DYNAMIC SPATIAL RESPONSIVENESS IN CONCERT HALLS

E Green Kahle Acoustics, Brussels, Belgium egreen@kahle.be
E Kahle Kahle Acoustics, Brussels, Belgium ekahle@kahle.be

1 INTRODUCTION

1.1 (Dynamic) Spatial Responsiveness

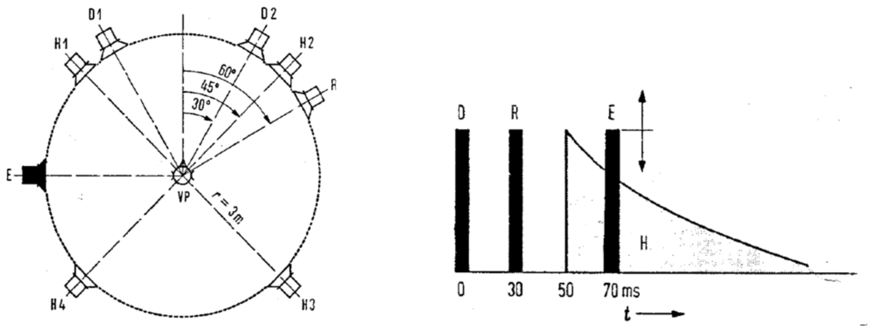
Musicians and conductors consider the concert hall part of their instrument – it is a tool for musical expression with the best concert halls enhancing the range of expression. The deliberate variation of loudness – in other words changes in musical dynamics – is a key ingredient in most musical composition and performance. Without both strong and subtle dynamic changes, much musical and emotional intensity is lost.

The connection between dynamics and spatial impression or room presence was observed in some of the earliest experiments in psychoacoustics. Marshall[1] writes in 1966 that "as a property of the hall, [spatial impression] relates to loudness attributes…for the listener, it generates a sense of envelopment in the sound and of direct involvement with it" with similar observations from Keet[2] and Barron[3]. Writing in 1978 Kuhl[4] states that (translated from German) "[spatial impression] is only, if at all, present in forte passages. Musical dynamics therefore not only influence the loudness, but with a sufficient spatial responsiveness, they increase the involvement of the room in the music. For the lay-listener, who is perhaps not even aware of this effect, this spaciousness is an unconscious yet pleasant experience." Dynamic changes can therefore generate an additional dimension of spatial responsiveness for the listener.

Since the work in the 1960s and 70s, the importance of dynamics and this factor in the enhancement of musical expression provided by the best-liked concert halls has been somewhat overlooked. Recently the group at Aalto University have taken up this topic anew[5,6,7,8] demonstrating that dynamic responsiveness is related to changes in instrument frequency responses and directivity at louder dynamic levels: due to the greater binaural sensitivity to high frequencies at lateral angles, an increased spatial impression results from the additional higher harmonics that are generated when instruments are played loudly. However, little-known research published by Wettschurek[9] in 1978 demonstrates that the connection between dynamics, loudness and spatial impression may be more fundamental and is related to how reflections from different directions are masked or unmasked at different overall loudness (dynamic) levels.

1.2 Research by Wettschurek

In Wettschurek's thesis[9], he describes an experiment to determine the perception threshold of a test reflection dependent on overall listening levels. The experiment is shown diagrammatically in Figure 1. A synthetic sound field was generated in an anechoic environment consisting of the direct sound, a 2-second reverberation starting at 50ms and two discrete reflections: one reflection being a fixed lateral reflection at 60° and 30ms delay and the other a test reflection at 70ms and of variable level. The direct sound, lateral reflection and reverberation levels were set to achieve a direct-to-reverberant ratio of 0dB. With anechoic speech used as source material, the overall sound level at the listening position ("listening level") was set to multiples of 5dB and then the level of the test reflection was adjusted to establish the perception threshold of this reflection.



D = direct sound as phantom source via loudspeakers D1, D2
R = fixed lateral early reflection, loudspeaker R
H = 2 second reverberation (500Hz value) starting at 50ms,
reproduced via 4 loudspeakers H1 – H4
E = 70ms reflection with adjustable level and location, loudspeaker E

Figure 1 – Diagrammatic representation of an experiment to determine the perception threshold of a reflection E in the presence of direct sound D, a lateral reflection R and reverberation H. After Wettschurek[9].

The results of this experiment, shown in Figure 2, exhibit a number of notable features;

- reflections from behind (“Hinten”) have the highest perception threshold at all listening levels;
- for all reflection directions, the perception threshold decreases almost linearly with increasing listening level until approximately 40dB;
- at listening levels higher than 40dB the sensitivity to reflections from the front (“Vorne”) begins to plateau. Sensitivity to reflections from behind (“Hinten”) plateaus at levels above approximately 60dB;
- above 40dB however, the sensitivity to reflections from the side continues to increase approximately linearly with listening level;
- by 80dB the sensitivity to reflections from the side is almost 10dB greater than for reflections from the front or rear which at this listening level have an almost equal sensitivity of 8dB below the direct sound level.

