PROCEEDINGS of ISMA 2019
International Symposium on Music Acoustics

13 - 17 September 2019 in Detmold, Germany

Timbre and Duration of Attack Depend on the Amount of Reverberation

Tor HALMRAST

University of Oslo/Musicology, Norway

ABSTRACT

The attack of a signal is of course best preserved if we hear the direct sound only, but that is not the case in a concert hall. Acousticians often remembers that long reverberation masks the entrance of the next onset, but more astonishing is the perceived and measured difference in timbre due to "smoothening"/"prolongation" of the attack, (also for a first note in a phrase). The paper will discuss general theory regarding how "diffuse field reverberation" influences the attack. The early response of a concert hall, is, however, seldom "diffuse". The paper discusses methods for measuring attack including Rise Time, Steepness etc. of the Integrated (Cumulative) Squared Step Response and also Spectral Flux, for both real halls and simulations (Odeon). Investigations were done for signals of different lengths and for different musical instruments which in itself have slow or fast note-onset. An important question: Is it possible to reduce the smoothening of the attack due to reverberation by adding early reflections? Preserving the attack is important also because listeners nowadays are used to recordings where any wanted amount of direct sound is mixed with a late, long and (too?) smooth, non-correlated, digital reverberation.

Keywords: Attack, Reverberation, Timbre

1 INTRODUCTION

When listening to musicians playing the same pieces of music in different acoustic settings in Stavanger concert hall (see 3.1), the perceived differences were astonishing. However, the measured overall changes in spectrum, level etc., were surprisingly small. Long reverberation of course increases the length of a tone, so that it masks the entrance of the next one, but probably most astonishing was the perceived difference in timbre due to changes of the attack also for a first note in a phrase.

The paper will discuss general theory of how an "ideal, diffuse field reverberation" influences on attack. The early response of a concert hall, is, however, seldom "diffuse". The paper will show measurements and simulations (Odeon) of attack with/without early reflections. Methods for measuring attack in concert halls is not as trivial as measuring Rise Time, Steepness etc. for ordinary electronic filters, and alternative approaches like Integrated (Cumulative) Squared Step Response and Spectral Flux is discussed. The importance of attack has been little discussed in concert hall design, but should be more and more important because listeners nowadays are used to recordings where any wanted amount of direct sound is mixed with a long and (too?) smooth, non-correlated, (digital) reverberation. It is generally assumed that a fast attack is more brilliant, sounds more like if there are more high frequencies present than a "fade in".

First we must remember that a signal of course preserves it attack best if we just hear the direct sound only. This paper will show how reverberation always smoothens/prolongs the attack and discuss how this affects signals of different lengths, different musical instruments which in itself have slow or fast onset of note, and if it is possible to reduce the smoothening of the attack by early reflections.

2 ATTACK IN ROOMS WITH/WITHOUT EXPONENTIAL DECAY

Following Schroeder[1], Jordan[2] discusses rise time etc. of concert halls. Assuming that the decay process in a hall follows the exponential function in eq.1, a corresponding (complementary) build-up process may be written as in eq. 2: (assuming a speed of sound of 344 m/s. I_0 is the intensity at zero time).

$$I_{t,decay} = I_0 e^{-\mathrm{kt}} = I_0 e^{\frac{-13.76t}{\mathrm{RT}}} \qquad (1) \qquad\qquad I_{t,buildup} = I_0\left(1 - e^{\frac{-13.76t}{\mathrm{RT}}}\right) \qquad (2)$$

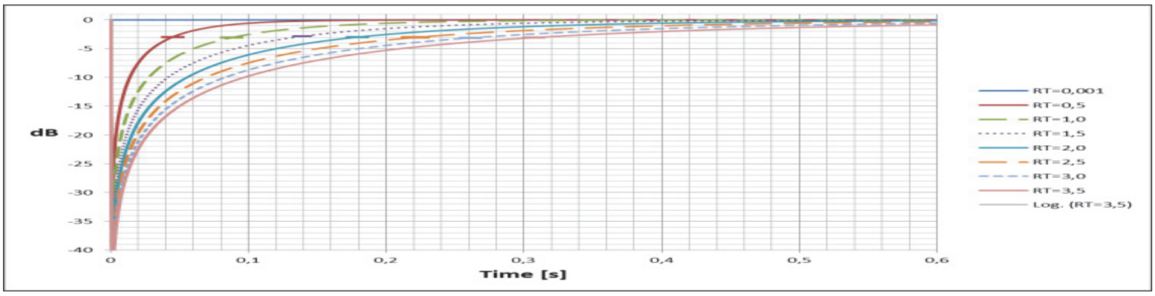

Figure 1 – Attack as a function of reverberation time following Eq.2

The value of **Rise Time** (*TR*) will correspond to the point of time where 50% of the total energy has arrived:

$$\frac{I_{TR}}{I_0} = 0.5 \qquad \text{which is shown to give:} \quad TR \cong 0.05\,RT\ [s] \tag{3}$$

