Perceptual Consequences of Direction and Level of Early Reflections in a Chamber Music Hall

Otavio Colella Gomes[1], Nils Meyer-Kahlen[2], Winfried Lachenmayr[1], Tapio Lokki[2]

[1] *Müller-BBM Gmbh, Helmut-A.-Mueller-Str. 1-5, 82152 Planegg/Munich, Germany, Email: otavio.gomes@mbbm.com*

[2] *Acoustics Lab, Dept. of Signal Processing and Acoustics, Aalto University, P.O. Box 13100, 00076 Aalto, Finland*

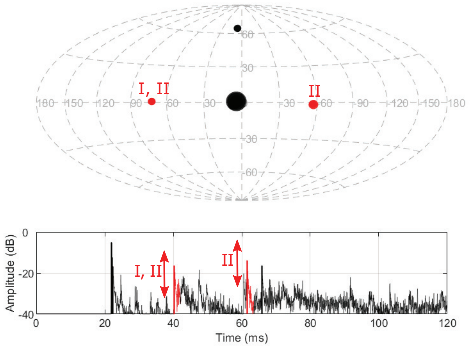
Introduction

Early reflections are important contributors to the sound of a concert hall. They are directly associated with clarity and spatial impression (SI). Early lateral reflections, especially, have a significant role in the perception of apparent source width (ASW) and in how the hall "wakes up" at louder orchestral dynamics [1]. The direction distribution of such reflections is found to be important for the acoustical quality as early as 1964 [2]. Naturally, the relation between the level and direction of such reflections and the perceptual consequences are of utmost interest for concert hall design. As central element at establishing such relations, the audibility threshold of changes to individual reflections should be determined.

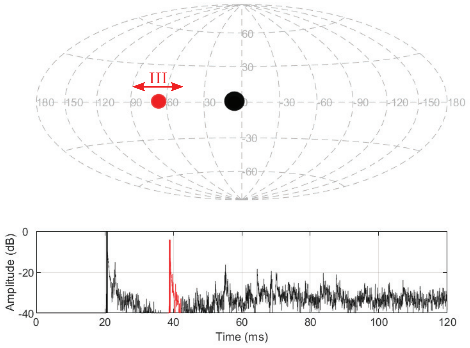
In the literature, several classic experiments concerning the audibility of reflections under controlled conditions are described, for example by Burgtorf, Seraphim and Schubert, which are summarized by [4] and [5]. In different experiments, they tested level thresholds depending on time and direction in anechoic environments, using different signals. Under these conditions, the threshold level depends linearly on the reflection time. In addition, lateral reflections have lower thresholds. As an example, the threshold of a single lateral reflection at 50 ms delay that is incident from 60° was determined to be about as much as 39 dB below the direct sound for continuous speech. Seraphim also studied multiple reflections at regular intervals and found that for four reflections with a spacing of about 18 ms, the threshold remains as high as 10 dB, until the last reflection has passed, after which it reduces with time again. Experiments with multiple reflections are more similar to listening in real concert halls, just as those used in the present test.

Olive and Toole have demonstrated that results from tests in anechoic rooms and using anechoic signals are not directly applicable to real world scenarios [5]. Using continuous sounds, adding slight reverberation or conducting the test in a listening room elevated the thresholds, and the strong time dependency found for discontinuous sounds in dry conditions is reduced. Conducting a test in a normal room with a music stimulus, reflection thresholds as high as 3 dB below the direct sound were found for a delay of 20 ms. They also determined the so called image shift threshold, i.e. the level at which the reflection is not only audible, but causes an image shift (change of source location).

Barron tested the effect of early reflections on SI as a function of direction [6] and he found that it peaks at 90 degrees. Barron also points out the dependence on absolute listening level. In [7], Barron displays the subjective effects of a single side reflections, and points out the SI and image shift changes for different levels and delays.



(a) MZW. Experiments I and II, modification in the level of reflections.



(b) HD. Experiment III, modification to the direction of the left lateral reflection.

Figure 1: Spatial Analysis of the measured SRIR of MZW and HD. The map shows the direction of the components. Below, the compound response according to Eq. (5).

Green summarizes audibility tests made for early reflections in synthetic environments [1]. Wettschurek [3] added an additional artificial reflection and artificial reverberation generated by a plate reverb via four loudspeakers. While one reflection was added at 30 ms at a fixed angle, the tested reflection was added at 70 ms, either in front, in the back or on the side. He repeated his experiments at different listening levels, showing that at lower levels, reflections are harder to detect, but that detectable thresholds saturate at higher levels. These early experiments were conducted with speech, and [1] also shows results for music, which exhibit the same dependence on level and direction. However, the overall thresholds are much higher for music.