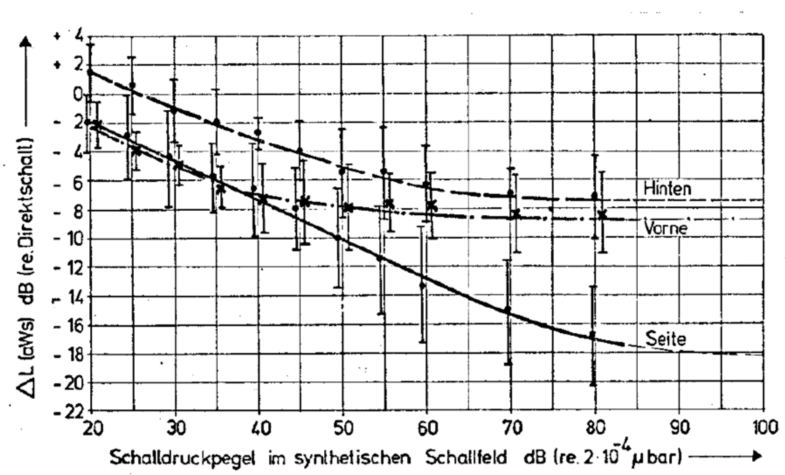


Figure 2 – Perception threshold relative to the direct sound for a 70ms reflection as a function of the sound pressure level at the listening position for 3 directions of arrival: Rear (“Hinten”, azimuth 180°), front (“Vorne”, azimuth 0°) and Side (“Seite”, azimuth 90°) (After Wettschurek[9], Figure 1.11).

Wettschurek's results provide a strong basis for understanding the increase of room presence with increasing listening level, independent of any changes in the source spectrum:

- At the lowest dynamic levels, most reflections would not be perceived and spatial impression would be extremely weak or non-existent. The direct sound would dominate with the sources on stage being clearly localizable;
- As level increases through a crescendo, reflections from the front and side having a lower threshold than those from the rear would be the first to be perceived, with the expected subjective effect being an initial increase in the Apparent Source Width (ASW), as observed and subsequently measured by Keet[2].
- As level increases further, the perception threshold for reflections from all directions of arrival continues to decrease. As a result, room presence increases, especially as reflections from behind also become audible – this corresponds with listening experience in concert halls and the comments of Marshall[1], Kuhl[4] and others.
- Above a listening level of 60dB, the thresholds for reflections from the front and rear plateau while the threshold for side reflections continues to decrease. If reflection paths from the side exist, then these will become increasingly perceived as level increases resulting in a subjective increase in room presence.

Wettschurek refers to this journey through a musical crescendo as the room "waking up". However, this assumes that similar perception threshold relationships exist for all reflection delays and for music as well as speech. Measurements to corroborate this were not published by Wettschurek.

2 REFLECTION THRESHOLDS WITH SPEECH AND MUSIC

2.1 Experiment with Speech

In order to understand how perception thresholds for reflections may vary for different delay times, directions of arrival and for music, Wettschurek's experiment was duplicated using a binaural method. The sound field shown in Wettschurek's experiment (Figure 1) was duplicated in a virtual Ambisonics system programmed using MaxMSP and SPAT[10] and then reproduced binaurally over headphones using the subject's own measured HRTF. Measurements of the overall listening level were made under the headphones using a calibrated sound level meter. For the first test, it was attempted to reproduce Wettschurek's experiment as closely as possible using anechoic speech and a 70ms delayed test reflection. The resulting thresholds are shown in Figure 3 with Wettschurek's measurements (as per Figure 2) adjacent for reference.

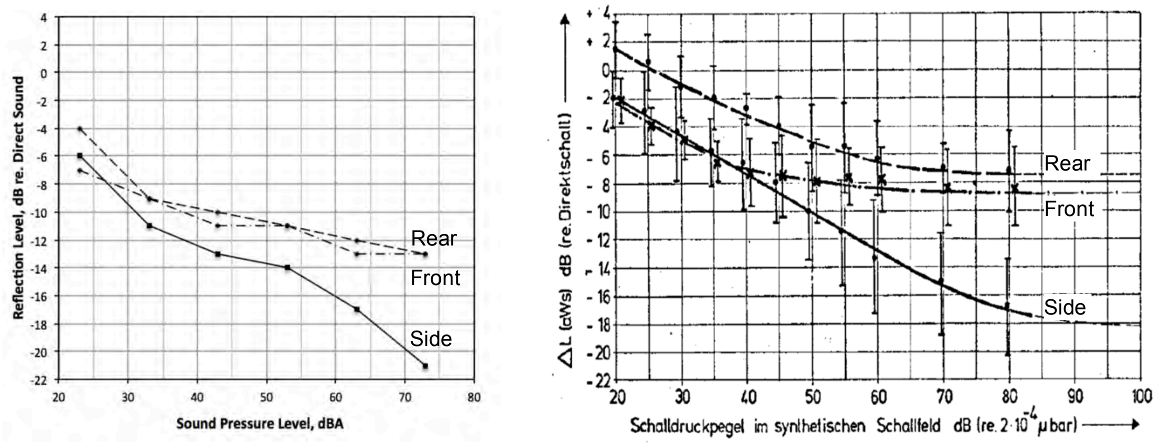


Figure 3 – Reflection perception thresholds (reflection level re. direct sound) for ***speech*** *for a 70ms delayed reflection. Left: measured using binaural reproduction. Right: measured using a loudspeaker system (after Wettschurek[9]).*

