Lokki, Tapio; Pätynen, Jukka

Auditory Spatial Impression in Concert Halls

Published in:
The Technology of Binaural Understanding

DOI:
10.1007/978-3-030-00386-9_7

Published: 01/01/2020

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Lokki, T., & Pätynen, J. (2020). Auditory Spatial Impression in Concert Halls. In *The Technology of Binaural Understanding* (pp. 173-202). Springer. https://doi.org/10.1007/978-3-030-00386-9_7

Auditory spatial impression in concert halls

Tapio Lokki and Jukka Pätynen

Department of Computer Science, School of Science, Aalto University, Espoo, Finland

Summary. This chapter discuss acoustics of concert halls, in particular from the viewpoint of binaural perception. The early reflections have a crucial role in the quality of sound, perceived dynamics, and timbre. In particular, the direction from which these reflections reach the listener is important due to the properties of human spatial hearing. The chapter has strong links to psychoacoustical phenomena, such as the precedence effect and binaural loudness. We discuss which aspects of a concert hall give listeners an impression of intimacy and a perception of proximity to the sound. Moreover, it is explained why a concert hall could change the perceived dynamics and spaciousness. Examples are presented with the measured data from real concert halls.

1 Introduction

Concert halls are buildings dedicated to performing and listening to non-amplified music. Audiences gather to these venues to have the best possible acoustical conditions to enjoy live music, but also to socialize. The acoustical conditions of concert halls have been studied scientifically more than a century, but major part of the literature ignore the human binaural hearing. Traditionally, the halls are studied with impulse response measurements to collect the objective data that can be compared between different concert halls. Moreover, most of these data are measured with omnidirectonal microphones, thus again ignoring some information that our binaural hearing could benefit from. The reason for such an approach originates from the idea that such measurements give technically valid and reliably reproducible results of the acoustical features. However, they do not tell us exactly how a hall sounds, as some properties of our binaural hearing are ignored.

The most natural way to study music perception in concert halls is to listen to concerts in-situ and gather opinions from the audience as well as from the musicians and conductors. This method has been popular since the fundamental work by Sabine (1900). Beranek (1962) has authored numerous articles and a few books, including the comprehensive technical data, on the acoustics of concert halls. Moreover, seminal work in the area has been published by Hawkes and Douglas (1971), Barron (1988), and Kahle (1995). While in-situ listening with one's own ears is the most natural way to evaluate acoustics, an inherent problem is that the performed music typically varies from hall to hall. Moreover, human auditory memory is hardly 10 seconds (Sams et al., 1993), meaning that truly reliable comparison of the acoustics of halls is practically impossible between concerts.

The major step to more detailed comparison was taken when binaural technology was adopted to room acoustics. Pioneering research was conducted in Germany by the groups in Göttingen (Schroeder et al., 1974) and Berlin (Kürer et al., 1969). They both involved dummy-head recordings, capturing the binaural sound in the studied concert halls. In addition, they both understood that the musical stimuli have to be the same in each hall to enable valid comparison. To excite the hall with music, the Göttingen group used two omnidirectional loudspeakers on the stage and they emitted an anechoic stereo recording of the 4th movement of Mozart's Jupiter symphony. For laboratory listening tests, the binaural recordings were reproduced in an anechoic room using two loudspeakers with a cross-talk cancellation technique to preserve the binaural cues. In contrast, the Berlin group followed the Berlin Philharmonic Orchestra on their tour and recorded live music with a dummy-head, while the orchestra was playing the same music program in dress rehearsals in unoccupied halls. The listening tests were later performed in the laboratory with headphone listening, again preserving the binaural cues. Both of these teams found interesting results on the perceptual aspects of concert hall acoustics, but neither of these really concentrated on auditory spatial impressions or the benefits of binaural hearing.

The auditory spatial impression and other perceptual factors related to binaural hearing have been investigated in several studies. Marshall (1967) introduced the concept of spatial responsiveness as he proposed that narrow halls with high ceilings have more of such quality. In contrast, a broad hall with low ceiling lacks spatial responsiveness. Based on these observations, it is clear that spatial responsiveness involves directional effects and, thus, binaural listening. In a review article, Marshall and Barron (2001) refer to the article by Kuhl (1978) (written in German), in which the connection between sound pressure level, lateral early reflections and the degree of spatial impression is discussed. The level-dependency was shown even before that by Keet (1968), who showed that increased sound level produces an impression of spatial widening of the sound source. However, since then the level of the orchestral music has been mainly ignored in research and concert hall acoustics research has concentrated to the analysis of impulse responses. Nevertheless, some objective parameters in ISO 3382-1 (2009) standard, such as j_{LF}, L_j or IACC, are applied to predict binaural properties of sound.

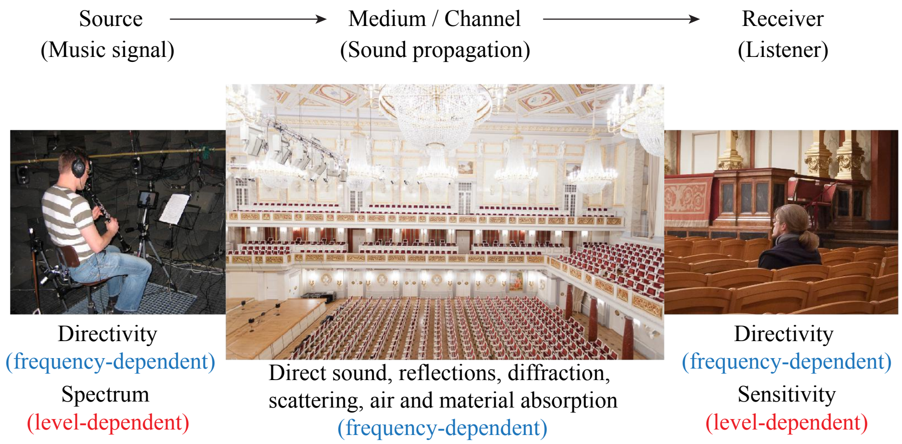

Fig. 1: The basic source — medium — receiver model in a concert hall.

1.1 The objective of this chapter

The excellent review by Marshall and Barron (2001) presents the research in the 1900's on spatial impression in concert halls. This book chapter concentrates mainly on the research performed in the last decade. The main objective lies in explaining the room acoustical as well as psychoacoustical reasons of the earlier research results. Therefore, binaural hearing, psychoacoustics and sound propagation in concert halls are discussed together. Figure 1 presents the connection of musical instruments to human spatial hearing from the authors' current perspective. In particular, to explain auditory spatial impressions, both the frequency and level-dependent aspects of music that propagates through a hall to listener's ears have to be linked together. Pätynen et al. (2014) were one of the first authors, who connected the well-known facts of dynamics-dependent spectra of orchestra instruments and the directional sensitivity of the human hearing to early reflections and their directions in the room impulse responses. The connection has been further discussed by Lokki and Pätynen (2015) and Lokki (2016).

Most of the presented results are based on the state-of-the-art auralization system that allows authentic reproduction of concert halls in laboratory conditions. The auralization of the concert hall measurements are accomplished using the process illustrated in Fig. 2. The symphony orchestra on stage is simulated with 33 calibrated loudspeakers connected to 24 channels. The details of the loudspeaker orchestra setup can be found in the previous publications of the authors (Lokki et al., 2011, 2012; Pätynen, 2011). The room impulse response from each of the loudspeaker channels is measured with a type 50-VI 3D vector intensity probe (G.R.A.S., Denmark) consisting of three co-centric phase-matched pairs of omnidirectional microphones arranged on the x, y, and z axes. The distance between the opposing capsules is 100 mm and the impulse responses are measured with 48 kHz sampling rate using the logarithmic sine sweep technique (Farina, 2000). The six impulse responses measured at a time are analyzed with the Spatial Decomposition Method (SDM) (Tervo et al., 2013) that estimates the direction of incidence for each sample in an impulse response in short time windows. Based on the spatial information, the impulse response in the topmost omnidirectional microphone is distributed to reproduction loudspeakers as convolution reverberators. The distribution of samples is performed with the nearest loudspeaker technique in order to emphasize the spectral fidelity of the high frequencies (Pätynen et al., 2014) at the slight expense of spatial accuracy. Such a choice is adopted based on the earlier results, which clearly shows the importance of timbral fidelity over spatial fidelity (Rumsey et al., 2005). Finally, the anechoic recordings (Pätynen et al., 2008) are convolved with all reproduction channel responses. The distribution of the instruments to stage loudspeaker channels is the same as earlier (Lokki et al., 2011) and when the process is repeated to all sources on the stage the end result is a realistic reproduction of an orchestra in a concert hall.

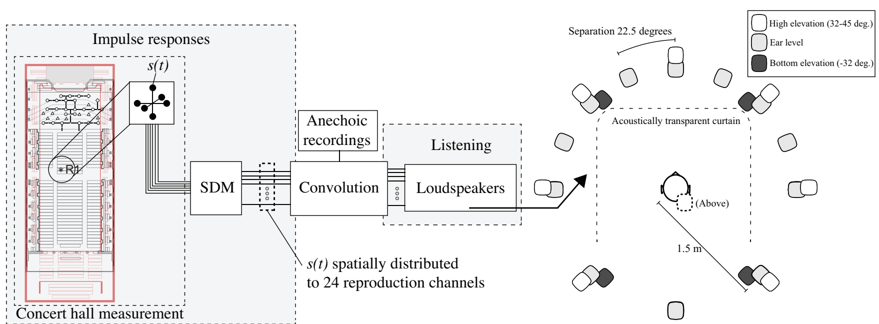

Fig. 2: The block diagram of the auralization with the loudspeaker orchestra measurements in the concert halls. The figure shows the method for a single source channel on stage and the process is repeated for all sources for auralizing the entire orchestra (Lokki et al., 2016).

2 Background and Motivation

Before we can discuss the perceptual aspects we need to illustrate the typical spatial room impulse responses measured in different concert halls. Two well known concert halls in Berlin, Germany, namely the Konzerthaus and the Philharmonie serve here as examples. The former is an example of a classical rectangular hall, which is often called as a ”shoe-box” hall. The latter is a prime example of contemporary design in which the orchestra is located in the center of the hall, and the audience is surrounding the stage on multiple terraces or blocks, hence the moniker "vineyard" hall. Naturally, there exists also other general typologies, but these two extreme architectural examples highlight the differences in spatial distribution of sound energy in a concert hall.