Here, we wish to take such experiments further by performing threshold tests by modifying auralizations of existing concert halls. In the first two experiments, the level of either one lateral, or both lateral first-order reflections was modified. In a third experiment, the direction of a lateral reflection was changed. The tools for this are provided by parametric spatial room impulse response processing, which allows us to detect, isolate and modify sound from specific directions, before rendering to a surrounding loudspeaker array. A similar approach was taken in [8], where hybrid auralizations of different halls were created, and recently by Müller [9], who tested the audibility of elevated reflections.

The present study allows comparison of found level thresholds to studies using less realistic conditions.

Method

The method for testing reflection level and direction thresholds under realistic conditions was based on processed spatial room impulse response (SRIR) measurements from two concert halls.

Measurements

For the experiments, SRIR from two different chamber music halls are used. Mozartsaal (MZW), in Vienna, is a shoebox-shaped hall with moderately high SI. Haydnsaal (HD), in Eisenstadt, is also a shoebox hall, with a strong lateral reflection from the left wall at the present measurement position.

For the measurements, the seats of the audience were covered with cloth simulating an occupied condition. A multichannel dodecahedral loudspeaker comprising of 12 individually controlled drivers was used. For the measurements selected in this test, the source was positioned on the center-left side of the stage, and the receiver array, a GRAS microphone array VI-50 was placed 7 m away from the source in the audience area. The microphone array has six channels, meaning that we obtain one room impulse response vector $\boldsymbol{h}_d(t)$ with six microphone responses for each of the source drivers $d \in \{1, ..., 12\}$.

Spatial Analysis and Rendering

For identification of the reflections, spatial analysis using Time-Difference of Arrival estimation (TDoA) was employed, as it is implemented in the SDM toolbox [13]. Directional analysis returns one direction estimate $\boldsymbol{\theta}(t)$ for each sample t, obtained by considering the response in a small time window. For processing, also an omnidirectional response $p(t)$ is required, which is extracted by using the response of the fifth microphone, the one at the top of the intensity array. Upon rendering, loudspeaker impulse responses $\boldsymbol{g}_l$ for loudspeaker at directions $\boldsymbol{\theta}_l$ are created by assigning each sample of the omnidirectional response to the nearest loudspeaker based on the instantaneous directional estimates

$$l_{\mathrm{NLS}}(t) = \arg\min_{l} ||\boldsymbol{\theta}(t) - \boldsymbol{\theta}_l(t)|| \tag{1}$$

$$g_l(t) = \begin{cases} p(t) & l = l_{\mathrm{NLS}}(t) \\ 0 & l \neq l_{\mathrm{NLS}}(t). \end{cases} \tag{2}$$

As assignment of samples in this way results in increased high frequency content, equalization to match the the original omnidirectional response was applied. As opposed to the equalization found in [13], frame-wise minimum-phase equalization filters were derived to avoid time-domain artefacts caused by the equalization.

Since in total, 12 responses are available, one for each driver, pointing to a different direction in the hall, a procedure similar to the reflection path tracing method described in [10] was employed. It enhances the directional analysis, making it easier to find the time and direction of particular reflections. For that, the directional estimates are computed for each driver separately, and the enhanced directional estimates $\hat{\boldsymbol{\theta}}(t)$ are then obtained by selecting the measurement conducted using the driver $\hat{d}$ that has the largest power

$$\hat{d}(t) = \arg\max_{l} \; p_d^2(t) \tag{3}$$

$$\hat{\boldsymbol{\theta}}(t) = \boldsymbol{\theta}_{\hat{d}(t)}(t) \tag{4}$$

$$\hat{p}(t) = p_{\hat{d}(t)}(t). \tag{5}$$

The enhanced estimates and the compound response $\hat{p}$ that uses the driver with maximal power for each sample are shown in Fig. 1. This procedure increases the salience of every reflection in the response and the directional estimates, and makes it easier to select time regions corresponding to them.

Modifications

With the directional estimates at hand, 2 ms long windows containing the target reflections were selected. These regions are marked in red in Fig. 1.

The threshold test consisted of three experiments, each representing a different modification:

I. Level of the left wall reflection in MZW

II. Level of the left and the right wall reflection in MZW

III. Direction of the left wall reflection in HD

The level modifications for test I and II were implemented by changing the level of the reflection in the corresponding time windows before rendering to loudspeakers, see Fig. 1a. The directional modification for test III was realized by modifying the rendering stage, see Fig. 1b. While the remainder of the response, without the reflection of interest, was rendered according to Eq. (2), movement of the target reflection was realized by means of panning. For this, the pair of loudspeaker containing the desired location was determined, and the reflection was panned between them fulfilling the tangent law, equivalent to 2D VBAP panning.