Although the perception thresholds with binaural reproduction are lower than those measured by Wettschurek, the result is encouraging as the general relationship between the frontal, rear and side reflections is similar. There are however clear differences: In the binaural case, the rear reflection threshold follows the front reflection closely, which could be attributed to known issues of median plane confusion with binaural reproduction. Furthermore, it is known that the source material can have a significant effect on measured thresholds. In addition, the primary subjects for the tests were experienced listeners who typically also exhibit lower thresholds than lay-listeners. Nevertheless, it is clear that the thresholds determined using binaural reproduction show similar trends to the (arguably more reliable) loudspeaker tests. For reference, the absolute level of the test reflection is in all cases significantly above the threshold of hearing, therefore the measured perception threshold is an effect of masking.

2.2 Reflection Perception Thresholds with Music

Experiments were carried out to determine reflection perception thresholds for music using an anechoic recording of solo cello. This source material was chosen for its even sound level throughout the recording and *legato* musical expression to contrast with the consonant-rich speech source used in the previous test. Tests were made for reflections delayed by 70ms, 50ms and 40ms. All other settings including the fixed lateral reflection, reverberation time and direct-to-reverberant ratio were maintained as previously. The results are shown in Figure 4 below.

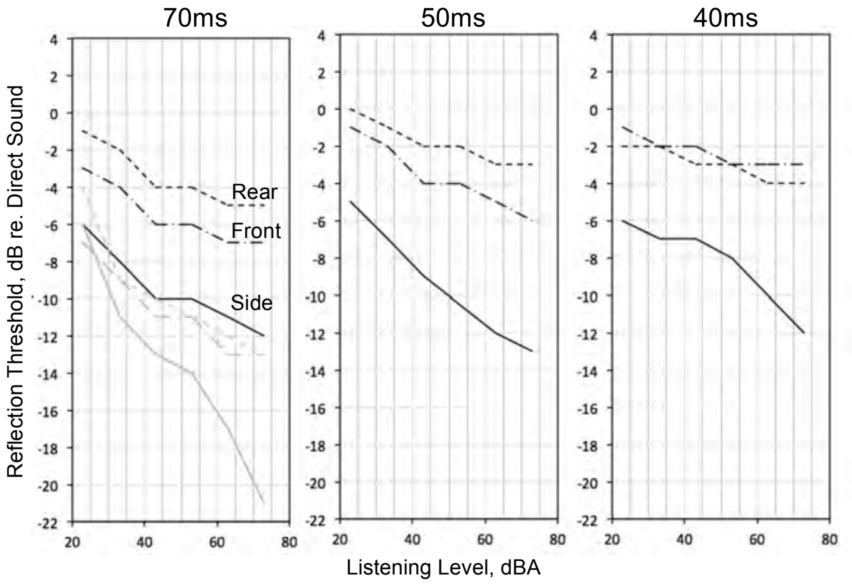


*Figure 4 – Reflection perception thresholds (reflection level re. direct sound) for **music**, dependent on reflection delay, reflection direction and overall listening level. For reference the binaurally measured reflection thresholds for speech and 70ms delay time are shown in grey on the left panel.*

Figure 4 shows that, as one might expect, the 70ms perception thresholds are in general higher for music, reflecting the longer "cognitive integration time" for musical material compared to speech. The 70ms reflection thresholds for side reflections are nevertheless 4-5dB below those for frontal and 6-7dB below those for rear directions. These are slightly less than with speech at higher listening levels, but the difference at low listening level (under 40dBA) is greater with music than with speech. At 50ms delay time, the threshold for side reflections is almost identical to that at 70ms, while the thresholds for frontal and rear reflections are higher. In general, over the range of listening levels measured (approx. 25-75dB), the sensitivity to side reflections increases by 6-7dB while the increase for frontal and rear reflections is 4dB at most, with an increase of only 1-2dB at 40ms delay time.

2.3 Subjective Effects of Changing Listening Level

While carrying out the perception threshold measurements, it was clear that reflections at different delay times and directions of arrival had very different subjective effects but also that the overall listening level had an influence on the subjective quality, even when the reflection level was close to the perception threshold. Three main changes in the quality of the sound were perceived: (1) tone colouration, (2) changes to the source spatial extent (reflection experienced as “fused” with the source) and (3) changes to the room presence (reflection experienced as separate to the source).

