Objective analysis of the dynamic responsiveness of concert halls

Tapio Lokki* and Jukka Päätynen†

Department of Computer Science, Aalto University, PO Box 13000, Aalto, Finland

Abstract: One of the central acoustical features of a concert hall is its ability to make sound sufficiently loud. Acoustics researchers often measure the objective parameter strength to investigate sound-amplifying properties of a hall. However, the strength is a linear variable, it does not reveal anything about the true dynamic responsiveness of a hall. The hall should render the music expressive with large dynamics and in this part the dynamic responsiveness plays an inseparable role. As an example, we analyze measurements from two concert halls combining the binaural sound levels with additional information on music and dynamics. These factors represent the spectral changes in the source signals as well as binaural hearing sensitivity according to the sound level. With such factors combined to the information obtained from the conventional impulse response, the dynamic responsiveness as well as the actual dynamic range experienced by the listener could be objectively measured. The presented analysis method shows the overall magnitude of differences in dynamic responsiveness that could be observed between concert halls.

Keywords: Auditorium acoustics, Room acoustics, Spaciousness, Dynamical responsiveness

PACS number: 43.55.–n, 43.66.–x [doi:10.1250/ast.41.253]

1. INTRODUCTION

Traditionally, the acoustics of concert halls is objectively evaluated by measuring room impulse responses and computing room acoustical parameters from them. Assuming a room with uniform temperature and negligible airflow, the acoustic response is linear and time-invariant, thus an impulse response completely describes the system between the source and receiver positions. Moreover, the ISO3382-1 measurement standard [1] defines the specification for the omnidirectional sound source and directivity of the microphones, being omnidirectional, figure-of-eight, or a dummy head for binaural parameters. From the engineering point of view, the standard is reasonable as carefully conducted measurements by different people can be compared, and the resulting room acoustical parameters allow the comparison of halls objectively.

The standard also suggests “subjective listener aspects” and relations of those subjective aspects to a few objective parameters. For standard measurements, the source is defined to be an omnidirectional point source and it is assumed to be linearly level dependent. Thus, looking at the source — medium — receiver communication model, depicted in Fig. 1, the measured impulse response explains only the medium part, not the real sources and receivers.

However, real music instruments have frequency dependent directivities and their spectra are level dependent, thus the standard omni-loudspeaker is quite far from reality. In addition, the receiver in measurements is again a linear microphone, which might have some basic directivity properties. As Fig. 1 shows, the subjective listening includes level dependent non-linear properties that are totally ignored in the measurements. In short, level dependent variables are the spectra of musical instruments and sensitivity of human hearing. Therefore, the ISO3382-1 objective room acoustical parameters fail to accurately represent subjective aspects observed by human listeners.

Recent research has showed that non-linear measures could better explain the sensations of listeners in a concert hall. Lee and Cabrera [2] showed that perceived reverberance was dependent on the listening level when subjects listened to measured impulse responses. Later, they replicated the listening tests with music convolved with monaural impulse responses and found out that level dependent reverberation time values correspond better with perceived reverberance [3,4]. Päätynen et al. [5,6] showed that perceived dynamics varies between halls, even if the measured room impulse responses are linear and time-invariant.

In this article, the level dependent changes in source spectrum and the sensitivity of a human receiver are connected together with measured room impulse responses. It is shown with real measured data that perceived

*e-mail: tapio.lokki@aalto.fi
†e-mail: jukka.patynen@aalto.fi
dynamics varies between halls with the identical stimuli. However, we do not propose any new level dependent parameter, but we hope that our example analysis would trigger more research on the dynamical aspects of multi-faceted perception of music in a concert hall.

