

# Impulse response of the Bayreuth Festspielhaus

Kai Huang<sup>1</sup>

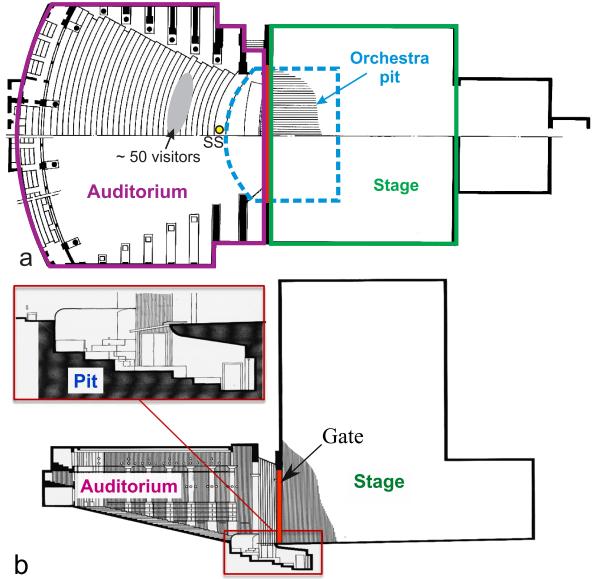
<sup>1</sup> Experimentalphysik V, Universität Bayreuth, 95440 Bayreuth, Germany, Email: kai.huang@uni-bayreuth.de

## Introduction

The Bayreuth Festspielhaus is unique for its special acoustic design by Richard Wagner in the 19th century [1]. It has also been well preserved since then for Wagnerian operas. The hidden orchestra pit, for instance, clears the way for the audience to focus on the stage. Together with the reflecting board that blocks the direct sound, the pit acts as a low-pass filter and sets the sound from the orchestra in perspective with the singers [2]. Due to the special design, the reverberation time (RT), which corresponds to the time scale for the sound pressure level to decay 60 dB, of the Bayreuth Festspielhaus is relatively large compared with the other opera houses of the same age [1].

Impulse response is often used to characterize room acoustics, because it probes all frequency components of a room. Based on the measured response, further characterizations on the strength, speech intelligibility, echoes, reverberation, and other features are straightforward to be quantified [3, 4]. Such a quantification facilitates room acoustics design and planning, so that incommunicable rooms such as the one that Sabine was facing more than 100 years ago [5] can be avoided. The source of an impulse can be a loudspeaker fed with a pulse signal, exploding objects such as air-balloons, or hand-claps. Following a standard protocol [6], we can determine RT and other room acoustic quantities from the recorded signals [3, 4]. The RT and other room acoustic properties of the Bayreuth Festspielhaus have been measured systematically for more than half a century [1]. Shortly before the recent renovation, Garai and colleagues have conducted a systematic characterization of the Festspielhaus in 2014 [7].

With the recent development of information technology, room acoustic characterizations are becoming more convenient using hand claps and the microphones from smart phones [8, 9]. An interesting follow-up question is: How well can the acoustics of a room be characterized using this approach in comparison to the conventional measuring protocol? Here, the impulse response of the Bayreuth Festspielhaus acquired by hand-claps and smart phone recordings (HCSP) is presented. The corresponding room acoustic parameters (RT and center time) are obtained with the ITA-Toolbox [10] and compared with literature values, which were collected in the same month as the present investigation (September, 2014) [7]. Based on the comparison, I discuss the advantages and typical drawbacks of the HCSP and provide a checklist for collecting sensible data from this approach.



**Figure 1:** Architects' drawings of the Bayreuth Festspielhaus adapted from Ref. [1]. (a) A floor plan with the borders of the auditorium, the stage and the orchestra pit, the location of the sound source (SS) marked. The recording devices are located in the region shaded in gray. The red line represents the stage gate, which was closed during the measurement. (b) Corresponding side-view sketch with a close-view of the orchestra pit.