It has long been known that one of the primary effects of reflections with short delay times is to change the tone colour[3]. This was indeed the case in this experiment: when listening for whether a short-delay-time reflection was perceptible, “comb filtering” was the primary quality that was identifiable. The comb filtering effect was strongest for reflections from the front and rear, while for a reflection from the side, depending on delay and listening level, tone colouration would be mixed with a simultaneous change in source spatial extent.

At 70ms delay time, the subjective effect of the test reflection, once it became perceptible, was generally a change in spatial extent or room presence for side reflections and perceived distance for frontal and rear reflections. Most interesting however was the change in subjective impression of a 70ms reflection with overall listening level. At a certain loudness level, the reflection could switch from “belonging to the source” to “belonging to the room”. There is increasing evidence to support the idea that our auditory system segregates sounds into various “auditory streams”: in concert hall listening one description is that there is a source (foreground) stream and a room (background) stream[11]. This experiment indicates that there are complex factors at play as to whether a sound is allocated to the source stream or room stream (or indeed to both streams, or no stream at all when below threshold): not only is delay time a factor (and that there is not a simple cut-off time at 80ms), but also reflection direction of arrival and the overall listening level. Depending on the particular combination of these factors, sounds arriving before 80ms may not necessarily be integrated into the source stream but could be segregated into the room stream. Since our subjective impressions of clarity, proximity and intimacy seem to relate to aspects of stream segregation, and seem to relate to our ability to segregate sound energy into streams at all, it is important to gain a better understanding of these relationships and auditory/cognitive processes. It should also be noted that the number of participants in the experiment described here was very small, so more research is required to understand this effect in more detail and for larger groups of listeners.

Another question that arises from this measurement is that of the effect of the “unperceived” sound energy. As mentioned above, the reflected energy which is below the perception threshold is above the absolute threshold of hearing (it is clearly audible if the direct sound and other reflected energy is switched off in the experiment). Does this “unperceived” sound energy have other subjective effects that are not picked up in this experiment, or does our hearing system consider it to be noise?

One conclusion that can be drawn from the above tests with music is that, if lateral reflections are present, then as loudness increases, one would perceive more lateral sound and the room presence would change in response to the dynamics, resulting in an additional spatial dimension connected to changes in loudness. However, if lateral reflections are weak or absent, the small differences in perception threshold with listening level for frontal and rear reflections indicate that the room presence would tend to remain static: the subjective perception of the room acoustic would be less responsive to musical dynamics. Whether a dynamic or more static room presence is preferable and whether this has a significant bearing on the overall assessment of acoustical quality is to an extent a matter of taste, but also requires experimental results from a much larger sample than was used here.

3 DETECTING DYNAMIC SPATIAL RESPONSIVENESS

Typical objective measures related to spatial impression such as the Lateral Fraction (LF) and IACC integrate lateral sound energy over a time window and so do not take into account the specific time, level and direction of arrival of sound energy. In order see whether the components of the sound field that may contribute to dynamic spatial responsiveness can be detected and visualized, and to compare measurements from different concert halls, a spatial filter algorithm has been developed based on the thresholds measured above.

3.1 Dynamic Spatial Response Filter

The *dynamic spatial response filter* algorithm uses as its input measured first-order Ambisonic B-Format 3D Room Impulse Responses (3DRIR) and processes these to reveal only the reflected sound energy that would be subjectively perceived at a given listening level. Although concert halls and music are of primary interest in this paper, since Wettschurek's experiments used a larger number of subjects and are therefore considered more reliable, these reflection perception thresholds have been used in the filter. To determine the thresholds for delay times other than 70ms and for directions of arrival other than those measured by Wettschurek, trends in our own measurements and interpolation have been used.

The filtering process is shown diagrammatically in Figure 5. Firstly, the X, Y, and Z channels of the first-order B-format 3DRIR are used to calculate the Direction Of Arrival (DOA; azimuth = θ, elevation = ϕ) for each sample in the impulse response – the algorithms developed at Aalto University[12] are used for this step.