Figure 3 illustrates the measured cumulative sound energy distribution, averaged over 24 source positions on the stage, in the time-frequency-space domain (Pätynen et al., 2013). The analysis shows typical acoustical conditions of a shoe-box and a vineyard hall and highlights the differences across hall types. Although, the illustrations here show only one seat in each hall the other seats have similar properties in both halls. The bottom row shows the average cumulative frequency response at 5, 30, 200, and 2500 ms after the initial direct sounds and the spatiotemporal energy distributions use the same color coding for analyzed time windows.

Direct sounds and adjacent scattering, i.e., the initial 5 ms of the acoustic response arrives from each source on the stage in frontal directions. In a shoe-box hall the stage floor is typically on the ear level of the audience at main parterre, thus the listener does not receive the stage floor reflection, in contrast to halls with inclined seating areas. In Fig. 3b it is seen that the seating rows behind the receiver position reflect sound within 5 ms time window. Fig. 3e shows that there are indeed no seats behind the measurement position in the Philharmonie, as the response was measured on the last row of one audience block. The frequency responses illustrate that in the shoe-box hall the direct sounds lack the low frequencies, but have considerably strong high frequencies. In contrast, in the vineyard hall with a raked audience area, the frequency response of the first 5 ms is quite different due to stage floor reflections.

Early reflections until 30 ms are visualized with dark blue color and they are integrated into the direct sound by human auditory system. Two main differences between the example shoe-box and vineyard halls are the shape of the frequency responses and the spatial distributions of early sound energy. First, the shoe-box hall provides prominent lateral reflections already inside this 30 ms time window, as illustrated by the triangular shape of the dark blue area in Fig. 3a. In addition, reflections from under the balconies are found in Figs. 3b-c. In the vineyard hall, the effect by the wall behind the measurement position is clearly seen in Figs. 3d-f. In addition, the reflectors above the stage contribute to cumulative energy. Nevertheless, there are hardly any side reflections, resulting in an distinct oval-shaped distribution of early energy in the lateral plane.

The second major difference lies in the frequency responses. As seen in panel 3g, the early reflections (between 5-30 ms) in the shoe-box hall strengthen the low frequencies below 200 Hz substantially, yet the middle frequencies up to 1 kHz remain at a relatively low level. Such a time-dependent filtering is caused by the seat-dip effect in halls with flat floor and open seats (Tahvanainen et al., 2015). In contrast, when the sound cannot pass under the seats due to the chair construction and raked floor, the low frequencies

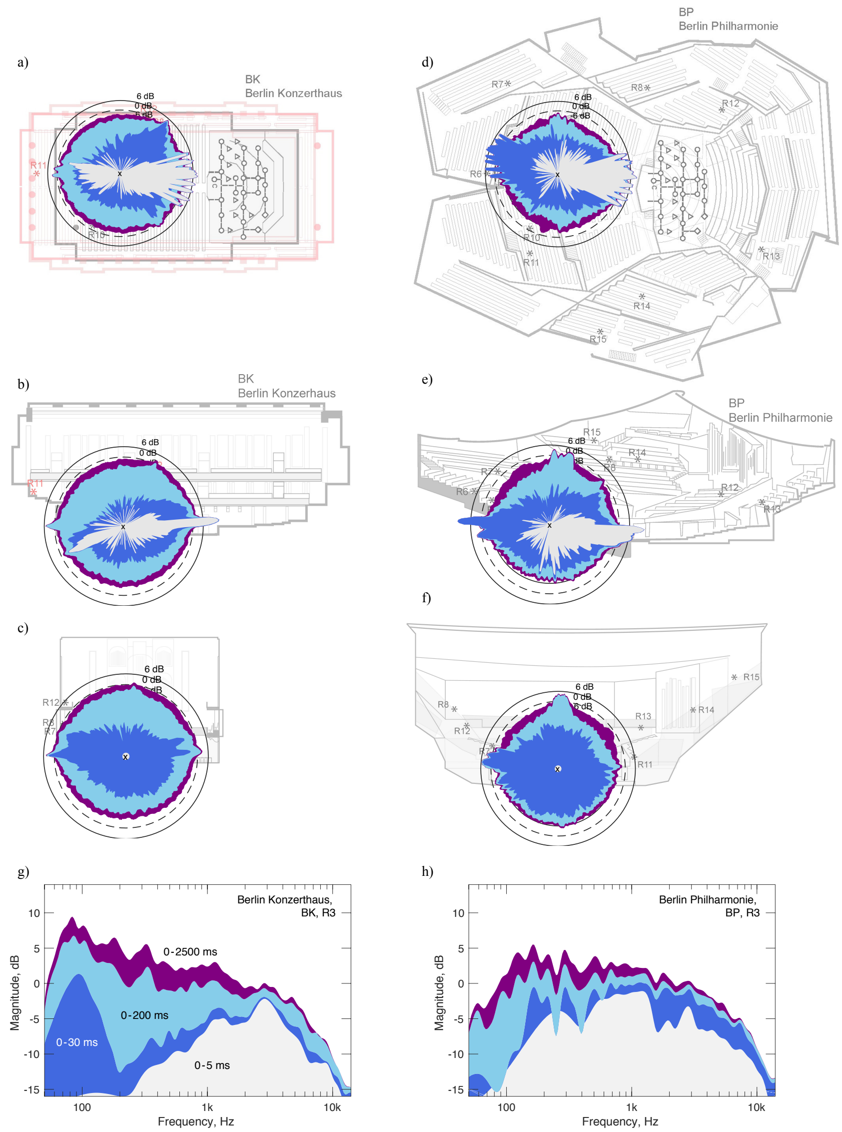

Fig. 3: Spatiotemporal and time-frequency analyses of Berlin Konzerthaus (left) and Berlin Philharmonie (right) concert halls at 15m from the orchestra. Spatiotemporal visualizations are shown in lateral (panels a and d), median (b and e), and transverse (c and f) planes in identical receiver positions. Bottom panels g-h visualize the temporal accumulation of the omnidirectional magnitude responses in respective positions. The time-windows of forward integration shown in panel g are common for all plots.

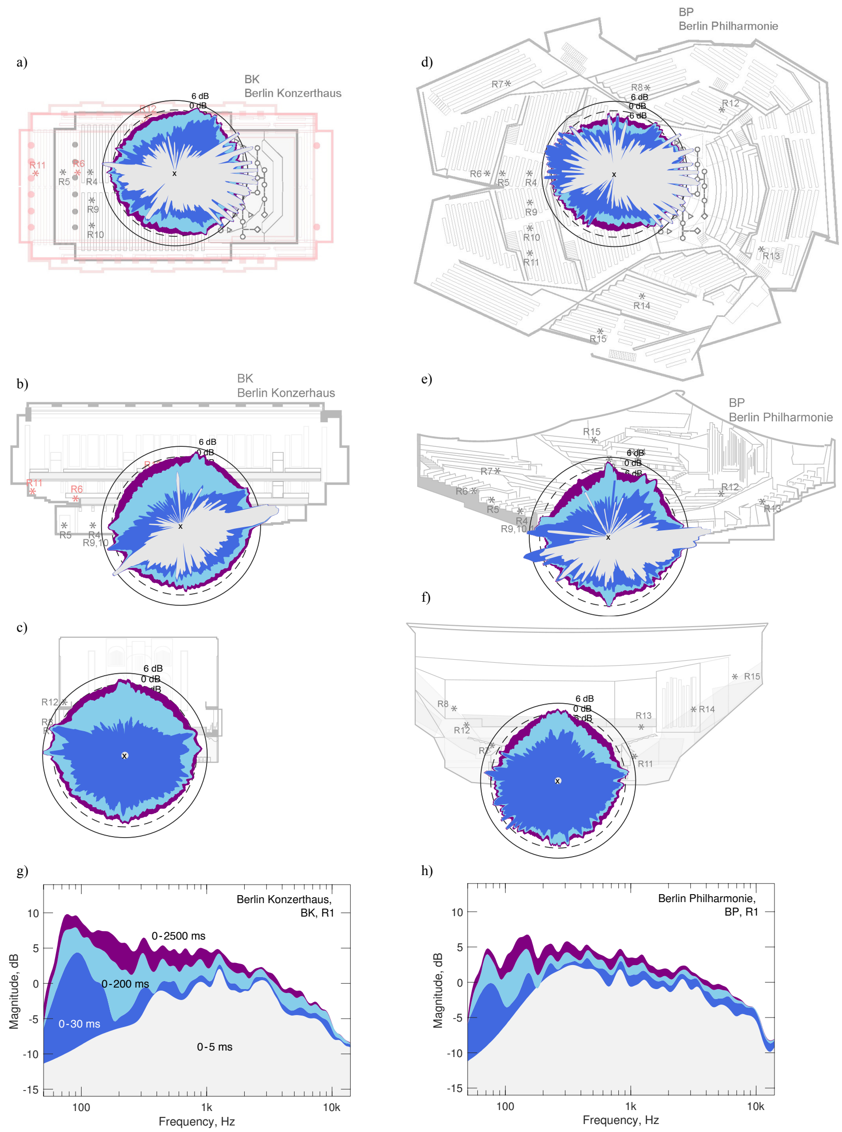

Fig. 4: Visualization of acoustic energy accumulation in two concert halls at a frontal receiver position at 7m from the orchestra. The analyses are respective to Fig. 3.

below 125 Hz are attenuated and sound energy increases only slightly between 5 and 30 ms. At higher frequencies in this particular seat in the vineyard hall, the cumulative energy increases mainly due to the strong reflections from the wall behind the measurement position. It seems that when the direct sound is strengthened by the reflection from the stage floor, relatively weaker early reflections have little contribution to cumulative sound energy after the strong direct sound. It could be argued that the direct sound accompanied with the floor reflection may mask the perception of early reflections. Similar ideas on masking were suggested by Marshall (1967, 1968), but hardly any research has been done on this complex perceptual phenomenon.

Later reflections between 30 and 200 ms increase the overall sound energy in both halls. In the shoe-box hall, the increase is particularly strong above 200 Hz, equalizing the frequency response to be more or less flat at 200 ms after the direct sound. Moreover, the energy in this time window reaches the measurement position almost evenly from all directions and the incident energy in the measured positions has a round shape in all three visualization planes. In the vineyard hall, the energy distribution is not as uniform, and visualizations reveal an array of reflections from the ceiling and reflectors above the orchestra. Although the cumulative energy increases the overall level, the frequency responses retain their shape along the same peaks and dips as earlier. Sound arriving from the high elevation angle of the ceiling likely interferes with the subsequent floor reflection at the microphone array.