2. THE LEVEL DEPENDENT SPECTRUM OF AN ORCHESTRA

When musicians play their instruments at different dynamic levels, the emitted sound power changes accordingly. In addition, the other discriminating and measurable changes are in timbre, in particular in spectral skewness [7]. Thus, the individual orchestra instruments have level dependent spectra, as shown already by Luce [8]. The largest variation in spectra exists with brass instruments, which excite much more harmonics when they are played with increased force. To estimate how the spectrum of the entire orchestra changes along music dynamics, the sound of the whole orchestra is analysed from an orchestral work by J. Sibelius (first movement of Lemminkäinen suite). The music was recorded with multiple close microphones on the stage of a real concert hall for a commercial recording [9]. For the analysis, instrument sections were mixed to balance and all tracks were merged to a mono recording, which has high signal-to-reverberation ratio due to the close positioning of the microphones. The entire piece is 956 seconds long and for the analysis it was segmented into one-second frames with 50% overlap. From each frame, the linear equivalent level ($L_{eq}$) was computed to detect totally silent frames. Figure 2 shows the histogram of levels indicating that there were 37 frames without any reasonable musical signal ($L_{eq} < 60$ dB) and those frames were omitted. Finally, magnitude spectra were estimated from the remaining over 1,850 one-second frames. Naturally, each frame has a distinctive spectrum as different instruments are in voice in each frame. Therefore, all magnitude values at each frequency bin (256 bins on the logarithmic scale on audible frequencies) were ordered to obtain a rough distribution of spectral magnitudes at different frequencies over the entire piece.

The result of the described analysis is illustrated as percentiles in Fig. 3. The novelty in this type of analysis is that given a sufficiently long music excerpt, the spectral magnitude distribution can be analyzed without accurately selecting specific passages in the music. In theory, the lowest curves represent a particularly quiet full orchestra sound and 2.5% curve is selected to represent pianissimo. Similarly, 25-percentile could correspond to a typical soft piano played by the entire orchestra. Respectively, the full orchestra fortissimo is represented with a curve of (97.5%) in the following analysis. The highest values (100%) were not thought to be representative as there might be unnatural peaks in some frames. To summarise, such analysis provides a statistical estimate for the dynamic spectrum change for the entire orchestra, including the typical orchestration.

To illustrate the frequency dependent dynamic range Fig. 4 depicts disproportionate differences between fortissimo and piano or pianissimo. When 0 dB level has been
set at 1 kHz, it is seen that below 200 Hz the differences in spectra are up to 15 dB. Respectively, above 1 kHz the level differences are between 5 and 10 dB up to 10 kHz. Below 100 Hz three individual peaks are seen, which are emphasised most probably by the musical key of the piece, with other music peaks would be different. Therefore, it can be concluded that below 100 Hz all the way down to 30 Hz the level difference between extreme dynamics is 10–15 dB larger than at 1 kHz. In pianissimo dynamic the lowest frequencies could include background noise, and the 25% curve likely represents more reliably the silent playing of the lowest frequency instruments. To conclude, this case study illustrates that the overall dynamic range of an orchestra, recorded close to instruments, varies between 50 to 70 dB, depending on the frequency.

3. SPECTRA OF THE HALL’S RESPONSE IN DIFFERENT MUSIC DYNAMICS

The room acoustics standard [1] proposes that the subjective level of sound could be estimated with the objective measure strength, $G$. As an example, Fig. 5 shows the values of $G$ in octave bands at two different concert halls in Berlin, Germany. Konzerthaus is a traditional rectangular hall with a high ceiling, a flat audience area with lightweight chairs, and side balconies. Philharmonie is the first hall of its kind having audience surrounding the stage, raked audience sections with solid seats and hardly any side walls to reflect lateral energy to the two measured positions at the main stalls. The impulse responses were measured from 24 source positions and the variation between sources are on average within 5 dB at each octave band. The measurement sources were studio monitor loudspeakers [10], and therefore there is a slightly more variation in $G$ values compared with standard omnidirectional sources.
The computed values show that Konzerthaus has higher strength on the lowest octave bands and values are lower on the more distant seat, as expected. In addition, the median $G$ values on different seats are closer to each other as in Philharmonie. As discussed earlier, the values of $G$ do not tell anything on the contribution of the halls on the musical dynamics. To be able to see any possible effect on actual musical dynamics, the spectral change according to the played dynamics and the sensitivity of human hearing has to be included in the analysis.

In the following objective analysis, low and high frequencies are considered separately, because they contribute to changes in perceived musical dynamics with two different mechanisms. The low frequency phenomena are connected with the spectral changes in music and the sensitivity of human hearing, while the high frequency phenomena are more complex combining binaural hearing and lateral early reflections (i.e. concert hall geometry) as well as auditory masking effects on the spectral changes in music.