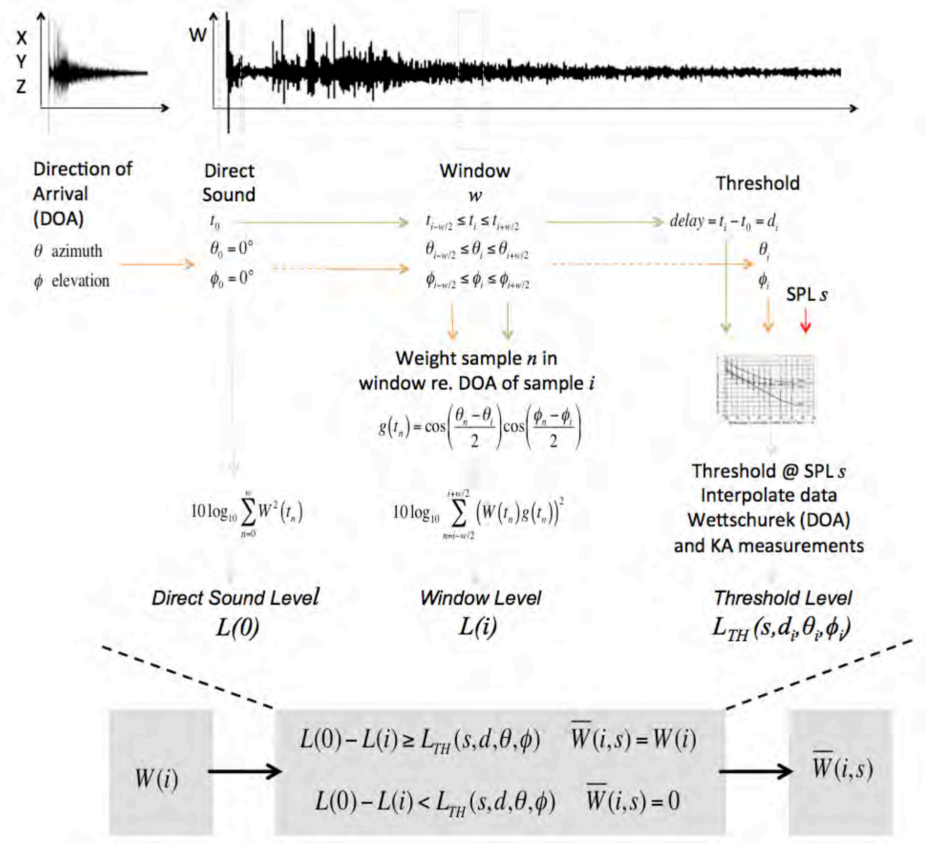


Figure 5 – Algorithm schematic for a Dynamic Spatial Response Filter

The direct sound level $L(0)$ is determined from the omnidirectional channel W for a user set window of w samples. Next, for each sample i of the impulse response, the sound level $L(i)$ is evaluated by integrating over a window of the same length w, centered on the sample of interest i. Since it is desired to collect energy in the window that can be considered to constitute a coherent reflection, each sample n in the window is weighted according to how similar the DOA is when compared to the sample i. Therefore, before integrating the energy in the window, a weighting g is applied to each sample n in the window according to its DOA relative to the DOA of the sample of interest. Samples with the same DOA as sample i are given a weighting 1, while samples arriving from the opposite DOA would be weighted 0:

$$g(t_n) = \cos\left(\frac{\theta_n - \theta_i}{2}\right)\cos\left(\frac{\emptyset_n - \emptyset_i}{2}\right) \tag{1}$$

The DOA and delay time of the sample i are used to establish the audibility threshold L_{Th} by interpolating the Wettschurek data and our own measurements at different delay times. If the level in the window $L(i)$ relative to the direct level L_0 is greater than the threshold L_{Th}, i.e. if the sample i can be considered to contribute to a perceivable reflection, then the sample i is passed through to the filtered output impulse response $\overline{W}$, otherwise it is "masked" (i.e. set to 0 in the output). This is repeated for all listening levels of interest s to give a series of modified 3DRIRs, one for each listening level.

3.2 Results from Measurements

The dynamic spatial response filter has been applied to 3DRIRs measured in the Nouveau Siècle concert hall in Lille, France. This hall started life in the mid-20th Century as a fan-shaped conference hall (Figures 6a & 6b, left) and was reconfigured in 2013, maintaining the stage and floor rake, to create a parallel-sided concert hall (Figures 6a & 6b, right). Subjective deficiencies for music uses with the fan-shaped conference hall were a lack of reverberation, envelopment and room presence, with the latter two attributed to insufficient lateral reflections. In the newly configured hall, the side walls were made parallel, while the addition of side balconies and modification of the ceiling design were intended to provide additional lateral reflections. Listening tests in the reconfigured hall indicate that an enveloping and present room sound has been generated, and that the acoustic now also exhibits (albeit moderate) dynamic spatial responsiveness. The question is, can these qualities be seen in the measured 3DRIRs?