From Jordan[2] we find that:

$$10log\frac{I_{t,buildup}}{I_0} \sim 10log(1 - e^{\frac{-13.76t}{\mathrm{RT}}}) \tag{4}$$

It is further shown that at a level -5 dB below the stationary level, we get a calculated value of Steepness, σ :

$$\sigma_{calc} = \frac{d}{dt}\left(10log\frac{I_{t(-5dB)}}{I_0}\right) = 0.0094\frac{13.76}{RT} \approx \frac{0.13}{\mathrm{RT}[s]} \ \ [\mathrm{dB\ /ms}] \tag{5}$$

Schroeder[1] states that for enclosures with nearly exponential reverberation, the time t_0 at which the sound intensity during the build-up process has reached a level 5 dB below steady state, is typically 1/40 of the reverberation time RT_{60}. Schroeder further states that a more convenient and accurate method of measuring steepness is *'measuring the echo amplitudes of the enclosure near RT_{60}/40 after excitation'.*

For standardised measurements of reverberation time, one often "forgets" the very first part of the decay, and starts the calculations after -5 dB decay. This is of course beneficial for the reproducibility of the measurement results, but we lose a lot of interesting information about possible coloration due to very early reflections etc., see Halmrast[3]. Regarding the shape of the build-up, Jordan[2] states: *'When one considers that the sound paths which are effective in determining the steepness are those which occur early in the build-up process and include the reflections with short time delays, it would seem possible to influence the value of steepness. If, for instance, reflecting surfaces were placed in some sound paths between source and receiver. ..... then this would correspond to a reduction of the effective mean free path in an early interval of the build-up process'.*

Several investigations conclude that measuring the Early Decay Time gives a better judgement of the perceived reverberation in a hall (especially for "running music"). Measurements including the first part of the decay, however, often give somewhat different results for different sender positions on stage. Such differences should in fact be considered important. (We should also remember that measurements on an empty stage often gives results that are repeatable, but not practically interesting).

3 LISTENING AND ANALYSING

3.1 Same music performed in different acoustic settings in Stavanger Concert Hall

For the IMS conference on musicology in 2016, the author had the possibility of playing with a group of musicians and recording the same short pieces of music in different acoustic settings in Stavanger Concert Hall (See Halmrast[4]). The same music was played on the same instruments for three settings in the Valen concert hall (a highly flexible hall, with both flexible absorbers on walls and flexible ceiling), and one "jazz-club"-setting in the flexible Zetlitz hall. The settings in Valen was named *2.Chamber* ($T60_{\mathrm{mid.frq}}$=1.9s), *4.Amplified* (a setting used for amplified events) ($T60_{\mathrm{mid.frq}}$=1.7s) and *6.Concert* ($T60_{\mathrm{mid.frq}}$=2.5s). The musicians were instructed to keep same strength for all settings (as good as possible). The recordings were done at the same position on the 8th row, slightly off-centre, with the same recording level for all settings. The main results are given in Halmrast[4]. The music was: classical trumpet a cappella followed by a string quartet (and trumpet) (material from Mahler 5th), jazz piano trio and rock guitar trio, both with trumpet soli. All music was performed without the use of any house amplification (only guitar and bass personal amplifiers with exact the same levels etc. for all the acoustic settings of the hall). Analysing the recordings, we found that the differences, both in level and overall frequency content were surprisingly small. This could of course be discussed regarding "performology", (how the musicians compensate for the acoustics, even when told not to do so), but for this paper, the most interesting result was that the "driest" settings sounded much more "brilliant"/high-frequency than the more reverberant settings. The note lengths were analysed, and the length of each note of course increases with increased reverberation time. For "the same" snare drum stroke, the length was 0,27 s in the dry *5.Zetlitz* setting, increasing to 0,40 (*2.Aplified*), 0,41 (*4.Chamber*) up to 0,44 s for the *6.Concert* setting. The increased length of each note due to reverberation of course gives that one note "masks" the attack of the next one (if the first one is not very, very short or the time between the notes are very long). This "masking" effect of reverberation is well known. The effect reverberation has on the attack also for the "the first note" is not so well known, and the main issue for this paper. (Close inspections on the spectrograms might indicate that in the most reverberate settings, the high frequencies "arrive later", giving additional "smoothing" to the attacks). (see Halmrast[4])

3.2 Convolution of "dry" recordings with impulse responses

We must remember that the general equations for the build-up of the attack in part 2 are only valid for long signals. Noise bursts; very short, longer and much longer, were convolved with Stavanger Concert Hall in most reverberant *6.Concert* setting. Figure 2 shows sound pressure level over time for increasing length of the noise burst. *Black:* dry noise bursts*, Lime:* Convolved with IR measured on stage (close reflections). *Blue:* Convolved with IR from stage to hall in Stavanger (most reverberant setting, *6.Concert*).