Finally, the **reverberation** beyond 200 ms increases the cumulative energy to its final state. In the Konzerthaus the increase is spatially more uniform than in the Philharmonie. Notable differences between these halls can be observed in the smoothness of the overall frequency responses, level of low frequencies, and spatial distribution of sound energy. Figure 3d reveals one distinct late reflection from the right side of the measurement position. Such a reflection might be heard as an echo or it might disturb orchestral balance by highlighting instruments in certain areas on the stage.

2.1 Auditory impressions with real orchestra in example halls

Among the most recent works on the perceived room acoustic quality, a study by Lachenmayr and Pätynen (2016) serves as a close counterpart to the early research by the Berlin group. In this particular study the authors recorded the Staatskapelle Berlin at the dress rehearsals on the consecutive nights in the Konzerthaus and in the Philharmonie, as the orchestra performed "Egmont Overture" by Beethoven in both halls. The recordings were captured with a four-channel pseudo-binaural arrangement with two channels on the sides of an absorbing sphere, and two additional rear surround channels with cardioid microphones. The sound was recorded simultaneously in two equidistant positions from the orchestra, resulting in four sound tracks to be compared. The reproduction utilized four loudspeakers of which two first were at $\pm 45^o$ angles, and the other two at $\pm 135^o$. The listening took place in a relatively dry listening room with additional absorbing baffle immediately in front of the listener's head to reduce crosstalk across the front channels.

Although the technical approach was comparable to the preceding work (Schroeder et al., 1974; Kürer et al., 1969), the authors adopted a different concept in the listening experiments. Instead of charting the acoustic quality in the traditional sense, the authors experimented with the potential *variation* in the auditory impression during changes in music dynamics. As established earlier, music is not a static signal, but in reality it contains continuous changes in many expressive aspects, such as instrumentation and dynamics. For this purpose, a 15-second excerpt with a gradual crescendo as the dynamic variation was isolated from the recordings for a listening test.

The listening test was a full paired comparison with two repetitions and the subjects had to choose the sample, which had more "impact". This "impact" was defined as having more influence, being more interesting, or more effective on oneself. The subjects were asked to initially listen to both crescendos completely instead of switching quickly back and forth. In addition to paired comparison, the subjects were asked to write down one or more descriptive adjectives that described the perceived difference between the samples.

Two different listening tests were completed by 18 and 10 subjects, respectively. The first test reproduced the recordings as such with the possible loudness differences and for the second test the samples were loudness-matched. With the original stimuli the result was clear, the subjects chose the Konzerthaus in both recorded seats having more "impact" than the Philharmonie. When the pairs were loudness-matched, listeners also reported greater impact in the Konzerthaus for the frontal seats, but found no difference between the seats in the back of the two halls. However, the seat closer to the orchestra was always chosen over the farther seat, regardless of the hall.

The elicited descriptive adjectives reveal more detailed information on the perceived differences between the halls and the seats. There were three main differences related to proximity, loudness (strength + dynamics + crescendo), and spatial impression (envelopment + spaciousness + width). The results were totally in-line with the earlier research done with rudimentary simulations (Marshall and Barron, 2001) and measured spatial impulse responses convolved with anechoic orchestral music (Pätynen and Lokki, 2016a,b). In the following sections these three aspects of concert hall acoustics are discussed in the light of the recent research results. We also try to make links to more traditional psychoacoustics research in an effort to increase the common understanding on binaural human perception.

Before concentrating to the perceptual phenomena, we briefly analyze the results of Lachenmayr and Pätynen (2016) in the light of the objective measurements. The measurement positions shown in Fig. 4 are close to the ones used in the recordings and are 8m closer to the orchestra than measurement positions in Fig. 3. The main difference between the frequency responses (Fig. 4g-h) between halls are below 1 kHz, the energy is accumulating quite differently as a function of time. In addition, in the Konzerthaus, there are 5 dB more low frequencies below 100 Hz. Figures 4c and 4f show clearly the spatial distribution of early sound energy. In the Konzerthaus, there are four lateral reflections from side walls and under the balconies. Moreover, the ceiling is quite high resulting later reflection from the ceiling than in the Philharmonie. The Philharmonie has also some lateral reflections, mainly on the left side, but also strong reflections from the reflectors above the orchestra. Thus, there is a clear difference in the spatial distribution of early energy. It is even possible that in the Philharmonie reflections in the median plane reach the listening position earlier than the (relatively weak) lateral reflections. Marshall (1967) suggested that such order of reflections might result in less spatial impression, which was also the result in this listening test. Based on the recent formulation (Pätynen et al., 2014), the lateral early reflections in the Konzerthaus conveys the high frequencies, emphasized in *fortissimo* playing, to the ear drums of a listener, again as found in this listening test. In conclusion, when the orchestra makes a large *crescendo* from *pianissimo* to *fortissimo*, the largest differences in sound pressure level occur at low and high frequencies (Lokki, 2016). Thus, the results by Lachenmayr and Pätynen (2016), obtained with the recordings of a real orchestra, are well supported by the time-frequency-space analysis of the measured impulse responses and the properties of human spatial hearing.

3 Early reflections that affect to proximity, intimacy and engagement

One major purpose of music as an art form is to tell stories, evoke emotions and touch the feelings of a listener. Therefore, it is not surprising that concert halls that sound intimate and engaging are often preferred. During the years researchers have called this aural aspect of a concert hall with different attributes, such as intimacy, proximity, presence, and engagement. As far as we understand, they all mean the same perception, which is suggested to have a major positive influence on preference (Lokki et al., 2012; Kuusinen et al., 2014).

Intimacy is probably the most used term (Beranek, 1992). Beranek's description of intimacy characterizes the listening attribute as closeness of communication between the listener and the orchestra. Moreover, Beranek (1992) defined an objective parameter, initial-time-delay-gap (ITDG), as "the time between the arrival of the direct sound from the stage to the arrival of the first reflection at a measuring point." Unfortunately, the current understanding is that ITDG does not correspond well with intimacy and ITDG has been misleading for many researchers (Hyde, 2018). For example, consider a typical shoe-box hall in which at front rows the ITDG is much longer than in the last rows in the audience area. Everybody can easily understand that a frontal seat feels much more intimate than the seat in the back, although the ITDG suggests vice versa. In addition, ITDG ignores the overall level (i.e. perceived loudness) and spatial location of first reflections, proposing that both a ceiling reflection and a side wall reflection gives the same intimacy.

If ITDG does not explain intimacy, what is then the possible reason for a sound source to sound proximate? Lokki (2014) showed that in their listening tests the most intimate halls were preferred. Naturally, sound pressure level is obvious as the louder the sound, the closer it is perceived. But perceived loudness do not explain everything, the loudest halls do not always sound the most proximate (Lokki, 2014). The spatiotemporal analysis of measured impulse responses at the listening positions revealed that sound is more proximate if there are strong lateral reflections that reach the listener before the ceiling reflection. If the ceiling reflection, or sound from reflectors above the orchestra, is heard before lateral reflections the sound is perceived more distant. The difference is visualized in Fig. 3 where in the Konzerthaus the lateral and under balcony reflections fits into the 30ms window after the direct sounds, but in the Philharmonie, the situation is the opposite; the ceiling reflection is earlier than (weak) lateral reflections. It is possible that early lateral reflections reduce the interaural correlation and that leads to less distant sound, as suggested by Kendall (1995), who discuss the relation of correlation and distance perception in stereo reproduction. Kuusinen et al. (2014) correlated many objective parameters with the listening test data and found out that the lateral early energy fraction (j_{LF} defined as the ratio of sound energies between 5-80ms and 0-80ms captured with figure-of-eight and omnidirection microphones, respectively) at high frequencies was associated with the perception of proximity. This is reasonable, as the research on binaural hearing has shown that due to the directivity of the human head, lateral reflections are louder than median plane reflections at the entrance of the ear canal. The binaural levels for sounds from different directions are illustrated in Fig. 5.

Lateral reflections are crucial for intimacy, as Beranek (1992) wrote, but he repeatedly indicated that intimacy in his terminology is the same perception as spatial impression for Barron and Marshall (1981). Indeed, they made a seminal work to show the importance of early lateral reflections, although there were some limitations in their methodology. Even though, Barron and Marshall (1981) studied early lateral reflections with different sound pressure levels, they used the same recording with different levels missing the natural spectral change of music (see Section 4). This might be one of the reasons that in their studies the spatial impression was not affected by high frequencies over 1.5 kHz, leading to too narrow conclusion that the most important frequencies are the four octaves from 125 Hz to 1 kHz octave bands. For some reason, concert hall acoustics researchers still use only these mid frequency bands in many occasions, although, e.g., Blauert and Lindemann (1986) already clearly showed that spaciousness increases with increasing bandwidth of the later reflections. In fact, they concluded that all sound fields with components in the spectral range above 3 kHz produce a larger horizontal width than those which lack these components.

Furthermore, low frequencies have an important role for intimacy and perception of strong bass is always connected to proximate sound (Lokki et al., 2012, 2016). Here, low frequencies are considered down to 30 Hz, not only the 125 Hz octave band as usual in concert hall acoustics research. The behavior of low frequencies is discussed more in Subsection 4.5, but it is worth emphasizing

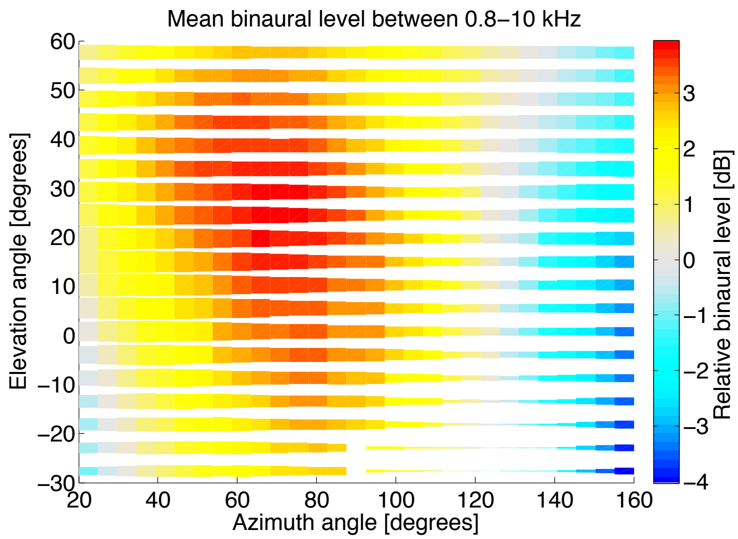

Fig. 5: Binaural level from regions on one side of the head at the frequency band of 0.8-10 kHz. The results show the characteristics of the binaural magnitude responses averaged over a region of ±15 degrees in azimuth and elevation angles centered at the nominal angle in relation to the mean response at the frontal region (±10 degrees azimuth/0-30 degrees elevation). Thickest and thinnest bars denote a variation range of 0.7 and 16 dB, respectively

that lateral reflections also render stronger bass than ceiling reflections do, as seen in Figs. 3 and 4.