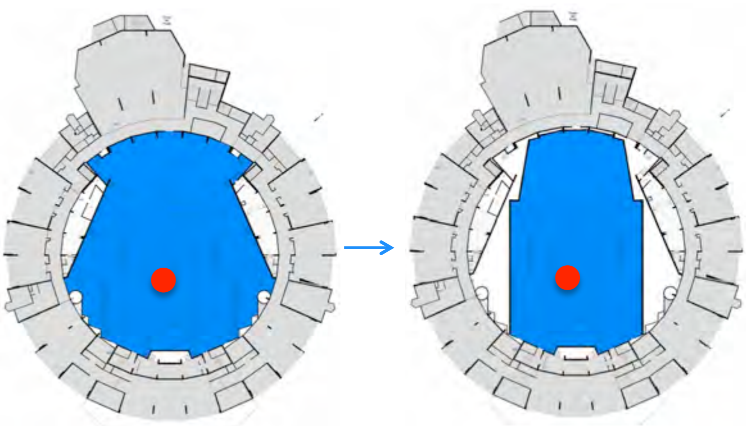


Figure 6a – Comparison plans of Nouveau Siècle, Lille. Left: fan-shaped conference hall before remodeling. Right: parallel-sided concert hall after remodeling. Measurement position for results in Figures 7 & 8 shown with red circle

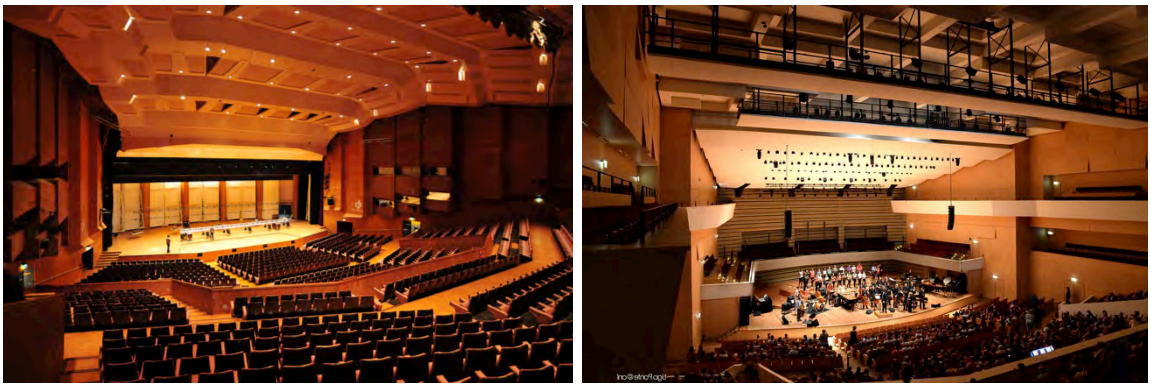


Figure 6b – Left: Lille Nouveau Siècle before remodeling. Right: after remodeling with new side walls, balconies, stage enclosure and canopy.

Figure 7 shows measured 3DRIRs for a seat 17m from the stage in the main seating area. The left-hand measurement is before the renovation and the right-hand plot after. The 3DRIRs are viewed from the top with the direct sound aligned to the “Front” direction. The arrival time is shown by colour with red shades indicating energy arriving up to 500ms, greens in the range 60-100ms and blue 5-50ms. While it is clear that after the remodeling there is more energy arriving from lateral directions, indicated by the higher sound level in the “Left” and “Right” directions, can filtering help to illustrate the subjectively perceived dynamic spatial responsiveness after remodeling?

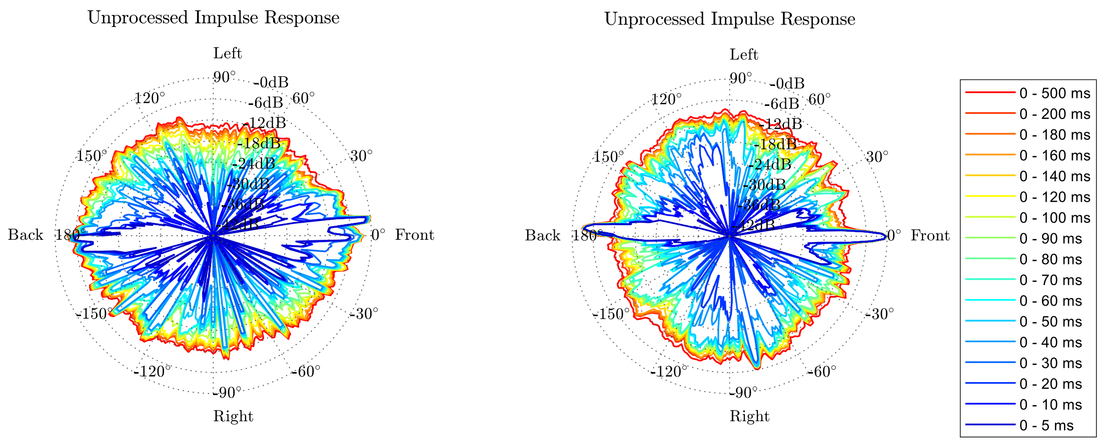


Figure 7 – Left: unprocessed impulse response for Nouveau Siècle before reconstruction and reconfiguration. Right: unprocessed impulse response for Nouveau Siècle after reconfiguration.