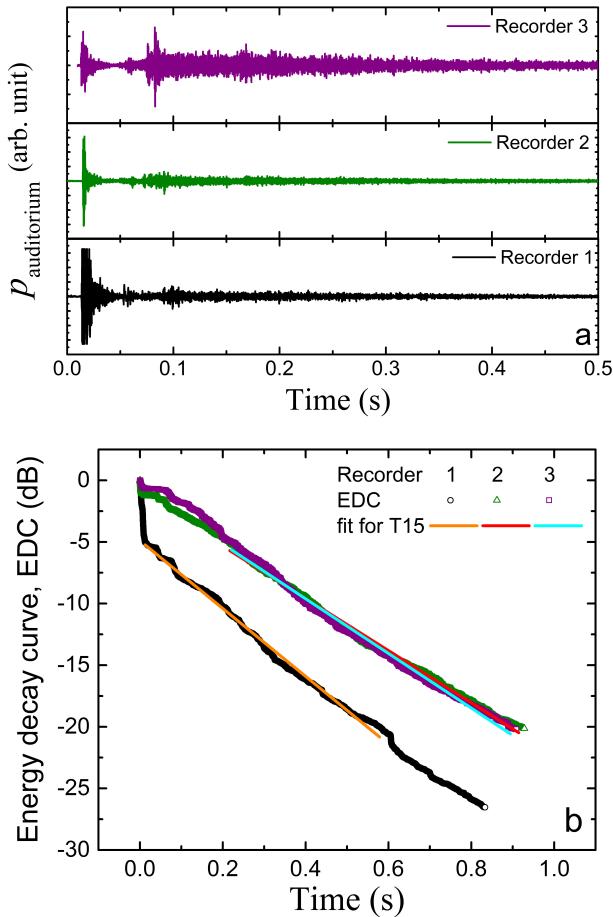
## Measuring Procedure

Figure 1 shows the sketches of the Festspielhaus adapted from the architects' drawings. The measurement was conducted during a visit with about 50 visitors on September 11, 2014. As marked in Fig. 1(a), the sound of five hand-claps was generated at location SS, directly in front of the orchestra pit. The sound was recorded by the microphones of three smart phones with a sampling rate of 44.1 kHz. The distances between the source and recorders were 5 ~ 10 meters, roughly double the reverberation distance [4]. In addition, five hand-claps with both SS and microphones in the orchestra pit (with shorter distances between SS and the recorders due to the limited space) and on the stage were also performed. Subsequently, the recorded signal of each individual clap was extracted manually for post processing.

The data analysis was performed with Matlab using the ITA-Toolbox [10], which provides a standard routine for room acoustic characterizations following ISO 3382 [6]. More specifically, the raw signals are filtered into different octave bands and subsequently the energy decay curve (EDC) for each band is obtained through applying the backward integration method on the squared envelope.

lope of each filtered signal [3]. From least square fits of the EDC in the semi-logarithmic plane, the reverberation time is obtained. Depending on the different ranges of data used in the fitting, EDT (early reverberation time), T15, T20, T30 (i.e., using the energy decay from  $-5$  dB to  $-20$ ,  $-25$ , or  $-35$  dB), etc., are obtained. In the present study, T15 and the center time obtained in the frequency range from 250 Hz to 4000 Hz, covering five octave bands, are presented. T20 and T30 are not used because they cannot be obtained for all frequencies due to the limited signal-to-noise ratio (SNR).

## Results and Discussion

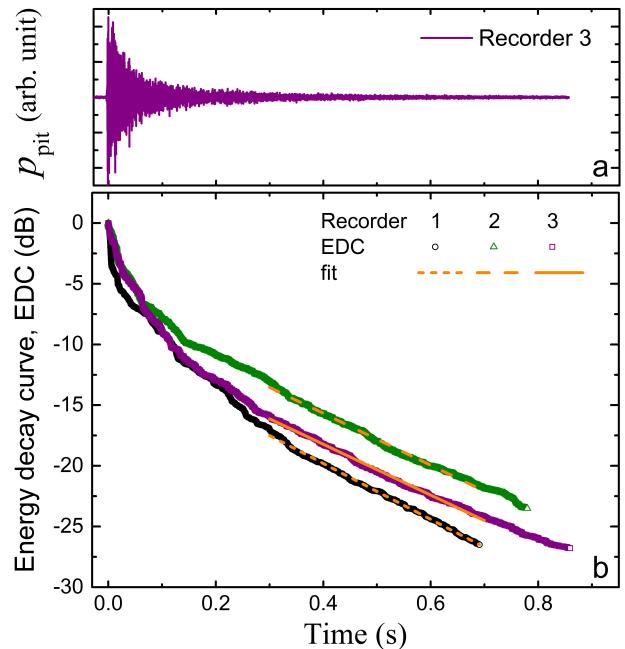


**Figure 2:** (a) Representative raw data captured by the three recorders in the Auditorium. (b) Corresponding energy decay curves for center frequency 1000 Hz. The solid lines in (b) are linear fits of the data in the semi-logarithmic plane to obtain T15.