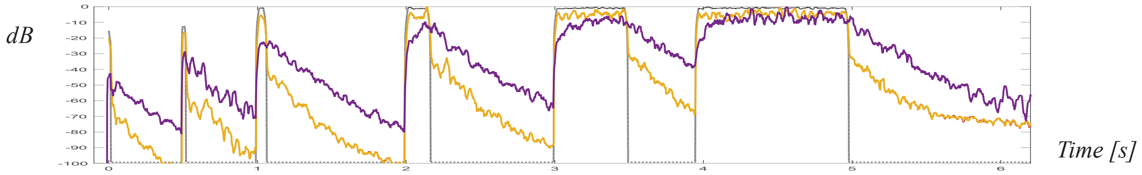

Figure 2 - Noise bursts of increasing length. *Black*=dry, *Lime*=short RT, *Blue*=Long RT

From the left in fig. 2 we see that very short notes are clearly detected without any prolonged attack, because they do not "build-up" in the reverberant room. For the longest reverberation (to the right), the attack for longer notes/bursts are severely prolonged. The decays, however, are similar for all the situations. This simple analysis shows that the room´s influence on the attack not only depends on the reverberation time, but also on the length of the signal (and of course also on the instruments own "onset time"/"attack time"/ "fade in").

The guitar is a good instrument for analysing attack. The guitar did not play any a cappella parts in our tests in Stavanger, but a dry guitar lick (without distortion pedal), was convolved with a very moderate reverberation (like in *5.Zetlitz*), and with the long impulse response measured in *6.Concert.* From figure 3 we clearly see how the reverberation smoothens and prolongs the attack.

Guitar *8 va basso*,

Tempo MM=100

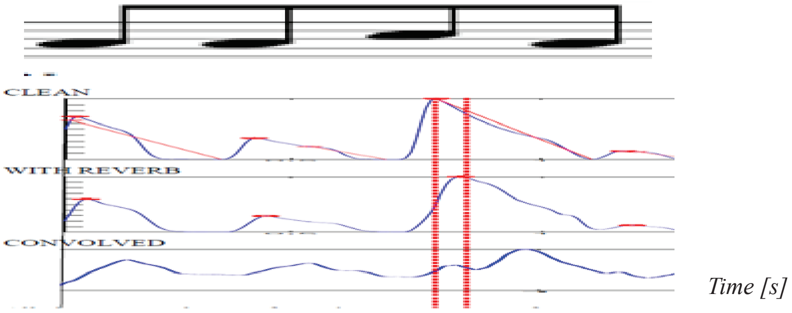

Figure 3 – Guitar Lick. *Upper*=clean/dry, *Middle*=Convolved short RT, *Lower*=Convolved Long RT

4 MEASURING ATTACK

At the moment, a lot of investigations on attack is done, for instance at the Ritmo[5] Centre at Dept. of Musicology at Univ. Oslo. However, most of this work is done regarding rhythm, finding the so-called *p-centre* of the musical event. The effect of attack on timbre is little discussed in literature, but a sharp (short) attack is usually said to be perceived as more "trebly". Hajda[6] mentions attack as: *´that period of the signal in which the global RMS amplitude is rising and the spectral centroid is falling after the initial maximum´……..'In general, three acoustical parameters repeatedly appear as correlates to dimensional solutions in timbre studies: 1. Amplitude-vs-time (temporal) envelope, usually expressed in terms of attack or rise times. 2. Spectral energy* *distribution across frequency components. 3. Spectral variance in terms of the amplitudes of frequency components.* Comment regarding 1*: Log-rise-time = log10(*$t_{max} - t_{thresh}$*), where* t_{max} *is the time from onset to maximum RMS amplitude and* t_{thresh} *is the time from onset to a threshold taken as 2% of the amplitude at* t_{max}*'.* Here we see yet another, similar, but not identical definition of Rise Time. (See also Vos[7]).