Finally, it has to be remembered that intimacy is strongly a multimodal sensation. Hyde (2003) wrote an excellent report considering vision in addition to aural intimacy. As vision is our primary sense, it could override aural intimacy in some cases. On the other hand, vision could also give us a baseline for the intimacy. For example, when sitting in the first row of a balcony visual cues suggest that orchestra is quite far away and we do not expect the music be really loud. Nevertheless, if music is loud, we might think that this is a great hall as we can hear the orchestra so well at such a distance. And vice versa, when sitting parterre closer to the orchestra, our expectation is to hear music well at certain level. However, in many halls that lack lateral early reflections the sound is quite weak at frontal seats and the impression is far from intimate, the sensation is more like watching the orchestra playing in front, and the music is not really touching us (Pätynen and Lokki, 2015). It is evident that more audiovisual studies are needed to find out the multimodal perception of intimacy. Modern technology helps to bring both immersive visuals and audio to the laboratory, as done in very recent study by Postma and Katz (2017). Interestingly, they found that subjects could be categorized into three subgroups; subjects who judged auralizations more acoustically distant with increased visual distance, subjects who judged auralizations louder with increased visual distance, and subjects whose audio judgment was uninfluenced by the visual stimulus.

Kuusinen and Lokki (2015) also discuss the combination of visual and aural percepts for intimacy. Moreover, they focused on intimacy from an auditory and psychological perspective and viewed it as a dynamic feature, which is heavily influenced by the manner how musical expressions are translated and even enhanced by the acoustics of the hall. If a hall can provide dynamically varying spatial cues, which for instance can induce a perception of looming during *crescendos*, the experience of intimacy would be elevated not only by a heightened emotional response to the music, but also by a feeling of deeper involvement with the space we are in.

3.1 The quality of early reflections

Robinson et al. (2013b) studied both with measured and simulated venues the effect of one side reflection on listeners' ability to separate two speech sources on stage. The measurements were done in a 600 seat theater with a binaural headset worn by a male subject. For simulations a simple concert hall model with 11 distinct reflections and late reverberation were used and in both cases the binaural auralizations were listened to with headphones. The listening tests were performed in two different laboratories and the task of the subjects was to say which one of the speakers, male or female, is on the left. In other words, the test investigated the spatial discrimination of multiple acoustic sources in a real and in a simulated room, in which the properties of early reflections were modified.

The results of the listening tests were close to each other in both laboratories. With the real hall, the results show clearly that if the proscenium splay surface, giving the first lateral reflection, was covered with a diffusor the subjects could not distinguish which talker is on the left. In the case of a lateral reflection being even from a flat or from an absorptive surface the task was easier. In other words, the experiment revealed that discriminating the lateral arrangement of two speech sources is possible at narrower separation angles when reflections come from flat or absorptive rather than diffusive surfaces. With the simulated hall all 11 early reflections were treated with measured or simulated diffusors, and again the result was the same, the separation of male and female speakers was easier when the reflections were from the flat surfaces and diffusors hinder the subjects' ability to hear which one of the speakers was on the left (Robinson et al., 2013a).

The above studies were done with speech as a stimulus, but it can be assumed that with musical stimuli the situation is the same. The diffusive architectural surfaces are widely applied in concert halls, but their perceptual consequences are not fully understood, and opinions in favor and against them exist (Oguchi et al., 2018; Kahle, 2018; Marshall, 2018). What the studies of

Robinson et al. (2013b,a) suggest is that diffusive early reflections make it harder to localize the sources, suggesting that diffuse early reflections might blend sources more. Interestingly, Meyer and Kuhl already in 1952 (according to Cremer and Müller (1982, p. 113)) found out that placing large reflectors at both sides of the proscenium in the opera house in Hamburg, the sound source seemed to expand laterally without losing its localization. Unfortunately, they did not continue this line of research further.

Lokki et al. (2011) investigated the perceptual consequences of the temporal envelope of the reflections. Often used diffusers in concert halls change the temporal envelope while reflections from a hard flat surface do not change the phases of the signals. The wave forms and their temporal envelopes of a harmonic signal at auditory bands are illustrated in Fig. 6. Lokki et al. (2011) suggested that with musical signals, as well as speech, the temporal envelope of reflections affect to how well the precedence effect works. They proposed that reflections from diffusors might partially break down the precedence effect, resulting in less clarity. On the other hand this means also that sound sources might be perceived wider and less defined, which is often considered as a desirable sensation as sound is then better blended. The effect of high frequency scattering was also discussed by Kirkegaard and Gulsrud (2011), who reported a harsh sound from diffusers and how the sound quality was better when the diffusers were covered with absorptive material or flat panels.

3.2 Summary of early reflections

Early reflections are really crucial for the quality of sound in a concert hall. Lateral reflections have been acknowledged for a long time to contribute positively to sound quality. However, often they are referred to as increasing the auditory source width, which is a misleading conclusion, as it is not always the wanted property for a sound source. Instead, we want to have proximate and engaging connection to music, produced by a musical instrument or a singer.

Lateral reflections, if they preserve the temporal envelope of signals, i.e., are reflections from flat or convex curved surfaces are integrated well to the direct sound in our hearing system due to the precedence effect. The phenomenon keeps the localization in the first wave front, but does not render the early reflections inaudible. Early reflections increase the overall loudness, color the sound and might change the perceived width of the source. As said, the temporal envelope preserving lateral reflections integrate to the direct sound best, increasing its quality and preserving the ability to localize. If such a reflection is coming from the median plane, i.e. from ceiling or reflectors above an orchestra, the sound quality might be reduced due to coloration, which is the same in both ears. Moreover, such ceiling reflection might increase the interaural correlation, which could increase the perceived distance of the source (Kendall, 1995).

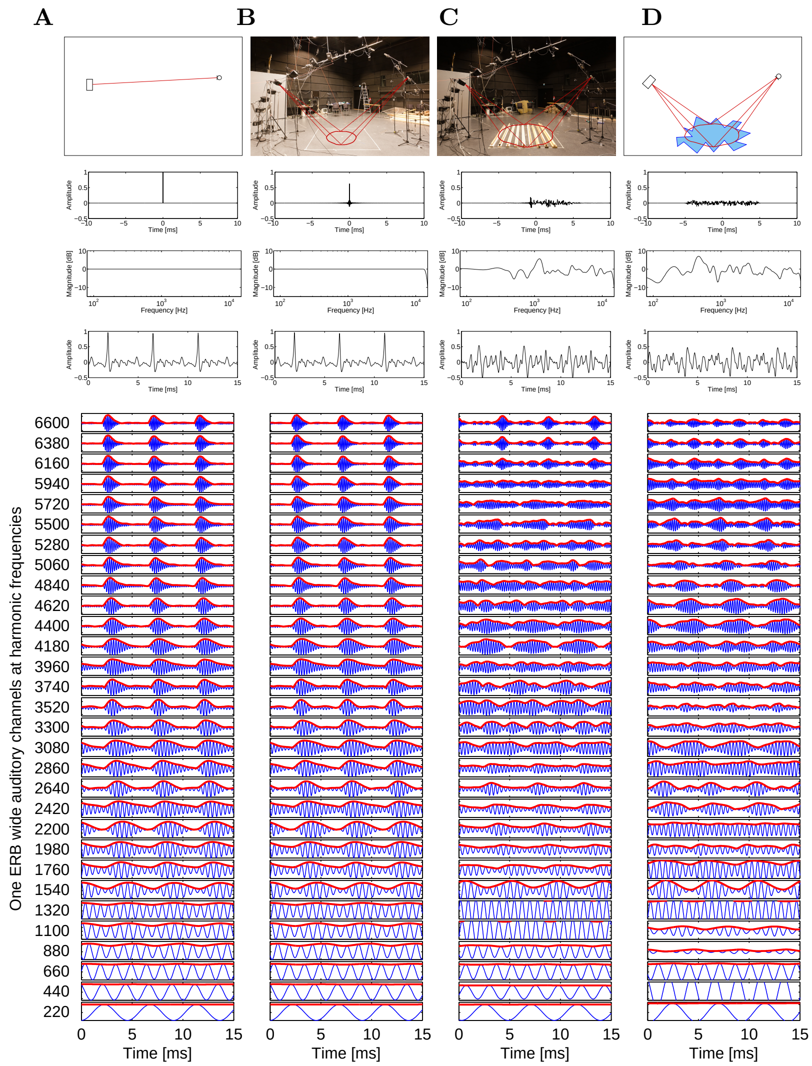

Fig. 6: From top to bottom: Illustration of a direct sound or a reflection, impulse response, frequency response, impulse response convolved with the trumpet sound of A3 (220 Hz), amplitude and envelope of convolved trumpet sound at ERB wide bands, i.e., the waveforms (blue) and envelopes (red) at the outputs of auditory filters on the basilar membrane. (**A**) Direct sound (**B**) Temporal envelop preserving reflection (**C**) Temporal envelope destructing reflection at high frequencies (**D**) Temporal envelope destructing reflection at all frequencies.

If the reflections are scrambling the phases of upper harmonics, i.e. reflections from heavily diffusing surfaces, the precedence effect might partially break down and such early reflections are not fully integrated to the direct sound (Lokki et al., 2011). Such reflections might increase the perceived width of the source to the detriment of less defined location of the source. As a result the instruments better blend together, but some listeners associate that to reduced clarity.

4 Time varying spectrum of music

The previous sections illustrated how the room acoustics function as a transmission channel for the information expressed by the music signal. In particular, the role of early reflections was discussed. Although music is much more abstract than speech, music often aims to convey expressions or emotions. European-influenced classical music offers composers a variety of key elements for expressiveness, such as pitch, note duration, timbre, and dynamics (Owen, 2000, p. 6). Without the intended variations in the music dynamics and tone color, the expressiveness is often diminished. This, in turn, would impede the listeners' experience and the possible emotional impact sought from the performance. After all, experiencing the music is for many listeners the foremost reason for visiting concerts.