Figure 2(a) shows the raw signals captured by the three microphones. Although all the smart phones use the same program (Smart Voice Recorder) for recording, the sound pressure differs one from another, owing to the different properties of the microphones and analogue-digital (AD) converters embedded, as well as the different distances to the SS. The following features can be learned from a comparison of the raw signals: (i) The direct and reflected sound can be clearly distinguished (separated at  $\sim 0.05$  s) for all recorded signals. (ii) the sound pressure of the direct sound may vary dramatically from one recording device to another. For instance, the maximum

sound pressure from recorder 1 is more than one order of magnitude larger than that from recorder 3. (iii) The direct sound is not always the strongest peak in the impulse response. The difference of the direct sound energy will lead to the scattering of the center time, which will be discussed at the end of this section.

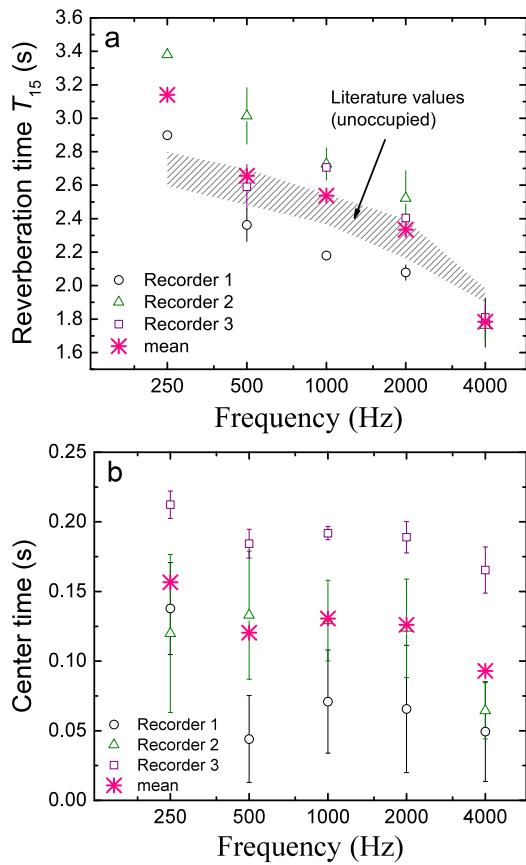
The cumulative energy decay curves obtained with the backward integration method [11] are shown in Fig. 2. For all three recordings, EDCs have a similar behavior: A stepwise initial drop followed by an exponential decay (note the semi-logarithmic scale). The initial drop of the energy level represents the influence of the direct sound. Because of the strong direct sound from recorder 1, the magnitude of the corresponding EDC drop, which represents the energy level of the direct sound, is much larger than the other two. Based on definition, RT is obtained through a fit of the exponential decay part of EDC. The agreement of the EDCs in this regime indicates that the exact locations and gain levels of the microphones play a minor role in obtaining RT, because those configurations only lead to different factors to the EDC, not the decay exponent. Because the sound energy generated by hand-claps is limited and the waiting time between subsequent claps is not sufficiently long, the conventional T30 cannot be obtained for all octave bands. Instead, T15 is used here. This difference can lead to additional uncertainty in determining RT. Moreover, as indicated by the EDC of recorder 1, too large direct sound energy may lead to a rapid drop of the energy level to below  $-5$  dB and consequently a shorter reverberation time.



**Figure 3:** (a) A sample recording obtained in the orchestra pit, illustrating the lack of clear separation between direct and reflected sound. (b) Corresponding EDCs at 1 kHz for all three recorders. Solid, dashed and short dashed lines correspond to the linear fits of the data in the semi-logarithmic plane, using the energy decay between 0.30 and 0.70 s.

As shown in Fig. 3(a), no clear distinguish between the direct and reflected sound can be made for the data col-

lected in the orchestra pit. This can be attributed to the relatively small volume and the complex geometry of the pit [see the inset of Fig. 1(b)]. In addition, the sound reflecting board, which was designed to block the direct sound to the stage, effectively mixes the directed and reflected sound together. Last but not least, the opening to the stage gives rise to additional influence from the stage and auditorium. Consequently, the EDCs shown in Fig. 3(b) exhibit a gradual decay instead of the rapid drop. Thus, using the conventional ways of determining RT is inappropriate, because the exponential decay starts at an energy level lower than  $-5$  dB. Thus, I fit the tails of the EDCs and obtain RT from the slopes of the fits. An average of the three EDCs yields an RT of  $1.38 \pm 0.08$  s for the octave band centered at 1000 Hz.



**Figure 4:** (a) Reverberation time of the auditorium for different octave bands. The error bars correspond to the standard deviation from five hand-claps. The asterisks are mean values from the different recorders. The region shaded with gray lines corresponds to the literature values [7, 12]. (b) Corresponding center time at different frequency.