A big problem when measuring timbre of attack is of course that the attack is fast, but a good frequency analysis requires measurement over a long time ("window"). The next problem is uncertainty about our hearing: How long time do humans actually integrate over, when the sound is changing? Using very small time windows for the analysis will separate the direct sound and the very early/early reflections, and indicate longer attack duration with early reflections than without. Often one state that up to 80ms adds clarity for music, but that cannot be the case if the signal is very short. Often 20 ms is suggested for the integration time. We have analysed both for 10 ms and 20 ms, as very early reflection often arrives within such time limits. Some researchers use 7 ms as the limit for "stage reflections" that is included in the direct sound, (even if they are not really part of the direct sound).

Meyer[8] states*:".....the perceived point of tone entrance lies about 10 dB below the final sound level, or the masking threshold (in the presence of pre-existing noise), and this is relatively independent of the speed of the attack. For very soft tones the point of attack can move as close as 7 dB to the final sound pressure level, i.e., it is sensed even later. For very loud tones, the tone entrance is already perceived at a sound pressure level of 15 dB below the final value".*

For the actual design of a hall, one should of course also pay attention to the direction of the reflection, and the fact that a single, distinct reflection gives audible coloration (comb filter with a distance between the dips (or CBTB, Comb Between Teeth Bandwidth) that is in the order of the critical band, which means delays in the region 5-25 ms, see Halmrast[3].

4.1 Measuring attack from recorded music. MIR and Spectral Flux

Measuring using the parameter Attack Time in common Music Information Retrieval/MIR- programs (like MIR Toolbox[9] and the new Mining Suite[10], and also parameters like Intensity (in Praat etc.) often includes some kind of low pass filtering ("long window"), giving that the attack is somewhat smoothed also in the actual analysis, so some of these methods are not quite useful for our task. (A new version of MIR Toolbox will incorporate improved measurements of attack time).

Spectral flux is a measure of how quickly the power spectrum of a signal is changing, comparing the power spectrum for one frame against the power spectrum from the previous frame. It is usually calculated as the *2-norm* (also known as the Euclidean distance) between the spectra. The spectral flux can be used to determine the timbre of an audio signal, or onset detection. There are numerous variants of Spectral Flux, and they all give different results and present the results in different ways. One version of Spectral Flux is incorporated in MIR Toolbox, and other versions can be found as plugins for Sonic Visualiser etc. Figure 4 shows the overall measurements of spectral flux from the recordings of the whole selection of music, for each of the acoustical settings. We see that the shortest reverberation time (*5.Zetlitz,* jazz-club setting), gives the highest Spectral Flux.

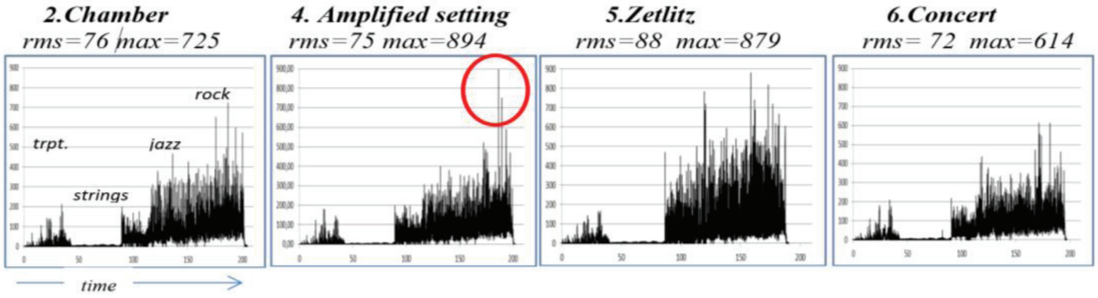

Figure 4 –Spectral Flux for same music in different acoustic settings in Stavanger Concert Hall
(The red circle indicates an accidental strong drum stroke)

4.2 Measuring attack from Impulse Responses

In electronics, **Rise Time** is the time taken by a signal (step function) to change from a specified low value to a specified high value. These values may be expressed as ratios or as percentages or dB values with respect to a given reference value. In analogue and digital electronics, the percentages are commonly the 10% and 90% of the output step height, however, other values are commonly used. (Examples of "other values" are -5 /-3 dB mentioned in part 2). We shall see that these parameters often are too "general", so we need to examine the attack more in detail. The method we found most convenient was to integrate the squared impulse response, which also might be called a Cumulative Step Response.