As in previous studies, participants should not be able to base the detection on the changed overall level. Therefore, normalization was conducted by assuming coherent summation of the signals, as this corresponded more closely to the measured increase of level in the reproduction room when changing the reflection level than power summation.The overall measured level in the center of the array for all experiments was 72 ± 1 dBA verified using a B&K 2250 sound level meter.

Listening Test Design

The test was conducted as an adaptive threshold test with fixed step-size. Every step was an ABX test. For experiments I and II, the reference (X) was a response from which the target reflection had been removed completely. Participants had to recognize this rendering amongst the other two, one of which was the response with the reflection added. For test III, the reference was the original response and one of the other two options was the modified response, where the location of the left lateral reflection was modified.

Upon the start of the test, the modification made to the response was well above the threshold, expected from pre-testing. For the level tests the modifications began with 12 dB above the original level of the reflection. For the direction test, the target reflection started at the center. After an initial burn-in phase, where each correct response leads to values closer to the reference, a 3-down-1-up rule was applied. This means participants had to give three correct answers, before reducing the amount of modification, but one incorrect answer immediately increased the amount of modification for the next trial. Further, following a recommendation by [11] for targeting the 73.93% point of an assumed psychometric function the down and the up step where not of the same size.The step down was set to only 84.2% of the up step. Note that [11] also warns about using adaptive tests of insufficient length, as convergence can not be assured. In our experiments we let participants do 8 reversals, which was on the limit of experimental time for a single stimulus.

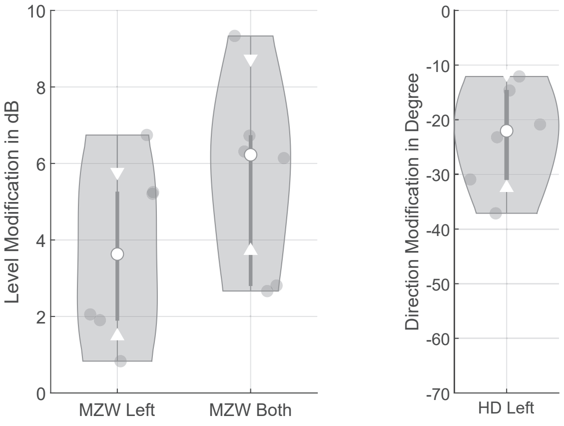
The stimuli for the tests was the 1st violin of the Trio from the String Quartet op. 76 n. 1 by F. J. Haydn [12], from the last note of bar 52 until the first note of bar 62. This is a soloist passage that presented a large range for the fundamental frequency (196 Hz to 1318.51 Hz), with a continuous musical sentence without interruptions and musically independent from the rest of the string quartet instruments.

Six participants (mean age 32.3y) took part in the test, all of which had previous experience in listening experiments. The test was conducted at the Aalto Acoustics Labs, in the listening room “Wilska”, that comprises of 47 Genelec 8331AP loudspeakers, surrounding the listener in an anechoic environment.

Results

The results of the adaptive threshold tests for experiments I and II, level modifications, are shown in Figure 2a. The first noteworthy observation is that the threshold for the level modification tests are positive in both experiments, and the median threshold for the case in which two reflections were modified was higher than when modifying one reflection. Please note, that the overall level was kept constant.

With regards to experiment III, Figure 2b, the angular threshold was between −35 and −12 degrees, implying that the very prominent side wall reflection could be further to the front while evoking the same percept as the real position.



(a) Experiment I and II (b) Experiment III

Figure 2: Listening Test Results

After completing the three threshold tests, participants were also asked to write down the auditory attributes to which they mostly attended when performing the test. The most frequently mentioned attributes were sound source location, ASW. Occasionally the reflection appeared as a separate auditory event. An image shift was reported by all participants in experiments I and III, but not in II. On the other hand, differences in SI, especially ASW, were audible in all three experiments, but mostly in experiment II.

Discussion

At first sight, the results show surprisingly high thresholds for the detection of reflections. Interestingly, the positive thresholds for experiments I and II implies that one or even two individual reflections can be completely removed from the spatial room impulse response without the participants detecting the change. In the case of one side reflection, the actual level has to be increased by 3.6 dB, until the reflection is audible. At this level, the reflection is only 2.11 dB below the direct sound, see Table 1. This threshold is significantly higher than values found in literature for anechoic experiments, but close to values from experiments in normal rooms by Olive and Toole [5].

What is striking is that when changing the level of both sidewall reflections, the threshold is even higher, 0.48 dB (left reflection) and 1.18 dB (right reflection) above the direct sound. This points to the fact that there are different perceptual consequences when changing reflections, which was highlighted by the listeners. The image shift effect, only reported in experiments I and III, could be easier to identify than changes in the ASW. An increase of SI is also reflected by the increase in early lateral energy.