Figure 8 below shows the results of applying the dynamic spatial response filter to the Nouveau Siècle measurements (Figure 7) for listening levels of 50dB, 60dB and 70dB. As with the unprocessed plots in Figure 7, results from before the renovation project are shown on the left results after renovation are on the right. At 50dB listening level, before renovation, perceivable sound energy is detected from the rear and from the reflective stage enclosure, along with a few other “spikes” in the range 0-40ms. Post renovation, although less sound energy would seem to be perceivable at this 50dB listening level, a bundle of sound energy from the right-hand side is clearly evident along with some later sound after 70ms: the beginnings of noticeable source-broadening and running reverberation could be expected.

At 60dB listening level, pre-remodeling, the reflections already perceivable at 50dB are supplemented by a bundle of lateral energy at around 80ms. However, in the renovated room, the right-side reflections present at 50dB are joined by reflections from the left-front, left-rear and right-rear all in the range 60-80ms: subjectively it would be expected that the sound is now becoming enveloping, along with a strong apparent source width.

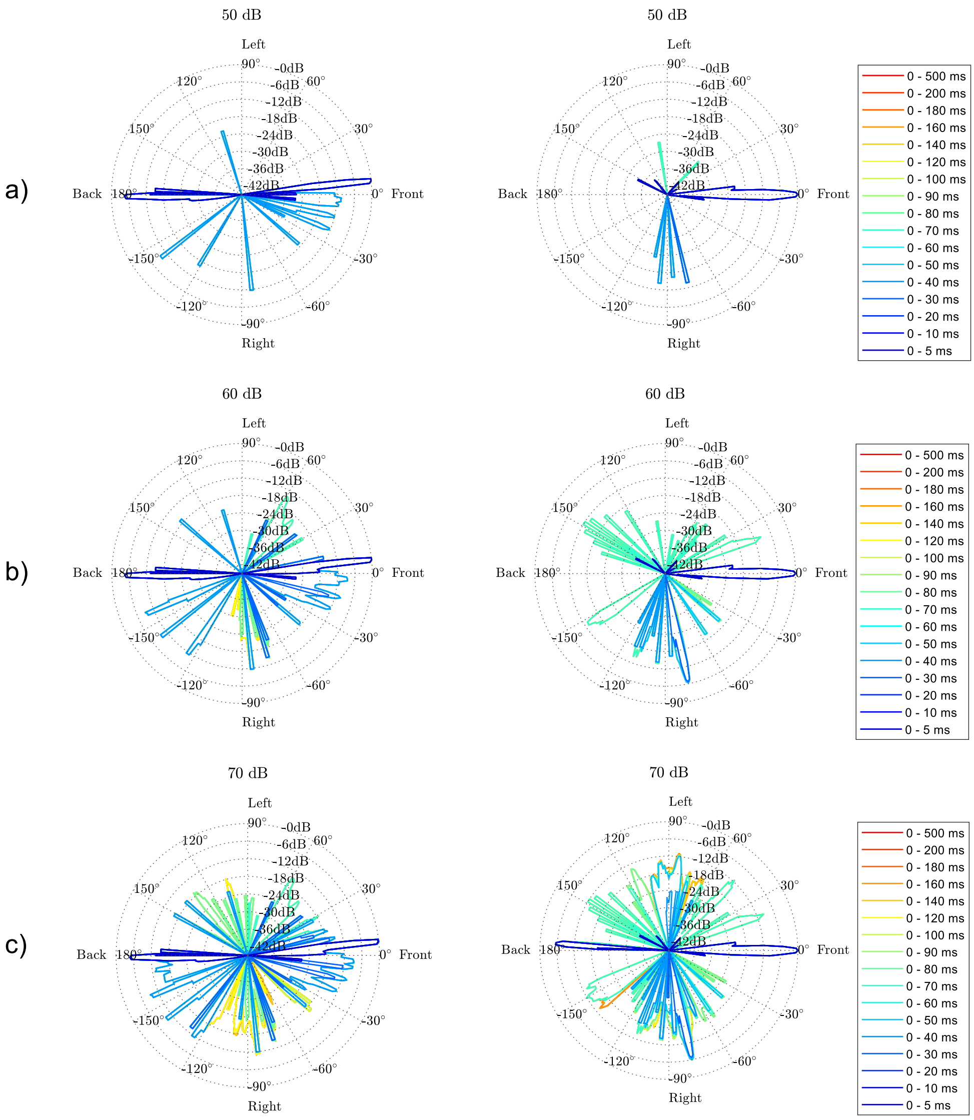


Figure 8 – Dynamic spatial response filter applied to measurements of Nouveau Siècle, Lille for different listening levels of a) 50dB, b) 60dB and c) 70dB. Left column: before renovation. Right column: after renovation.

At 70dB listening level, the spatial response in both cases is filled out, with energy in the range 100-200ms becoming apparent. However, in the renovated room, a new bundle of lateral energy around 50ms and from the left becomes apparent. In the renovated room, there is therefore a significant difference in the perceived spatial response due to the change from 60 to 70dB.