With this background the importance of transmitted expressiveness becomes evident. Still, a survey on the related literature on room acoustics perception reveals that the music is mostly considered quasi-stationary excitation signals, and that the time-varying property related to the expressiveness — a key aspect of music — is typically left without consideration. Of the several aspects of expressiveness, the recent research has concentrated on the music dynamics due to its relatively straightforward interpretation and simple measures. In order to include the aspect of music dynamics into the overall concept of perception of acoustics, we need to briefly discuss the properties of orchestra instruments.

The sound of most musical instruments is based on harmonic vibrations and pressure waves that are excited by the musician at a desired force and style. Depending on the type of the instrument, the amplitudes of harmonic overtones can vary strongly with the excitation. The higher harmonics are weakest at the minimum level of excitation, and the overtones become richer as the dynamics is increased. This effect applies to practically all instruments of a typical orchestra. The dynamic spectrum of individual instruments have been reviewed by e.g. Luce (1975) and Meyer (2009). These studies demonstrate that the most prominent dynamic spectrum effect is present with the brass instruments, where the magnitude slope of the overtone frequency envelope shows extreme variations between opposite playing dynamics. For many instruments, the amplitude of high frequency overtones varies more than the amplitude of the fundamental. Consequently, the spectral content of the mu-

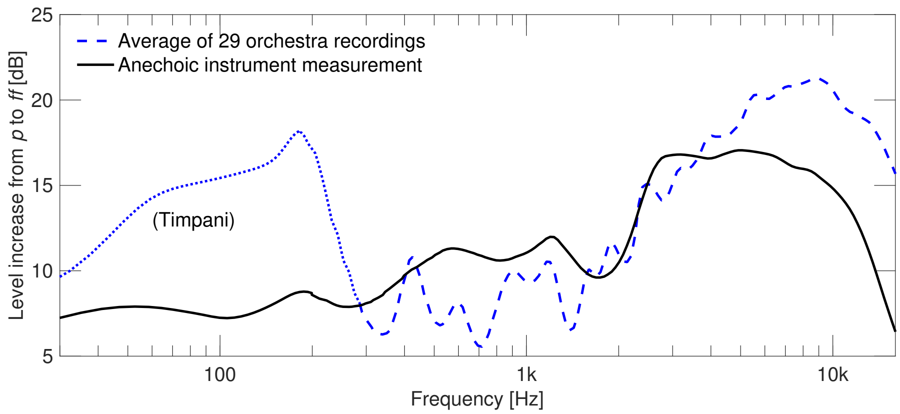

Fig. 7: Spectral change of full orchestra dynamics based on two datasets.

sic signal varies disproportionally more at the high frequencies if the same pitches are played in different dynamic levels.

Two examples of orchestra dynamics are illustrated in Fig. 7. The diagram includes the spectral analyses of two datasets on orchestra instrument signals. The dotted line employs data collected from 29 commercially published recordings of Bruckner's Symphony no. 4, II movement, bars 19-26. The particular passage includes a notable *crescendo* from indicated *pianissimo* to *fortissimo* with full orchestra. The excerpt is particularly useful for analyzing dynamics since the harmony and note pitches remain unchanged during the entire passage with only small variation in the rhythmic pattern. The played notes of the orchestral parts extend from $B\flat 2$ to $F6$, which corresponds to the approximate frequency range of 120-1400 Hz. The dashed curve in Fig. 7 shows the level difference between softest and loudest dynamic levels. As discussed above, the frequency range from 300 Hz to the highest fundamental of 1400 Hz is increased by 5-10 dB from *pp* to *ff*. In contrast, the frequency band consisting only of overtones gains up to 20 dB level increase during full-orchestra *crescendo*. The low-frequency emphasis is attributed to the strongly emphasized timpani tremolo near the *crescendo* peak.

The solid black line included in Fig. 7 presents the results from anechoic orchestra instrument measurements by Pätynen et al. (2008). In this dataset, separate notes of A-major triads were recorded with each instrument, spanning two octaves of the typical playing range of each instrument. All notes were recorded in indicated dynamics of *pianissimo* and *fortissimo*, and the maximum spectra of all notes were estimated from the signals. The number of each instrument in a typical orchestra complement of 83 players was simulated in the overall spectrum. The result yields a similar trend as the first example, as the lowest frequencies up to 300 Hz show an average *pp*-to-*ff* level increase of 7 dB, middle frequencies up to 2 kHz gain around 10 dB, and the region of overtones reach up to 16 dB addition.

While these examples illustrate the spectral effect of full-orchestra dynamics, such situations can be conservative considering the overall variation of dynamics in orchestral works. Contrasting dynamics with full orchestra occur relatively infrequently, and more often expressiveness is realized with variations in instrumentation and the texture of the instrument parts. For example, in lower overall dynamics only some instruments or sections are in voice, but other instruments, such as brass or percussions, join in for more powerful segments (Rimsky-Korsakov, 1922). This effect strongly emphasizes the contrast between soft and strong passages. Therefore it is feasible to analyze the distribution of frequency contents over a longer duration of music which contains a larger variety of instrumentation. For this purpose, the authors analyzed an orchestral recording of entire first movement of Sibelius' Lemminkäinen suite. The recording was derived from close microphone signals captured for a commercial recording on a concert hall stage. The mix-down of the fairly dry signals of different sections give a representation of the orchestra sound over a 15-minute piece. The analysis in Fig. 8a presents the spectral variation as the distribution of occurrence for each frequency. In practice, the entire piece was spliced into one-second frames with 50% overlap, and magnitude spectrum of each non-silent frame was stored. 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 logarithmic scale) were ordered to obtain a rough distribution of spectral magnitudes at different frequencies over the entire piece. The resulting graph (Fig. 8a) shows the percentiles of each frequency bin, and thus the rough spectrum distribution over time. We can then theorize that lowest 5-percentile represents a particularly quiet full orchestra *pianissimo*, or 25-percentile a typical soft *piano* played by part of the orchestra. By comparing these percentile curves to the high full orchestra *fortissimo* curves of (97.5%), the analysis provides a statistical estimate for the dynamic spectrum change for the entire orchestra, including the typical orchestration. These estimates are found in Fig. 8b. The general trend shows a distinct similarity to the results in Fig. 7. In comparison to 300-1000 Hz, the frequency range of overtones is emphasized up to 10 dB. Since this approach takes also into account the variation in music texture, the low-frequency addition is much more prominent than in the investigations shown above. It should also be underlined that the largest level differences occur below 100 Hz, even down to 35 Hz.

Together the presented examples demonstrate that the spectral content of natural music signals is not constant, but it varies heavily with music dynamics. While the varying sound waves, regardless of the level and spectrum, are propagated linearly in rooms, the spectral sensitivity for level and direction of binaural hearing becomes an essential non-linear part of the entire source-medium-receiver model in Fig. 1.

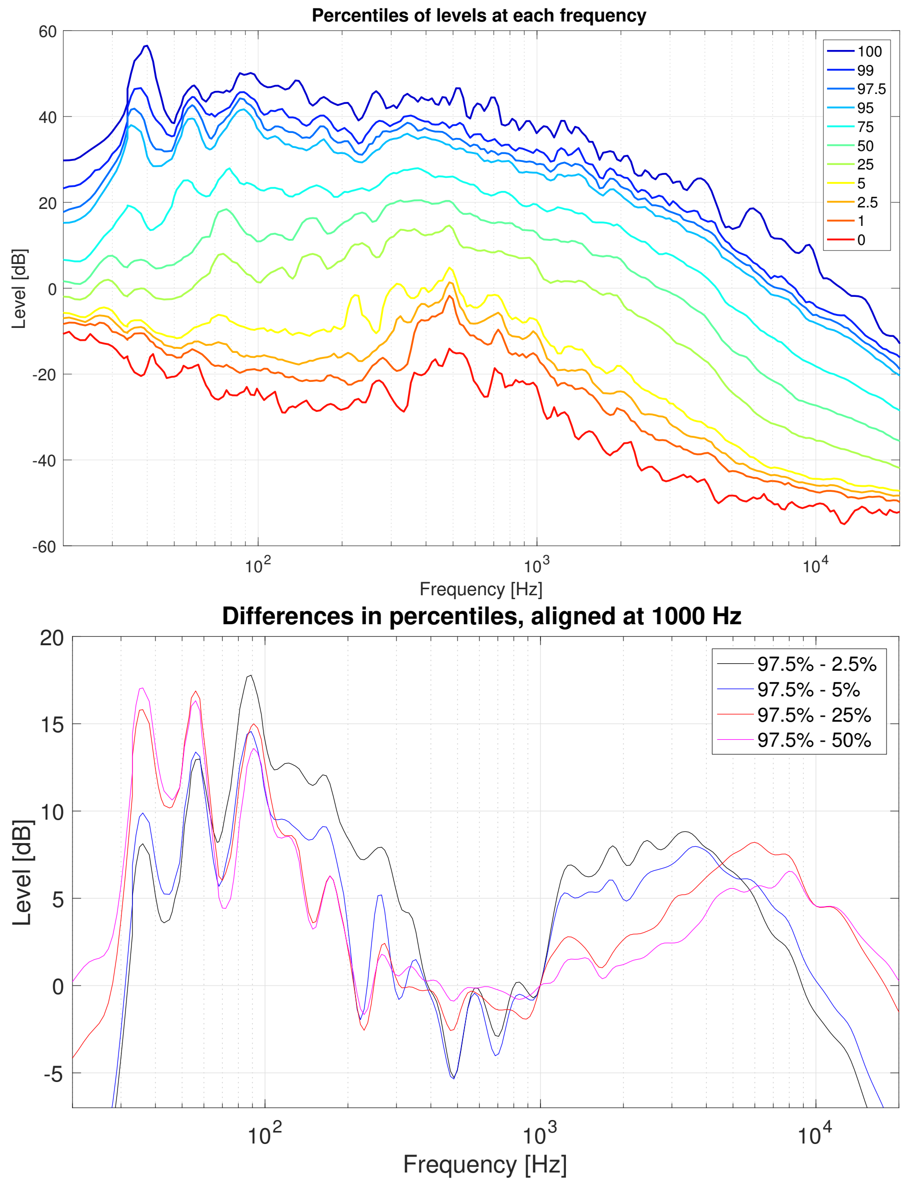

Fig. 8: a) Distribution of orchestra spectrum during a 15-minute piece (Sibelius: Lemminkäinen suite, I movement) over one-second time frames as percentiles. b) Differences between representative *fortissimo* curve and various softer dynamics normalized to 1 kHz. The data is analyzed from a mix-down of multi-channel near-field recording at a concert hall stage.

