_{60} = \frac{0.161A}{Sa}$$

## 7 REFERENCES

### REFERENCES

- [1] Durand R Begault. 3-d sound for virtual reality and multimedia. *Computer Music Journal*, 19(April):99, 1995.
- [2] F Alton Everest and Neil A. Shaw. *Master Handbook of Acoustics, 5th Edition*, volume 110. focal press, 5th edition, 2009.
- [3] Helmut Haas. The Influence of a Single Echo on the Audibility of Speech. *J. Audio Eng. Soc.*, 20(2):146–159, 1972.
- [4] M Karjalainen and H Jarvelainen. More about this reverberation science: Perceptually good late reverberation. *Audio Engineering Society Convention 111*, pages 1–8, 2001.
- [5] Siegfried Linkwitz. THE MAGIC IN 2-CHANNEL SOUND REPRODUCTION - WHY IS IT SO RARELY HEARD ? *Reproduced Sound 2015*, pages 1–22, 2015.
- [6] Oculus. Oculus connect: Introduction to audio in vr - youtube, October 2014. (Accessed on 11/21/2016).
- [7] T. Rossing. *Springer Handbook of Acoustics*. Springer Handbook of Acoustics. Springer New York, 2007.
- [8] F Rumsey. *Spatial Audio*. Taylor & Francis, 2012.
- [9] Bruce Wiggins. *An investigation into the real-time manipulation and control of three-dimensional sound fields*. PhD thesis, University of Derby, 2004.