These plots indicate that, both before and after renovation, the perceivable room presence increases with listening level – however, it is clear that subjective room presence and envelopment should be much stronger after renovation since the strength of perceivable lateral reflections is around 6dB higher than before renovation. This magnitude of difference is not so evident in the full unprocessed impulse responses shown in Figure 7. Furthermore, before renovation the strongest reflections were clustered around the source or arrived from behind, and this attribute did not change with listening level, indicating an unresponsive spatial impression and room presence. After renovation, the spatial zone around the source is relatively free of reflected sound energy (a technique used in live sound mixing to enhance clarity): this can be attributed to the absorption of the new choir seating and the partially absorbing upstage wall.

Overall, through processing 3DRIRs with the dynamic spatial response filter, features of the impulse responses that indicate how the room presence is perceived at different listening levels can be extracted and visualized. In the case of the Nouveau Siècle concert hall, the extracted features correspond well with the subjective impression of changing room presence as loudness increases. The filter is currently relatively crude, based on a small dataset of thresholds and a very basic algorithm for attributing measured sound energy to coherent reflections. It is hoped that further work on both of these fronts will reveal further insights.

4 CONCLUSIONS

The responsiveness of a concert hall's acoustic to changes in loudness or dynamics has been identified as a key ingredient for high quality acoustics. Connections between musical and acoustical attributes – in this case where loudness changes generate subjective changes in spatial impression or room response – are believed to be important for excellent music acoustics. Research into such connections has however been rare. Expanding on the work by Wettschurek in the 1970s, measurements of the perception thresholds of early reflections have been made using a binaural virtual acoustics system. Both speech and music were used as a source and various reflection directions of arrival and delay times were tested. These measurements indicate that for music, perception thresholds for reflections from the front and behind vary little with overall listening level while for side reflections, the perception threshold decreases substantially with increasing listening level. This indicates that during a crescendo, increasing numbers of side reflections should become perceivable – the hall "wakes up" – while this is not expected for frontal or rear reflections (although there are other subjective factors, positive and negative, associated with reflections from in front and from behind). The findings correspond with listening experience, with those halls lacking lateral reflections also tending to exhibit poor dynamic spatial responsiveness.

It was also observed in the listening tests that, depending on the overall listening level, delay and direction of arrival, reflections could be allocated to the source auditory stream (fused with the source) or experienced as part of the room stream i.e. separate to the source. This is an important factor in our perception and rating of acoustical quality, and therefore warrants detailed further research with much larger groups of listeners than were used here. The question of how the "below threshold" sound energy influences our subjective impression is also pertinent.

Based on measured reflection perception thresholds, 3D room impulse responses have been filtered in order to visualize changes in perceivable reflected energy at difference listening levels. Initial tests of the so-called *dynamic spatial response filter* using measurements made in the Nouveau Siècle concert hall in Lille, France indicate that such filtering can aid in the discrimination of impulse responses from concert halls that exhibit dynamic spatial responsiveness. With refinement of the measured thresholds and filter algorithm, it is hoped that such analysis can become a reliable tool to aid in the prediction of situations exhibiting high acoustical quality.

5 REFERENCES

1. H. Marshall, A note on the importance of room cross-section in concert halls, Journal of Sound and Vibration 5(1) 100-112. (1966)
2. W.V. Keet. The influence of early lateral reflections on the spatial impression. Proceedings of the Sixth International Congress on Acoustics Vol. 3, E53–E56. (1968)
3. M. Barron., The subjective effects of first reflections in concert halls – the need for lateral reflections, Journal of Sound and Vibration 15(4) 475-494. (1971)
4. W. Kuhl. Räumlichkeit als Komponente des Raumeindrucks, Acta Acustica 40. (1978)
5. J. Pätynen et al, Concert halls with strong lateral reflections enhance musical dynamics, PNAS 11(12). (2014)
6. J. Pätynen, and T. Lokki, Perception of music dynamics in concert hall acoustics, Journal of the Acoustical Society of America, vol. 140, no. 5, pp. 3787-3798. (2016)
7. J. Pätynen, and T. Lokki, Concert halls with strong and lateral sound increase the emotional impact of orchestra music, Journal of the Acoustical Society of America, vol. 139, no. 3, pp. 1214-1224, (2016)
8. J. Pätynen, and T. Lokki, Towards a common parameter characterizing the dynamical responsiveness of a concert hall, Journal of the Acoustical Society of America, vol. 141, no. 5, (2017)
9. R. Wettschurek, Über die Abhängigkeit raumakustischer Wahrnehmung von der Lautstärke, PhD Thesis Technical University of Berlin. (1976)
10. SPAT acoustic simulation tool by IRCAM, http://forumnet.ircam.fr/product/spat-en2.
11. D. Griesinger, The Psychoacoustics of Apparent Source Width, Spaciousness and Envelopment in Performance Space, Acta Acustica, vol. 83. (1997)
12. https://www.mathworks.com/matlabcentral/fileexchange/56663-sdm-toolbox