4. **ONE WAY TO ESTIMATE MUSICAL DYNAMICS AT LOW FREQUENCIES**

As an example, the frequency responses of two halls were computed as averages over all source positions used in the measurements, thus the presented responses are averages of 24 one-third octave smoothed frequency responses. Moreover, presented responses are measured with a dummy head and the binaural responses (left and right ears) are merged together with power summation [11]. Finally, the magnitude responses are multiplied with the spectra of the orchestra in different played dynamics (Fig. 3) and the results of the combined frequency responses are depicted in Fig. 6. Thus, the curves show the frequency response at the eardrum of a listener in different musical dynamics. As the $G$ values indicated, Konzerthaus has more sound energy on both seats, in particular at low frequencies. Note that due to the used loudspeakers in measurements the low frequencies roll off steeply below 70 Hz.

To see the possible differences in perceived dynamics between these two halls, the hall times $pianissimo$ spectra can be divided from the hall times $fortissimo$ spectra (naturally on magnitude scale the division is a subtraction). However, such operation does not make sense as in each hall the spectra of different dynamics are multiplied with the same spectra of a hall. In other words, when a hall amplifies both $pianissimos$ and $fortissimos$ the dynamic range is untouched. Nevertheless if the spectra in Fig. 6 are first mapped to equal loudness contours, which correspond the sensitivity of human hearing at different levels, the loudness values in $sones$ are obtained. Now, they indicate the loudness at all frequencies in each hall. As a result, the frequency responses of loudness in $sones$ are depicted in Fig. 7 and it is seen that the perceived dynamic range is indeed larger, in particular at low frequencies. The difference comes from the fact that the equal loudness contours are denser at low frequencies, i.e., human hearing expands the perceived dynamic range below 500 Hz.

Finally, it makes sense to subtract the loudness values in different dynamics and Fig. 8 compares the range of dynamics in both halls and at both positions. Moreover, the differences are plotted both at loudness (sones) and at sound pressure levels (desibels) for comparison, although the range in both halls is obviously the same with sound pressure levels. It can be seen that in these two halls the dynamic range in loudness is unequal below 200 Hz. The difference is approximately five $sones$ at maximum, thus clearly noticeable but not huge. However, it should be emphasised that in Konzerthaus the $G$ values at low frequencies are already higher, meaning that music is louder than in Philharmonie. Therefore, the additional
perceived dynamics is even expanding the dynamic range of bass instruments. When conventions in composition are taken into account, this difference becomes more substantial because bass instruments, such as gran cassa, tuba, trombones, and double basses are usually only in voice at forte/fortissimo passages and in many delicate pianissimo passages they are not playing at all. Therefore, the dynamic range of bass instruments is from total silence to forte, which is on much higher level in Konzerthaus than in Philharmonie and emphasized by loudness expanding mechanism in human hearing. Here, the example was only from two different halls, but there are similar variations between other halls that we have studied recently [6,12].

5. ONE WAY TO ESTIMATE MUSICAL DYNAMICS AT HIGH FREQUENCIES

The above example showed that musical dynamics at low frequencies is affected by the sensitivity of human hearing. However, above 1 kHz the equal loudness contours are almost linearly spaced at different levels, thus no differences are seen in Fig. 8. On the seat in 11 m the level of high frequencies is practically the same in both halls, but on the seat in 19 m Konzerthaus has more total energy above 2 kHz than Philharmonie (see Figs. 6 and 7). A plausible explanation for the difference is that Konzerthaus, being a reasonably narrow rectangular hall, provides much more early lateral energy, and that is seen in the frequency responses which were computed from binaural responses.

Nevertheless, the high frequencies definitely make a contribution to the musical dynamics as presented earlier [5,6]. In fact, we have already proposed one objective way to compute the binaural dynamic responsiveness (BDR) [5]. It uses the basic principle of combining the spectral characteristics of orchestral sound in different dynamics, the impulse response of a room and the directional sensitivity of hearing. However, the formulation for BDR substituted equal loudness contours, used here, by a more complex auditory masking model [13].
The change in orchestral spectrum at high frequencies can be seen in Figs. 3 and 4. Overall, the instruments, in particular brass instruments, excite much more harmonic overtones when they are played at higher level. Therefore, the high frequencies are emphasized more than middle frequencies in *forte* playing — hence, the relative balance between frequency regions carries important cues on the musical dynamics. A concert hall has a crucial role in conveying these high frequencies to the ears of the listeners as high frequencies tend to be absorbed fast by air and surface materials. Thus, to maximize high frequency energy, the surfaces providing early reflections should attenuate minimally high frequencies. Finally, the shape of the human head modifies the sound spectrum differently for different directions, emphasizing high frequencies reaching the listener from the side [14]. These three components are relevant to musical dynamics at high frequencies although they cannot completely alone explain perceived dynamics.