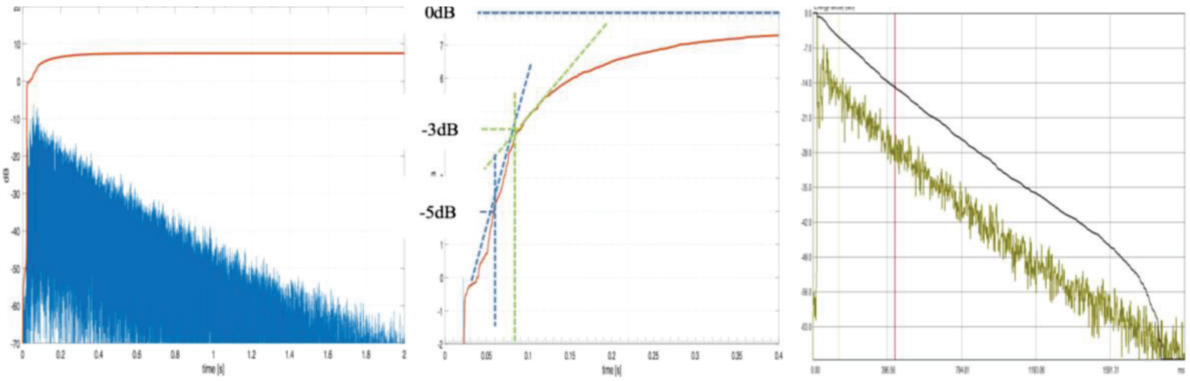

Figure 5 – *Left:* Imp.Resp. and Int.Imp.Resp. *Middle:* Variation of Steepness during attack phase. *Right:* Imp.Resp. and Schroeder curve.

Figure 5 shows measurements in a shoebox hall with reasonable exponential decay, (the University Aula, Oslo, empty). To left we see the squared impulse response (up to 2s, in blue) and the integral of this (red) in the left pane. The pane in the middle shows a zoom in of this Integrated Impulse Response (up to 0.4s) for. To the right is shown the common energy decay and the Schroeder curve.

The reverberation time (T30 and T20) for this hall is app. 2.2 s (empty, mid freq.). According to Jordan (in Part 2), Steepness should then be 0.13/RT= 0.13/2.2=0.059 dB/ms. From the middle pane in the figure above, we see that the steepness is changing during the attack, but we find a typical value around the -5 dB point of 0.087 dB/ms, and 0.033 dB/ms around the -3 dB point. The mean value of these might actually correspond quite well. From the same figure, we find that Rise Time (up to -5 dB) is 38 ms. For the build-up unto -3 dB, we get a Rise Time of app. 60 ms. The calculated rise time, TR according to Jordan (up to -5 dB), should be 0.05 x 2.2 = 0.110 s= 110 ms. As a conclusion, both Steepness and Rise Time give measured values that deviate from the equations for exponential decay (and build-up). The hall is a moderately small regular shoebox, so this hall, if any, should be assumed to have exponential decay. One reason why even this hall differs from the equations might be that it actually has some (nice) early reflections. Keeping in mind that the measurements highly depends on where on the build-up curve we analyse these parameters, we should not relay too much on either Rise Time or Steepness before we have more measurements from different halls. For now, we should inspect the shape/curvature of the Integrated, Cumulative Squared Impulse Responses itself in order to find the interesting issues of attack.

5 ROOM ACOUSTIC MODEL

A very simple, large hall was modelled in Odeon, with sidewalls and rear wall behind audience as reflective, highly diffusing/scattering. (See Halmrast[14] for details). Dimensions: L x W x H = 46 x 31 x 20 m. Receiver: 30 m distance, 1m off centre, height 2 m. Three different settings of a "box" around the source, open only towards the hall, were analysed, called; *Transparent*, *Reflecting* and *Absorbing*. Especially the two last ones will be discussed here. Figure 6 shows both the impulse responses and the Integrated Imp.Resp. up to 320 ms. (PS! For the first part of the Imp. Resp., both curves are identical, so the red curve is covered by the black one).

Comparing the common acoustic parameters for *Refl.Box* and *AbsorbBox,* the close reflections give an increase in C50 and C80. Spectral centroid for the Impulse Responses is also increased from 7165 Hz to 7504 Hz due to the close reflections. Adding the close reflections changes EDT more than RT. (For this simple simulation, it seemed like the close reflections gave an increase in EDT for 500 Hz and below, but a decrease for 1 kHz and above. This should be investigated further, together with close examination of the EDT algorithm in use, as the first reflection(s) actually might be stronger than the direct sound, and make confusions in choosing the exact time of arrival of the direct sound). From analyses of Spectral Flux, we clearly see that the (changes in) Spectral Flux is larger for the black curves/Refl.Box.