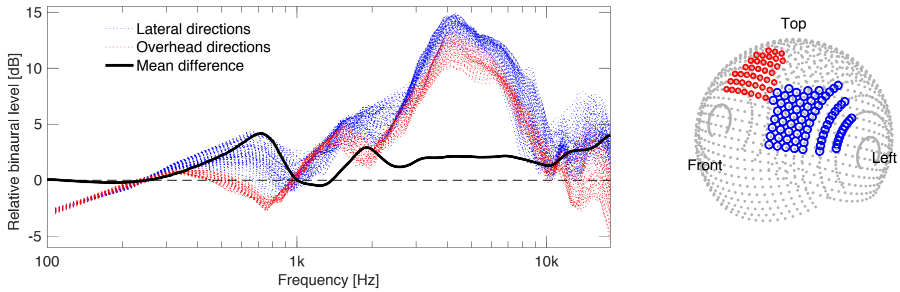

Fig. 9: Example of binaural level of two typical directions for early reflections. Each magnitude response curve is an average over the CIPIC database of 45 HRTF measurements. The solid curve illustrates the average level advantage of lateral incidence (blue directions) to median plane reflections (red directions).

4.1 The dynamic variation at high frequencies combined with binaural hearing

The acoustic effects of the head and outer ear are well-known through the research on binaural technology. One of the earliest reports of directional sensitivity dates back to the 1930's (Wilska, 2011), and the modern concept of head-related transfer functions (HRTF) has been developed through several studies overviewed by Møller (1992). Typically the mid and high frequency bands are perceived more sensitively with the ipsilateral ear when incident sound arrives from outside the frontal directions. Although the shadowing effect of the head can reduce the sensitivity on the opposite ear, the perceived effect with both ears' magnitude response combined is stronger for lateral incidence than with frontal sound. The magnitude of this effect is illustrated in Fig. 5, where the mid and high frequency binaural gain is the most consistent around azimuth angles of 40-80 degrees.

The directional effect of binaural hearing plays a significant part when combined with the propagation paths provided by the room geometry and dynamic spectrum of the sound source. Given a frontal source, the room geometry yields the directions for the reflected sound, which is accumulated by the respective binaural magnitude response. Therefore, the reflection directions are instrumental in the binaural magnitude response of the room. Typical directions for early reflections in concert halls are the lateral angles from the side walls, and median plane directions from ceiling or overhead reflectors. Detailed analysis of the binaural gain for such reflection directions in concert halls is depicted in Fig. 9. By comparing the average binaural gain between sets of lateral and median plane angles it is evident that the lateral reflections yield emphasis on frequency regions of 400-1000 Hz and 1.7-10 kHz. Moreover, there are always multiple lateral reflections, but only one ceiling reflection, as shown in Figs. 3 and 4.

The dynamic variation of the signal spectrum brings in the decisive component in the overall picture. Dynamic variation is emphasized at high frequencies and lateral reflections (40 to 80 degrees) lead to binaural gain in the same high frequency bands. Therefore, lateral reflections may increase the perception of dynamic variation. That was indeed proposed by Wettschureck (1976), who studied with speech the sensitivity of hearing for one reflection at 70 ms. The results were clear showing that sensitivity of hearing is lower for late reflection from the side than from behind or front of the listener when the listening level is high. At low listening levels the sensitivities were more or less the same. Green and Kahle (2018) obtained similar results with music stimuli. Thus, it might be that audibility of reflections is a function of the listening level and when level is raised the lateral reflections become more audible, increasing the perceived dynamics. However, more psychoacoustic research is needed with real music at different dynamics to understand better level dependent aspects of human spatial hearing.

The earlier analyses of measured concert hall acoustics serve as visual examples of this concept. Figures 3a and 3d demonstrate how in certain halls the early response between 5-30 ms provides substantially more energy through reflections from the lateral angles. As illustrated in Fig. 9, lateral reflections emphasize frequencies in the 700Hz to 1kHz range, as well as high frequencies to some extent, more so than frontal reflections due to the directionality of binaural hearing.

Another advantage in this respect lies in the second-order lateral reflections via the bottom surfaces of side balconies. This effect can be observed in the early energy along the transverse plane in Figs. 3c and 4c. Viewed as through the listener toward the stage, the conventional lateral reflections from the side walls are joined by additional pair of reflections from moderately elevated directions. Such reflections may also complement the overall timbre by providing additional early energy with slightly different HRTF spectra.

The given examples of concert halls also show that in some cases the raked audience receives a direct sound which is amplified by the floor reflection (compare Figs. 3b and 3e). Correspondingly, distinct early reflections can be observed from ceiling, canopy, or reflectors (see Figs. 4d and 3d). In contrast to early lateral energy, reflections from the median plane may even reduce the binaural dynamic effect as the median plane incidence does not benefit from the directional emphasis of the dynamic-related frequency bands.

4.2 Objective metrics on the dynamic responsiveness with binaural hearing

The proposed phenomenon of the sound field affecting the perceived music dynamics has been lately studied from objective and subjective viewpoints. Until recently, differences in dynamic effects in concert hall acoustics were only anecdotal references. For instance, Beranek has characterized this as the hall supporting both quiet and powerful dynamics: "*listening is enhanced* *immeasurably by the dynamic response of the concert hall*" (Beranek, p. 509). Meyer, for one, has stated that the quality of *forte* is a sign of an acoustically good hall, while sound in quiet dynamics can be acceptable in otherwise poor halls as well (Meyer, 2009, p. 199). Importantly, these remarks not only suggest the existence of a non-linear effect but also connect responsiveness of the hall to dynamics with subjective preference and increased listening pleasure.

The degree of the responsiveness to music dynamics by different concert hall acoustics was explored by Pätynen et al. (2014) in a study which combined the components of source, medium, and receiver as illustrated in Fig. 1. The dependency of the music signal spectrum was derived from anechoic orchestra instrument measurements. The dynamic spectra of different instruments were mapped to the source positions respective to an orchestra layout in concert hall measurements. The spectra conveyed through the direct sound and the early reflected sound to a binaural receiver were analyzed separately. The excitation of the left and right ears by the respective spectra were estimated with the model by Moore and Glasberg (1987), and the total binaural excitation was subsequently calculated with the binaural summation formula in the manner of Sivonen and Ellermeier (2006). In short, the adopted approach provides an objective metric for the auditory excitation by the orchestra sound in varied dynamic levels in different parts of the spatial room impulse response. This method was applied to ten European concert halls which were measured with the identical calibrated system.

The main results revealed that the two hall typologies (rectangular shoebox, or non-rectangular) varied prominently in the proportion of the binaural excitation between the direct sound and the early reflections in contrasting music dynamics, hence the label "binaural dynamic responsiveness" (BDR). Expectedly, the effect was observed at the higher frequencies where the orchestra spectrum varies a lot between dynamic levels. When compared to the direct sound, the auditory excitation of the early reflections was greater in rectangular rooms than in non-rectangular rooms. In essence, this outcome provided an explanation for the dynamic effect by concert hall acoustics. As discussed earlier, preceding studies have identified the perceptual effect of the spatial responsiveness of the acoustic space (Kuhl, 1978; Marshall and Barron, 2001). Those effects can also be easily connected with the recent results, as the rectangular rooms are often characterized by the early reflections in the lateral plane which become proportionally more audible with increasing music dynamics. Further discussion on the geometrical features and their effect on the sound field related to music dynamics are presented by Pätynen and Lokki (2015) in an article which concentrates the inspection on few halls with different designs.

4.3 Perceptual attributes of music dynamics variation in room acoustics

Whereas the objective approach has proposed means to quantify the dynamic responsiveness, controlled listening experiments have aimed to chart the perception of the music dynamics in different acoustic conditions. A subjective listening test exploring the perceptual attributes in varied music dynamics in concert hall acoustics employed a setting similar to the study with objective metrics. The presented music stimuli consisted of short full orchestral excerpt containing a sudden, yet musically feasible increase in the music dynamics while the instrumentation and orchestral texture were kept constant. The signal was created by concatenating bars 41-43 (in piano) and 53-55 (in fortissimo) from the anechoic recordings of Bruckner's 8th symphony, II movement. Auralizations with room impulse responses from various concert hall measurements were reproduced to the listeners through a spatial 24-channel loudspeaker array in an acoustically treated listening space. The listening test employed paired comparison augmented by simultaneous free attribute elicitation. Hence, the subjects had to first decide which one of the presented two stimuli appeared to have a wider overall contrast between the contrasting music dynamics. Additionally, the subjects gave short descriptions on which perceptual degrees the stimuli changed differently. Together this data provides insight on the overall perceptual dynamic responsiveness as well as the perceptual qualities of music dynamics in concert halls.

The results reported by Pätynen and Lokki (2016b) suggest that the foremost perceptual attribute differentiating the rooms' responsiveness to music dynamics is the dynamic range itself. Out of circa one thousand trials, approximately one-fourth of the compared pairs demonstrated the dynamic range as the discriminating perceptual attribute reported by subjects. Another substantial attribute describing the effect by varied dynamics was the varying width of auditory image. The comparisons between six halls revealed that the traditional room geometries (i.e. shoe-boxes) tend to provide higher degree of perceived contrast between music dynamics, and the effect becomes emphasized with increased receiver distances. These findings are consistent with earlier discoveries with objective metrics by Pätynen et al. (2014). Of individual rooms, Vienna Musikverein and Berlin Konzerthaus appeared to exhibit particularly strong dynamic loudness and spatial effects among the concert halls included in the experiment (Pätynen and Lokki, 2016b). Correlation analysis for disentangling the salient connections between perceived dynamic responsiveness and traditional objective room-acoustic parameters showed that the strength (G) at high frequencies and the inverse of early binaural coherence [1-IACC] predicted best the high degree of dynamic effects. Contrary to expectations, the early lateral energy fraction did not show substantial correlation with the perceptual effect.