Experiment	Modification Type	Median Modification	Level Difference to Direct Sound (dB)	Δ DRR (dB)	Δ DRR$_{\text{Early}}$ (dB)	ΔJ_{LF}
I	Level Left	+3.63 dB	-2.11	-0.28	-0.63	0.039
II	Level Both	+6.23 dB	+0.48 / +1.18	-0.70	-1.49	0.097
III	Direction Left	$-22.04°$	-	-		-0.003

Table 1: Determined threshold and change of acoustics parameters at the threshold. Note that the median modification refers to the change made with respect to the original response, difference in direct-to-reverberant ratio Δ DRR, difference of ratio between direct and early energy Δ DRR$_{\text{Early}}$, and difference in early lateral energy ΔJ_{LF} are calculated with respect to the reference used in the experiment, i.e. depicts the changes of parameters that participants were exposed to. DRR, DRR$_{\text{Early}}$ and J_{LF} at mid frequencies (average of 500Hz and 1kHz octave bands).

When comparing the direct-to-reverberant ratio (DRR) in octave-bands of reference and probe, the largest differences were found at high frequencies (-2 dB at 4 kHz vs. -0.3 dB at 250 Hz). This shows that for further experiments using this method, it needs to be considered that a 2 ms window might not be sufficient to capture the full reflection at low frequencies.

Conclusion

We have presented experiments to test the threshold of level and directions of early reflections in auralizations of two chamber music halls.

The experiments show that the median level modification with respect to the original response are positive, i.e. that the level of either one or two reflections actually needs to be increased with respect to the original response in order for a change to become noticeable. Furthermore, modifying one or two reflections had significantly different thresholds with reported different perceptual effects. The different thresholds are likely explained by the different perceptual cues that the experiments evoked.

The directional threshold test has shown that when approaching the true direction of this lateral reflection from the front, over 20 degrees of deviation can remain unnoticed. This indicates that, in this experiment, this reflection could come from a broad range and still evoke the same perceptual responses.

As we have observed, even prominent early reflections can be removed from the response without being noticed.This work highlights that an impulse response should not be thought of as a collection of individual reflections, each of which with its own perceptual effect. Rather, it is the interaction of all reflections, as the acoustic field builds up, that needs to be considered.

Future work includes also experiments with more participants, and a wider range of test signals, in order to verify the found high thresholds. Finally, with more data at hand, auditory models could be applied.

Acknowledgments

This research was supported by the EU's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 812719.

References

[1] Green, E., Kahle, E.: 'Dynamic Spatial Responsiveness in Concert Halls', Acoustics 1(3) (2019), 549–560

[2] Sessler, G. M., Atal, B. S., Schroeder, M. R., West, J. E.: Acoustical Measurements in Philharmonic Hall (New York). J. Acoust. Soc. Am. 36(5) (1964), 1011–1011.

[3] Wettschureck, R.: Über die Abhängigkeit Raumakustischer Wahrnehmung von der Lautstärke. Dissertation. Berlin, Germany, 1976

[4] Kuttruff, H.: Room acoustics (pp. 159-163), 6th ed. CRC Press, London, England, 2016.

[5] Olive, S. E., and Toole, F. E.: The detection of reflections in typical rooms. J. Audio Eng. 37 (7/8) (1988), 539–553

[6] Barron, M., and Marshall A.H.: Spatial Impression Due to Early Lateral Reflections in Concert Halls: The Derivation of a Physical Measure. J. Sound Vib. 77 (2) (1981), 211–232

[7] Barron, M.: The subjective effects of first reflections in concert halls-The need for lateral reflections. J. Sound Vib. 15(4) (1971) 475–494.

[8] Haapaniemi, A. and Lokki, T.: Identifying concert halls from source presence vs room presence. J. Acoust. Soc. Am. 135(6) (2014), EL311–EL317

[9] Müller, L., Perceptual Differences Caused by Altering the Elevation of Early Room Reflections. Master's Thesis. Chalmers University, Göteborg, Sweden, 2021

[10] Tervo, S., Pätynen, J., and Lokki, T.: Acoustic reflection path tracing using a highly directional loudspeaker. IEEE WASPAA (2009), 245–248

[11] García-Pérez, M. A.: A cautionary note on the use of the adaptive up–down method. J. Acoust. Soc. Am, 130 (4) (2011), 2098–2107

[12] Colella Gomes, O., Lachenmayr, W., Thilakan, J., and Kob, M.: Anechoic Multi-Channel Recordings of Individual String Quartet Musicians. I3DA (2021)

[13] Tervo, S.: SDM Toolbox. `https://se.mathworks.com/matlabcentral/fileexchange/56663-sdm-toolbox`