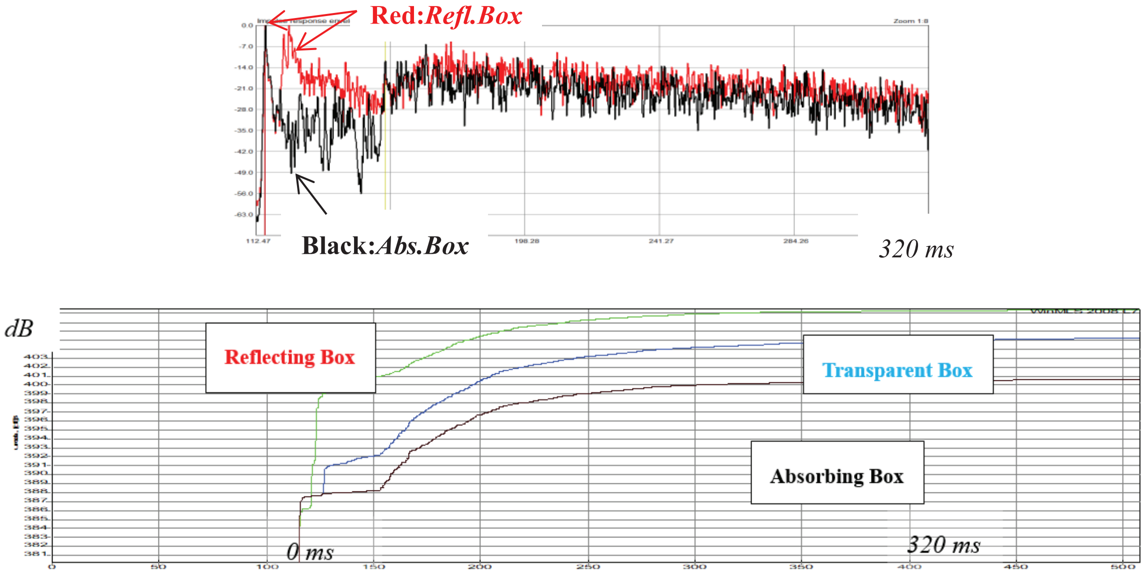

Figure 6 – Upper pane: Impulse response with/without early reflections (Odeon)
Lowe pane: Integrated Impulse Response. *Red*= with early reflections, *Black*=without early reflections:
(*Blue*= with transparent box, not discussed)

5.1 Note Length and Musical Instruments own attack time/build-up

Figure 7 shows a short musical phrase that was played by different instruments, “dry” and convolved with the measured impulse response from the most reverberant setting in Stavanger Concert Hall (*6.Concert*). Figure 7 shows a comparison of the attacks for an instrument with very short sounds; a xylophone, compared with a “slow reacting” and long sustaining instrument; a bowed violin. We see that for the convolved curves, the attack is preserved as rather short for the xylophone, but prolonged for the violin. We also see that for the violin, the decay (and sustain) of one note masks the entrance of the next. Listening to these examples, each note is heard separately for the xylophone, and the first attack is rather well preserved, but the violin has a got an added build-up for the first note, and the second and especially the third note is almost not perceived as a new attack, but is masked by the first notes. (Total length is 1.5 s. Division is 5 dB).

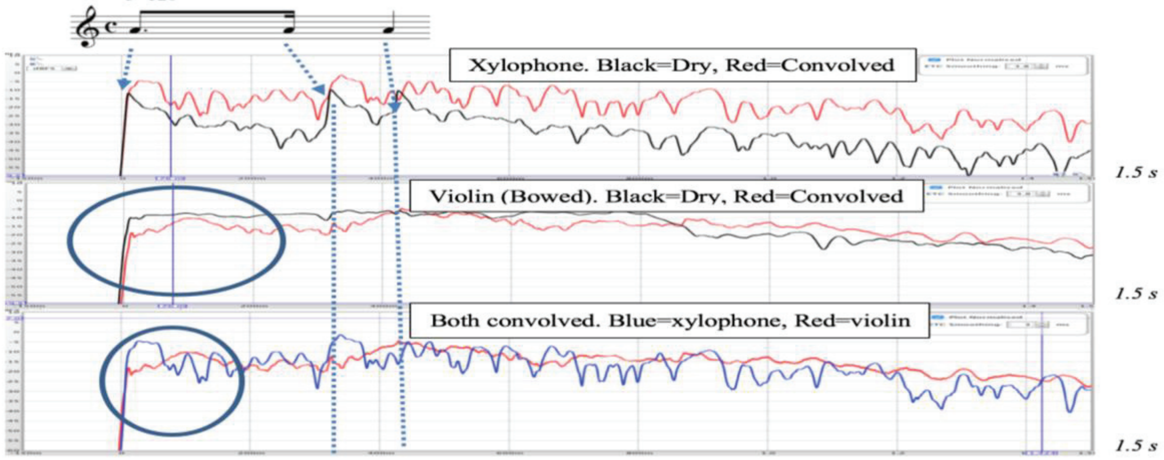

Figure 7 – Analysis of xylophone and violin playing the same music. Dry and convolved with long RT

Figure 8 shows Energy Time Curve (ETC) of the first part of the phrase, (up to 60 ms) for xylophone. *Black*=Dry recording, *Green*=Convolved/Odeon/Refl.Box and *Blue*=Convolved/Odeon/Abs.Box. The leftmost pane is for very short ETC-smoothing (0.2 ms), the middle pane is for 5 ms, and the pane on the right is for 10 ms ETC-smoothing. (Division is 2 dB).