One important aspect of human spatial hearing is the sensitivity of hearing for early reflections. Wettschureck [15] studied one reflection at 70 ms in a simplified simulation with speech and varied the direction of that reflection. His results suggest that when such a late reflection is coming from the same direction with the direct sound (or behind the listener), and the overall level is raised, the audibility of the reflection plateau around $-8\,dB$ relatively to the direct sound. However, when the reflection is coming from the side the audibility threshold is lower at high listening levels, even $-16\,dB$ relatively to direct sound. Green and Kähle [16] replicated this test with music stimulus and varied also the delay of the reflection. They obtained very similar results and concluded that perception thresholds for reflections from the front and behind vary little with overall listening level while for lateral reflections, the perception threshold decreases substantially with increasing listening level. To conclude, the audibility of reflections is a function of the listening level and when the level is raised the lateral reflections become more audible, increasing the perceived dynamics and the responsiveness of the hall to increased playing level.

Finally, the contribution of auditory masking [13] is still not fully understood, but it plays a critical role in perceived dynamics. A strong direct sound may mask some early reflections, thus potentially attenuating even more the important high frequencies. In contrast, when the direct sound is not so strong, e.g. due to the elevated stage or long distance between musicians and listeners, then the early reflections are masked less and they could contribute more to the overall perceived level of high frequencies, which are excited only in *fortissimo* passages. Such auditory masking effects need much more research with music as an excitation signal, before the definitive conclusions can be made. Nonetheless, the authors of this paper are convinced that auditory masking has an important role in the audibility of high frequencies, which are crucial for large perceived dynamics.

6. DISCUSSION AND FUTURE WORK

The presented results demonstrate that the inclusion of sensitivity of human hearing influences on the transmission of low-frequency musical dynamics in concert hall acoustics, even though the variations between halls are relatively small at sound pressure levels. As observed in the shape of the equal loudness contours, perceptual sensitivity to sound pressure level variations is pronounced in comparison to middle frequencies. This factor is omitted when evaluating the measured standard strength parameters, as they do not depend on the level. Whereas middle frequency values for room-acoustic amplification effect can be considered as such, the low frequency values should receive much more attention in concert hall acoustics research. From the perspective of loudness perception, the effect of each gained or lost decibel in low-frequency $G$ is more prominent than at the middle frequencies, as illustrated in Fig. 9.

Figure 9 does not emphasise the importance of high frequencies for musical dynamics, as the equal loudness contours are more or less linearly spaced between 1 and 10 kHz. However, as Fig. 4 illustrates there are at least 6 dB difference in silent and loud passages at high frequencies compared with middle frequencies. Therefore, the ability of the hall to convey this frequency region well to the farthest seat is really important, suggesting that high frequency objective strength values should be measured and evaluated carefully. Many times it is seen that results at 4 kHz octave band are not considered, mainly due to the uncertainty of measurements, which are inherited from the uncertainty of measurements, which are inherited from the...
unequal directivity of a dodecahedron loudspeaker [17]. As engineers, we need to design a novel standard loudspeaker, which would enable to measure high frequencies with omnidirectional source reliably up to 10 kHz or even higher.

Here, the simple loudness model, based on equal loudness contours, was used to estimate the perceived loudness. However, there are much more elaborated perceptual models available and they should be tested and embedded in the proposed analysis framework. Naturally, at high frequencies a binaural loudness model would be required, but luckily such models also exist.

7. CONCLUSIONS

Traditionally, the level of music in a concert hall is estimated by measuring the room acoustical parameter Strength, which is the total energy of an impulse response. The strength values are reasonable at middle frequencies and they enable to compare concert halls to some extent. However, at low and high frequencies such linear room acoustical parameter fails to predict the perceived loudness of music as both spectrum of the orchestra and sensitivity of human hearing have non-linear characteristics. Therefore, the perceived musical dynamics in a concert hall cannot be estimated with any existing room acoustical parameter. Here it is shown with real measured data how the variation in loudness, i.e. dynamics, is different in different concert halls and on different seats. The above calculations hopefully pave the way for more research in the field so that new objective room acoustical parameters, which would describe the perceived dynamics, would be able to be defined in the near future.