4.4 Dynamic responsiveness and emotional impact of listening

Earlier sections in this chapter presented the communication of dynamic variations as one ingredient for the emotional impact in music listening. Another study by Pätynen and Lokki (2016a) assumed a more general perspective on the concept of music dynamics. Whereas the study described above aimed to explore the perception of music dynamics, the second set of experiments focused on the emotional impact produced by listening to orchestral music in different acoustic conditions. The employed listening test methodology departed from the conventional experiments by applying psychophysiological measurements with participants during focused listening. In laboratory conditions with multi-channel spatial sound reproduction, measurement of electrodermal activity, i.e. variations of skin conductance due to autonomic nervous system activation, proved to be a feasible technique. The subjects were presented a sequence of 12 auralizations of a total of six concert halls via convolutions of anechoic orchestra material and spatial room impulse responses. The music signal was a positively looming passage from Beethoven's 7th symphony, first movement. Bars 11-18 of the piece begin softly with alternating woodwind chords and ascending major scales with strings, and eventually culminates in a prominent full-orchestra crescendo to the tonic A-major fortissimo. With its easily approachable tonal development and texture, and as one of the principal works in the orchestral literature, the passage is regarded a suitable excerpt for comparing the possible emotional effects between orchestra performances. The duration of each auralization was approximately 30 seconds, and they were presented in randomized order with 15 s silence in between without any listener interaction.

During the entire experiment of circa 12 min the skin conductance response (SCR) was recorded synchronously with the presented audio stimuli. Following the conventions for analysing this kind of measurement, the intensity of emotional responses in different acoustic conditions could be ranked according to the recorded psychophysiological data. Similarly to the study on perceived music dynamics, the acoustic characteristics of rectangular halls showed a distinct advantage also in eliciting stronger emotional responses. Of individual rooms, the two highest mean SCR responses were found in the front positions at Vienna Musikverein and Berlin Konzerthaus. Listening to the performance in a non-rectangular hall instead of a shoebox room appeared to have a negative effect on the emotional intensity comparable to approximate doubling of the listening position's distance from the orchestra.

In order to gather more evidence for the emotional impact measured through psychophysiological responses, the set of 12 auralizations were also presented to the same subjects in a more traditional listening experiment with paired comparisons (Pätynen and Lokki, 2016a). Listeners were asked to choose the more impactful stimulus, and these self-reported results show the same general pattern as the SCR results.

To summarize, the binaural hearing combined with non-linear spectral excitation of musical instruments and different acoustic propagation paths and directions due to concert hall designs yields a complex setting. The communication of music dynamics elicits a wide range of perceptual attributes linked to the constantly varying nature of music. Therefore, the perception of room acoustics does not remain static for any signal, but instead it is influenced by the music itself. With the recent experiments on these topics, previously presented claims and impressions have found support from the research findings.

4.5 The dynamic variation at low frequencies combined with the seat-dip effect

Figure 8a illustrates a considerable amount of energy at low frequencies, even below 40 Hz, in the symphonic music. In addition, the large dynamic changes are the strongest at low frequencies (see Fig. 8b), often due to the orchestration as presented earlier. Therefore, it is reasonable to briefly discuss the low frequencies in concert halls, although, at low frequencies the binaural hearing does not play any role for auditory impression. Nevertheless, Marshall and Barron (2001) mention that the perceived width of the sound source is wider if the music is loud and the bass is strong.

As the sound travels from the stage over the seats at near grazing angles (below 15^o) at low frequencies an excess attenuation has been measured already 50 years ago by Sessler and West (1964); Schultz and Watters (1964). The phenomenon is called as the seat-dip effect, and it is prominent for the direct sound and some early reflections. The seat-dip effect is a combination of several phenomena, but mainly diffractions from the seat rests and reflections from the floor interfere destructively with the direct sound (Ishida, 1995). Bradley (1991) suggested that the main frequency of the attenuation depends on the dimensions of the seats. Tahvanainen et al. (2015) confirmed Bradley's results by analyzing measured data from 10 different concert halls and they found also that the inclination of the floor and the seat type affect to the range of attenuated frequencies. Based on these measurements, the seat-dip effect can be presented with two profiles, wide band attenuation centered around 150-300 Hz, and narrow band attenuation centered around 100 Hz (Tahvanainen et al., 2015). These two cases are indeed seen in Fig. 4g-h. In the Konzerthaus, with flat audience floor and lightweight open seats, the main attenuated frequency is about 200 Hz and the attenuation spans up to 1 kHz. However, the frequency response is filled in by the later reflected energy so that the attenuation due to the seat-dip effect has disappeared already 200ms after the direct sound. In contrast, the Philharmonie with raked audience area and seats that do now allow the underpass of sound, has narrow seat-dip around 100 Hz and it is not corrected at all with the later reflected energy.

Together with the early reflected energy, the seat-dip in these two halls most probably contributes the frequency response below 100 Hz as well. In both halls it can be seen a positive interference at about one octave lower than the seat-dip frequency. In the Konzerthaus, the seat-dip frequency is one octave higher than in the Philharmonie and therefore this positive interference is also at higher frequencies. In addition, the emphasis on frequencies below 100 Hz is much stronger in the Konzerthaus, and we assume that this boost is related to the seat-dip effect. Such low frequency behavior is very important for the dynamic variation in music, as illustrated in Fig. 8b, the largest dynamic differences are between 35 and 100 Hz. Therefore, it is reasonable to assume that in the Konzerthaus the dynamics at low frequencies is larger than in the Philharmonie (Lokki and Pätynen, 2019).

5 Late reverberation contributes to loudness, envelopment, spaciousness, and timbre

As discussed in Section 3, it is clear that early reflections have a crucial role on the perceived acoustics of a concert hall. Indeed, Haapaniemi and Lokki (2014) found that the characteristics of a hall are recognized within the first 80 ms of an impulse response. They investigated measured real concert halls with multichannel auralization system so that while the late reverberation was same in each rendering the first 80ms was from different halls. Subjects had to choose out of four different renderings, which hall was the same as a reference and the recognition rate was close to 100%. If the first 80ms was kept constant (i.e. from one hall) and the late reverberation tails were from different halls, the subjects considered the task much harder, however still the recognition rate was about 80%. The main attributes that subjects used in recognizing the halls were timbre and auditory width.

Regardless of the big role of early reflections the characteristics of late reverberation, i.e., length of decay, spectral coloration, and spatial distribution, are very important for the quality of music. Indeed, reverbration ties music segments together and blends instruments together, making music more enjoyable. When people are comparing different concert halls, they always pay attention to reverberance and aspects that the reverberation causes, such as envelopment, loudness, and width to some extent (Lokki et al., 2016). Reverberation also color the sound and in particular the strength of high frequencies in reverberation influences to perceived brightness, dynamics and brilliance (Pätynen and Lokki, 2015). Moreover, often listeners perceive a certain "warmth" if there are enough low frequencies (down to 20 Hz) and a reasonably long low frequency decay.

From the binaural point of view, the major perceptual effect of late reverberation is related to envelopment, i.e., how well a listener feels surrounded by music. Again, in our studies (e.g. Lokki et al. (2016)) we have found out that the hall type has great influence on the envelopment. In the shoe-box halls with flat floor audience areas, there is typically good envelopment and late reverberation is uniformly distributed around the listener, see Figs. 3 and

4. In contrast, in vineyard halls with raked audience areas the envelopment is reduced and late reverberation is less uniformly spread. The perceptual impression is often that in the halls with highly raked audience areas the listeners are ”looking at the music” (Pätynen and Lokki, 2015) while in the flat floor halls the listeners feel ”inside the music”.

There is not much research on the importance from the perception point of view of the spatial distribution of the late reverberation. Recently, Lachenmayr et al. (2016) studied the effect of the direction of reverberation to the perceived envelopment. The results were not very clear, but they propose that when reverberation from the side or above is reduced the feeling of envelopment is also reduced. The frontal or rear reverberation had minor, although important effects on envelopment. In the second listening test, the listeners adjusted one late component of the sound field blindly to the preference level. The results showed that when the hall lacks reverberant energy in a certain direction, subjects raised energy at that particular direction to a level so that reveberation is more or less uniform in the end. For example, in the Philharmonie the lacking reverberation behind the listener was compensated more than in the Konzerthaus.

Another contribution to the directional late reverberation has been present by Kahle (2016). He concluded that excessive reflections and reverberation from frontal directions can have a negative influence on orchestral balance and on on-stage hearing conditions for musicians. Kahle (2016) also describes several halls in which the openness and quality of sound was increased when the back wall of the stage and choir balcony was covered with absorptive material. Thus reducing the level of frontal reflections and reverberation increased the quality of sound both on the stage and in the audience area.

6 Conclusions

Concert halls are often studied by measuring impulse responses and computing room acoustical parameters based on the measurements. By definition, the impulse responses are linear and time-invariant, and the spatial aspects of sound field could be investigated using directional microphones or a dummy head as a measurement device. However, they do not tell anything about the level dependent phenomena, such as source broadening according to level or the effect of room acoustics to the perceived dynamics of music.

This chapter explains in detail, why the traditional means of analyzing concert hall acoustics are insufficient. It is explained how the level dependent spectra of musical instruments and sensitivity of human hearing are combined to room acoustics. In particular, the phenomena related with binaural hearing and perception of music in a concert hall are explained.

Acknowledgement

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

References

M. Barron, A.H. Marshall, Spatial impression due to early lateral reflections in concert halls: The derivation of a phisical measure **77**(2), 211–232 (1981)

M. Barron, Subjective study of british symphony concert halls. Acta Acustica united with Acustica **66**(1), 1–14 (1988)

L. Beranek, 2nd edn. (Springer, New York, USA)

L. Beranek, *Music, Acoustics, and Architecture* (John Wiley Sons, New York, NY, USA, 1962)

L.L. Beranek, Concert hall acoustics-1992. The Journal of the Acoustical Society of America **92**(1), 1–39 (1992)

J. Blauert, W. Lindemann, Auditory spaciousness: Some further psychoacoustic analyses. Journal of the Acoustical Society of America **80**(2), 533–542 (1986)

J. Bradley, Some further investigations of the seat dip effect. Journal of the Acoustical Society of America **90**(1), 324–333 (1991)

L. Cremer, H. Müller, *Principles and Applications of Room Acoustics* (Applied Science Publishers, London, England, 1982)

A. Farina, Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique, in *the 108th Audio Engineering Society (AES) Convention*, Paris, France, 2000. preprint no. 5093

E. Green, E. Kahle, Dynamic spatial responsiveness in concert halls, in *the Tenth International Conference On Auditorium Acoustics*, vol. 40 (Institute of Acoustics, Hamburg, Germany, 2018), pp. 226–236

A. Haapaniemi, T. Lokki, Identifying concert halls from source presence vs room presence. Journal of the Acoustical Society of America **135**(6), 311–317 (2014). http://dx.doi.org/10.1121/1.4879671

R.J. Hawkes, H. Douglas, Subjective acoustic experience in concert auditoria. Acta Acustica united with Acustica **24**(5), 235–250 (1971)