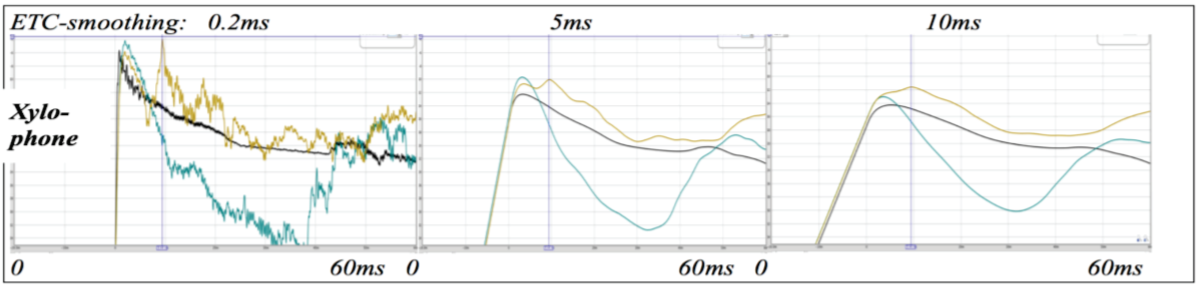

Figure 8 – ETC (Energy Time Curve) of the same xylophone for different smoothing: 0.2 ms, 5 ms and 10 ms, *Black*=Dry recording, *Green*=Convolved/Odeon/Refl.Box and *Blue*=Convolved/Odeon/Abs.Box

Figure 8 clearly shows the importance of the decisions we have to make when analysing small differences in short attacks. If we assume that our ear has an integration time of 10 ms (or more), the setting with close reflections, *Refl.Box*, will be perceived with the "clearest" attack.

For the bowed violin in figure 9 it is problematic to see if close reflections give any influence, because the build-up of a violin tone itself is so slow. The following figure shows the same 60ms, but now for the "slow", bowed violin convolved with *Refl.Box* (Black) and with *Abs.Box* (red), and we see that the influence on the attack is much smaller than for the xylophone, and that the issue of integration time is not that important for the "slow" violin.

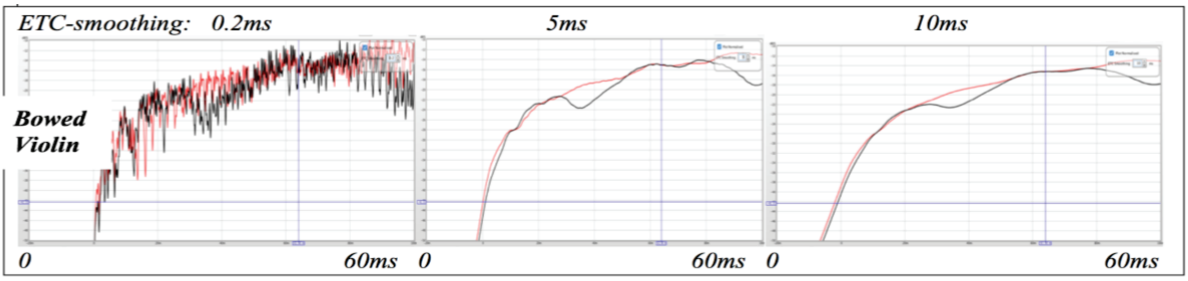

Figure 9 – Energy Time Curve of the same violin (arco) for different smoothing: 0.2 ms, 5 ms and 10 ms *Refl.Box* (Black) and with *Abs.Box* (red

6 MEASUREMENTS IN A CHURCH

Figure 10 shows the attack (Integrated Impulse Response) for sender position *Without*(aisle) and *With* close reflections (pulpit) in a medium sized church. (see Halmrast[14] for details).