ACKNOWLEDGEMENT

This research was supported by the Academy of Finland, project nos. 296393 and 289300.

REFERENCES

[1] ISO 3382-1, Acoustics—measurement of room acoustic parameters—part 1: Performance spaces, International Standards Organization (2009).
[2] D. Lee and D. Cabrera, “Effect of listening level and background noise on the subjective decay rate of room impulse responses: Using time-varying loudness to model reverberance,” Appl. Acoust., 71, 801–811 (2010).
[3] D. Lee and D. Cabrera, “Accounting for listening level in the prediction of reverberance using early decay time,” Acoust. Aust., 40, 103–110 (2012).
[4] D. Lee, D. Cabrera and W. Martens, “The effect of loudness on the reverberation of music: Reverberance prediction using loudness models,” J. Acoust. Soc. Am., 131, 1194–1205 (2012).
[5] J. Pätynen, S. Tervo, P. W. Robinson and T. Lokki, “Concert halls with strong lateral reflections enhance musical dynamics,” Proc. Natl. Acad. Sci. USA (PNAS), 111, 4409–4414 (2014).
[6] J. Pätynen and T. Lokki, “Concert halls with strong and lateral sound increase the emotional impact of orchestra music,” J. Acoust. Soc. Am., 139, 1214–1224 (2016).
[7] S. Weinzierl, S. Lepa, F. Schultz, E. Detzner, H. von Coler and G. Behler, “Sound power and timbre as cues for the dynamic strength of orchestral instruments,” J. Acoust. Soc. Am., 144, 1347–1355 (2018).
[8] D. A. Luce, “Dynamic spectrum changes of orchestral instruments,” J. Audio Eng. Soc., 23, 565–568 (1975).
[9] Finnish Radio Symphony Orchestra, https://www.ondine.net/?lid=en&cid=2.2&oid=5545, Compact Disc (accessed 24 April 2019).
[10] T. Lokki, J. Pätynen, A. Kuusinen and S. Tervo, “Concert hall acoustics: Repertoire, listening position and individual taste of the listeners influence the qualitative attributes and preferences,” J. Acoust. Soc. Am., 140, 551–562 (2016).
[11] V. P. Sivonen and W. Eilmerier, “Directional loudness in an anechoic sound field, head-related transfer functions, and binaural summation,” J. Acoust. Soc. Am., 119, 2965–2980 (2006).
[12] T. Lokki, J. Pätynen, A. Kuusinen and S. Tervo, “Disentangling preference ratings of concert hall acoustics using subjective sensory profiles,” J. Acoust. Soc. Am., 132, 3148–3161 (2012).
[13] B. C. J. Moore and B. R. Glasberg, “Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns,” Hear. Res., 28, 209–225 (1987).
[14] T. Lokki and J. Pätynen, “Lateral reflections are favorable in concert halls due to binaural loudness,” J. Acoust. Soc. Am., 130, EL345–EL351 (2011).
[15] R. Wetschereck, Über die Abhängigkeit raumakustischer Wahrnehmungen von der Lautstärke (On the dependence of room acoustic perception by sound level), PhD thesis, TU Berlin, Germany (1976).
[16] E. Green and E. Kahle, “Dynamic spatial responsiveness in concert halls,” 10th Int. Conf. Auditorium Acoustics, Vol. 40, pp. 226–236, Hamburg, Germany, October 4–6 (2018).
[17] T. Knüttel, I. B. Witew and M. Vorländer, “Influence of ‘omnidirectional’ loudspeaker directivity on measured room impulse responses,” J. Acoust. Soc. Am., 134, 3654–3662 (2013).

Tapio Lokki Dr. Tapio Lokki is an associate professor in the department of Computer Science, Aalto university School of science, Finland. His research team is investigating auralization, spatial sound reproduction, binaural technology, and novel objective and subjective evaluation methods, as well as physically-based room acoustics modeling methods. Particularly, the interest has been in concert halls, in which the team has developed new measurement techniques, analysis methods for spatial impulse responses, and sensory evaluation methods to understand the perceptual differences between concert halls.

Jukka Pätynen Dr. Jukka Pätynen is a post-doctoral fellow in the department of Computer Science, Aalto university School of science, Finland. His research interests are concert hall acoustics and music. In particular, he is investigating the emotions raised by music and the contribution of performance space to emotions.