J. Hyde, Acoustical intimacy in concert halls: Does visual input affect the aural experience? (multisensory integration and the concert experience), Technical Report 010301, Paul S. Veneklasen Research Foundation, Santa Monica, California USA, 2003

J.R. Hyde, Discussion of the relation between initial time delay gap (ITDG) and acoustical intimacy: Leo Beranek's final thoughts on the subject documented, in *the Tenth International Conference On Auditorium Acoustics*, Hamburg, Germany, 2018, pp. 245–253

K. Ishida, Investigation of the fundamental mechanism of the seat-dip effect – Using measurements on a parallel barrier scale model. Journal of the Acoustical Society of Japan (E) **16**(2), 105–114 (1995)

ISO 3382-1, *Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces*, International Standards Organization, 2009

E. Kahle, Acoustic feedback for performers on stage – return from experience, in *International Symposium on Musical and Room Acoustics (ISMRA)*, La Plata, Buenos Aires, Argentina, 2016

E. Kahle, Halls without qualities – or the effect of acoustic diffusion, in *the Tenth International Conference On Auditorium Acoustics*, Hamburg, Germany, 2018, pp. 169–173

E. Kahle, Validation d'un modèle objectif de la perception de la qualité acoustique dans un ensemble de salles de concerts et d'opéras, PhD thesis, 1995

W.V. Keet, The influence of early lateral reflections on the spatial impression, in *Proc. 6th International Congress on Acoustics*, vol. 3, Tokyo, Japan, 1968, pp. 53–56

G.S. Kendall, The decorrelation of audio signals and its impact on spatial imagery. Computer Music Journal **19**(4), 71–87 (1995)

L. Kirkegaard, T. Gulsrud, In search of a new paradigm: How do our parameters and measurement techniques constrain approaches to concert hall design? Acoustics Today **7**(1), 7–14 (2011)

W. Kuhl, Räumlichkeit als Komponente des Raumeindrucks. Acta Acust United Ac **40**(3), 167–181 (1978)

R. Kürer, G. Plenge, H. Wilkens, Correct Spatial Sound Perception Rendered by a Special 2-Channel Recording Method, in *the 37th Audio Engineering Society Convention*, 1969

A. Kuusinen, T. Lokki, Auditory distance perception in concert halls and the origins of acoustic intimacy, in *The 9th International Conference on Auditorium Acoustics*, Paris, France, 2015

A. Kuusinen, J. Pätynen, S. Tervo, T. Lokki, Relationships between preference ratings, sensory profiles, and acoustical measurements in concert halls. Journal of the Acoustical Society of America **135**(1), 239–250 (2014). http://dx.doi.org/10.1121/1.4836335

W. Lachenmayr, J. Pätynen, Influence of acoustics on emotional impact of music in Konzerthaus and Philharmonie Berlin, in *DAGA 2016, 42. Jahrestagung für akustik*, Aachen, Germany, 2016, p. 79

W. Lachenmayr, A. Haapaniemi, T. Lokki, Direction of late reverberation and envelopment in two reproduced Berlin concert halls, in *the AES 140th International Convention*, Paris, France, 2016

T. Lokki, Tasting music like wine: Sensory evaluation of concert halls. Physics Today **67**(1), 27–32 (2014). http://dx.doi.org/10.1063/PT.3.2242

T. Lokki, Why is it so hard to design a concert hall with excellent acoustics?, in *the Second Australasian Acoustical Societies' Conference (Acoustics 2016)*, Brisbane, Australia, 2016. Invited Plenary lecture

T. Lokki, J. Pätynen, The acoustics of a concert hall as a linear problem. Europhysics News **46**(1), 13–17 (2015). http://dx.doi.org/10.1051/epn/2015102

T. Lokki, J. Pätynen, Objective analysis of the dynamic responsiveness of concert halls. Acoustical Science and Technology (2019). Submitted.

T. Lokki, J. Pätynen, S. Tervo, S. Siltanen, L. Savioja, Engaging concert hall acoustics is made up of temporal envelope preserving reflections. Journal of the Acoustical Society of America **129**(6), 223–228 (2011)

T. Lokki, J. Pätynen, A. Kuusinen, S. Tervo, Concert hall acoustics: Repertoire, listening position and individual taste of the listeners influence the qualitative attributes and preferences. Journal of the Acoustical Society of America **140**(1), 551–562 (2016)

T. Lokki, J. Pätynen, A. Kuusinen, H. Vertanen, S. Tervo, Concert hall acoustics assessment with individually elicited attributes. Journal of the Acoustical Society of America **130**(2), 835–849 (2011)

T. Lokki, J. Pätynen, A. Kuusinen, S. Tervo, Disentangling preference ratings of concert hall acoustics using subjective sensory profiles. Journal of the Acoustical Society of America **132**(5) (2012). 3148-3161

D.A. Luce, Dynamic Spectrum Changes of Orchestral Instruments, in *Audio Engineering Society Convention 51*, 1975

A.H. Marshall, A note on the importance of room cross-section in concert halls. Journal of Sound and Vibration **5**(1), 100–112 (1967)

A.H. Marshall, Levels of reflection masking in concert halls. Journal of Sound and Vibration **7**(1), 116–118 (1968)

A.H. Marshall, On the architectural implications of diffusing surfaces, in *the Tenth International Conference On Auditorium Acoustics*, Hamburg, Germany, 2018, pp. 618–624

A.H. Marshall, M. Barron, Spatial responsiveness in concert halls and the origins of spatial impression. Applied Acoustics **62**(2), 91–108 (2001)

J. Meyer, *Acoustics and the Performance of Music* (Springer, New York, NY, USA, 2009)

H. Møller, Fundamentals of binaural technology. Applied Acoustics **36**(3), 171–218 (1992)

B.C.J. Moore, B.R. Glasberg, Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. Hearing Research **28**, 209–225 (1987)

K. Oguchi, M. Quiquerez, Y. Toyota, Acoustcal design of Elbphilharmonie, in *the Tenth International Conference On Auditorium Acoustics*, Hamburg, Germany, 2018, pp. 89–96

H. Owen, *Music theory resource book* (Oxford University Press, New York, NY, USA, 2000)

J. Pätynen, A Virtual Symphony Orchestra for Studies on Concert Hall Acoustics, PhD thesis, Aalto University School of Science, 2011

J. Pätynen, T. Lokki, The acoustics of vineyard halls, is it so great after all? Acoustics Australia **43**(1), 33–39 (2015)

J. Pätynen, T. Lokki, Concert halls with strong and lateral sound increase the emotional impact of orchestra music. Journal of the Acoustical Society of America **139**(3), 1214–1224 (2016a)

J. Pätynen, T. Lokki, Perception of music dynamics in concert halls. Journal of the Acoustical Society of America **140**(5), 3787–3798 (2016b)

J. Pätynen, V. Pulkki, T. Lokki, Anechoic recording system for symphony orchestra. Acta Acustica united with Acustica **94**(6), 856–865 (2008)

J. Pätynen, S. Tervo, T. Lokki, Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses. Journal of the Acoustical Society of America **133**(2), 842–857 (2013). http://dx.doi.org/10.1121/1.4770260

J. Pätynen, S. Tervo, , T. Lokki, Amplitude panning decreases spectral brightness with concert hall auralizations, in *Proc. 55th Audio Eng. Soc. conference, Helsinki, Finland* (Audio Eng Soc, New York, NY, USA, ???, 2014). Paper no. 49

J. Pätynen, S. Tervo, P.W. Robinson, T. Lokki, Concert halls with strong lateral reflections enhance musical dynamics. Proceedings of the National Academy of Sciences of the United States of America (PNAS) **111**(12), 4409–4414 (2014)

B.N.J. Postma, B.F.G. Katz, The influence of visual distance on the room-acoustic experience of auralizations. The Journal of the Acoustical Society of America **142**(5), 3035–3046 (2017). doi:10.1121/1.5009554

N. Rimsky-Korsakov, *Principles of Orchestration (tr. E. Agate)* (Editions Russes de Musique, Berlin, Germany, 1922)

P. Robinson, J. Pätynen, T. Lokki, The effect of diffuse reflections on spatial discrimination in a simulated concert hall. Journal of the Acoustical Society of America **133**(5), 370–376 (2013a). http://dx.doi.org/10.1121/1.4798648

P. Robinson, J. Pätynen, T. Lokki, H.S. Jang, J.Y. Jeon, N. Xiang, The role of diffusive architectural surfaces on auditory spatial discrimination in performance venues. Journal of the Acoustical Society of America **133**(6), 3940–3950 (2013b). http://dx.doi.org/10.1121/1.4803846

F. Rumsey, S. Zielinski, R. Kassier, S. Bech, On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality. Journal of the Acoustical Society of America **118**(2), 968–976 (2005)

W.C. Sabine, Reverberation: Introduction. The American Architect (1900)

M. Sams, R. Hari, J. Rif, J. Knuutila, The human auditory sensory memory trace persists about 10 sec: Neuromagnetic evidence. Journal of Cognitive Neuroscience **5**(3), 363–370 (1993). doi:10.1162/jocn.1993.5.3.363

M.R. Schroeder, D. Gottlob, K.F. Siebrasse, Comparative study of european concert halls: correlation of subjective preference with geometric and acoustic parameters. The Journal of the Acoustical Society of America **56**(4), 1195–1201 (1974)

T. Schultz, B. Watters, Propagation of sound across audience seating. Journal of the Acoustical Society of America **36**(5), 885–896 (1964)

G. Sessler, J. West, Sound transmission over theatre seats. Journal of the Acoustical Society of America **36**(9), 1725–1732 (1964)

V.P. Sivonen, W. Ellermeier, Directional loudness in an anechoic sound field, head-related transfer functions, and binaural summation **119**(5), 2965–2980 (2006)

H. Tahvanainen, J. Pätynen, T. Lokki, Analysis of the seat-dip effect in twelve european concert halls. Acta Acustica united with Acustica **101**(4), 731–742 (2015)

S. Tervo, J. Pätynen, A. Kuusinen, T. Lokki, Spatial decomposition method for room impulse responses. Journal of the Audio Engineering Society **61**(1/2), 16–27 (2013). http://www.aes.org/e-lib/browse.cfm?elib=16664

R. Wettschureck, Über die Abhängigkeit raumakustischer Wahrnehmungen von der Lautstärke (On the dependence of room acoustic perception by sound level), PhD thesis, TU Berlin, Germany, 1976

A. Wilska, Studies on directional hearing; Untersuchungen über das Richtungshören, Technical report, 2011. 978-952-60-4061-5 (electronic). http://legacy.spa.aalto.fi/publications/WilskaThesis/