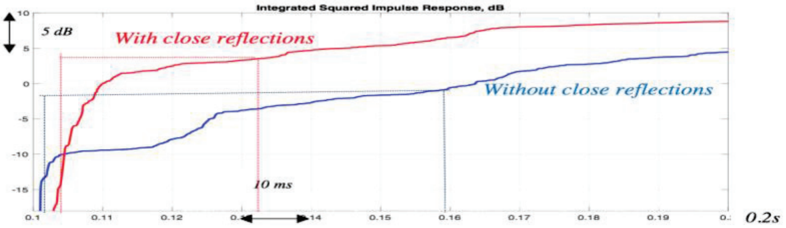

Figure 10 – Integrated Impulse Responses with/without close reflections in a church.

We see that the overall attack is faster with the close reflections. If we should want to measure the Steepness, we would need to closely define which part of the curve is interesting (especially for the blue curve). Quick estimations of Rise Time (to – 5dB) gives: $TR_{WithCloseReflections} = 30ms$, and $TR_{WithoutCloseReflections} = 58\ ms$. Schroeder´s "rule of thumb" for exponential decay in Part 1 gives: $2.4s/40 = 60\ ms$ so, for this quick test, the agreement is nice for the measurement without close reflections, and the close reflection practically halves this attack time. However, until we have more analyses and simulations, we should inspect the shape of the attack curve in detail and compare with the typical note length, and not rely too much on parameters like Steepness and Rise Time.

7 CONCLUSION

Reverberation "smoothens/prolongs" the attack. For rooms with an exponential decay, the build-up of the attack can be determined from the decay. However, most interesting halls do not have exponential decay, especially in the early part, and we need to investigate the attack more in detail.

Integrated, Squared Impulse Response (Cumulative Step Response) is a good way of investigating attack. Parameters like Rise Time and Steepness can be derived from this, but for many interesting situations with early reflections, the curvature of the attack is non-uniform and we should inspect the actual shape of the curve rather than trust these parameters. It is shown from computer simulations and measurements in a church that **early reflections** in might help preserving time for halls with long reverberation times.

The effect reverberation (and early reflections) have on attack is highly dependent on the type of signal:

- **Extremely short, bright signals** can be so short that they do not "build up" due to the reverberation.
- **Instruments with long internal build-up/onset** (like bowed strings) get a nice development from the smoothed attack due to long reverberation time in the hall.
- **For most instruments in between these groups,** (especially piano with its special "resonance" phase) there is a problem of "smoothing" due to prolonged attack from the reverberation.

This can be reduced by introducing early reflections. For the design of halls, however, we must also remember that distinct early reflection might give comb filter coloration, and take the direction and delay times of the early reflections into consideration (See Halmrast[3,14]). Attack is highly important for today´s listeners, who are used to listening to recordings where any wanted amount of direct sound is mixed with a (too?) long, diffuse reverberation.

REFERENCES

1. Schroeder, M. R. *'Complementarity of Sound Buildup and Decay'*. JASA 40, 549 (1966)
2. Jordan, V. L. *'Acoustical Design of Concert Halls and Theatres'*. Applied Science Publishers, London (1980) p.11 and 206-211. (and J. Audio Eng. Soc. 13, 98-103 (1965))
3. Halmrast, T *'Orchestral Timbre: Comb-filter coloration from reflections"*. Journ. Sound and Vibration, 232(1),53-69, (2000)
4. Halmrast, T '*Analysis of Music Performed in Different Acoustic Settings in Stavanger Concert Hall'*. Int. Congress on Sound and Vibration, ICSV24, London (2017), paper 398
5. Ritmo, Centre for Interdisciplinary Studies in Rhythm, Time and Motion. https://www.hf.uio.no/ritmo/
6. Hajda, J. M. *'The Effect of Dynamic Acoustical Features on Musical Timbre'* in *'Analysis, Synthesis, And Perception Of Musical Sounds:The Sound Of Music'* ed. J.Beauchamp. Ch.7, p.254
7. Vos, J. and Rasch, R. (1981): *'The perceptual onset of musical notes',* Percept. Psychophys. 29, S. 323
8. Meyer, J. *'Acoustics and the Performance of Music'* Springer, p.12, (2009)
9. MIR Toolbox. https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox
10. Mining Suite. https://www.hf.uio.no/imv/english/people/aca/temporary/oliviel/
12. Halmrast, T. *'Acoustics in Between: Perception of Sound in Rooms Beyond Standard Criteria'.* Psychomusicology: Music, Mind, and Brain, Am. Psychological Association 2015 Vol25, No3, 256–271
13. Halmrast, T *'Attack the Attack'* IoA proceedings, Hamburg 2017, Vol.40, p 335-345
14. Halmrast, T. (2019) *'Sam Philips' Slap Back Echo, Luckily in Mono".* Journ. Art of Record Production http://urn.kb.se/resolve?urn=urn:nbn:se:kmh:diva-3124