The reverberation time obtained from different hand-claps and recorders is shown in Fig. 4(a). A comparison between the mean value obtained from the current measurement and the literature values [7, 12] demonstrates that RT can be extracted from the simple HCSP approach, at least in a certain frequency range. However, due to the lack of high repeatability, different hand-claps yield slightly different RTs (represented as errors). Thus, an average of the obtained RTs from different clapping events and recorders is necessary for a sensible characteriza-

terization. The error from the exponential fits can be ignored as it is typically much smaller than the uncertainty from different clapping events.

The deviation from the literature values at low frequencies ( $\leq 250$  Hz) can be attributed to the following three reasons: (i) The sound energy does not always decay to below  $-20$  dB within the recording time period because of the slow decay rate for low frequency sound. Therefore, the statistics is not sufficient for an accurate quantification. (ii) The strong influence of the direct sound may lead to large fit error because the exponential decay may start at an energy level smaller than the standard  $-5$  dB. (iii) The frequency response of MEMS (microelectromechanical systems) microphones, which are typically used in smart phones, has a roll-off at low frequencies and a peak at  $\sim 15$  kHz due to the design of the chamber geometry [13]. Consequently, the reliability of data obtained at low frequencies may suffer from the low SNR.

RT obtained at different locations is compared in Table 1. Quantitative agreement between the data from the auditorium and from the stage is found for frequency  $\geq 500$  Hz, suggesting that the characterization of RT is weakly dependent on where the hand-claps and recorders are located. Due to the same reason described above, there is also an overestimation of RT at 250 Hz for the RT obtained on the stage.

**Table 1:** Mean values of the reverberation time obtained in the auditorium (A) and on the stage (S), unoccupied.

	250	500	1000	2000	4000	(Hz)
A	3.1	2.7	2.5	2.3	1.8	(s)
S	3.5	2.6	2.5	2.1	1.8	(s)

Figure 4(b) shows the center time obtained from the same measurement. It is defined as the first moment of the squared impulse response  $t_s = \int_0^\infty [p(t)^2] dt / \int_0^\infty [p(t)] dt$  [3]. As  $t_s$  characterizes the balance of direct and reflected sound, it is expected to vary with the locations of the recorders and sound sources, as well as with the source signals generated. Therefore, the data scattering is strong in comparison to the RT results. As described above, the influence of direct sound on recorder 1 is the largest among the three, therefore the corresponding  $t_s$  obtained is the smallest for most frequencies. An average over the results from different frequencies yields  $0.125 \pm 0.023$  s, which also compares fairly well with a previous measurement [7]. This value suggests that the speech intelligibility is  $\geq 80\%$  [3] if the speaker is standing close to the proscenium.

## Conclusions and Outlook

To summarize, room acoustic characterizations are conducted for the Bayreuth Festspielhaus using hand-claps as sound source and smart phones as recording devices. The reverberation time obtained with the ITA-Toolbox agrees quantitatively with another measurement [7] taken place in the same month within a certain frequency range. Possible reasons for the data scattering among different

recorders and hand-claps, as well as deviations from the literature values are discussed. This investigation demonstrates the possibility of using amateur measurement devices for monitoring room acoustics, provided that the following precautions are properly taken care of:

- The sound generated by hand-claps should be as loud as possible to have sufficient initial sound energy, otherwise the accuracy for individual measurements will suffer from the low SNR.
- Sufficient waiting time (at least  $1.5t_s$  with  $t_s$  the expected RT) between individual claps is necessary for extracting the response at low frequencies, because the low frequency components decay slower than the high frequency ones.
- Multiple measurements (hand-claps, recording devices and locations) are needed for better statistics.

In addition to RT, the center time obtained with the same recordings shows strong scattering because of its dependence on the positions of the sound source and recording devices. Therefore, further analysis on the spatial distribution of  $t_s$  is needed for a better comparison with the other measurements.

The possibility of quantifying room acoustics properties with easily accessible devices helps, for instance, to monitor an opera house in occupation on a more regular basis and to provide instantaneous feedback on building open-air theaters with the help of real-time analysis tools [8, 9].

## Acknowledgments

I acknowledge the organizers of the 19th Dynamics Days Europe for arranging the visit to the Bayreuth Festspielhaus. I would also like to thank Erik Werner, Thomas Müller, and Stephan Messlinger for their kind help in audio recording.

## References

- [1] Beranek, L.: Concert Halls and Opera Houses. Springer, New York, 2010
- [2] Polack J. D. and Retbi M.: Wagner and the Cinema: A cognitive approach to the acoustics of the Bayreuth Festspielhaus. 8th International Conference on Auditorium Acoustics (2011), 198-206
- [3] Kuttruff, H.: Room Acoustics, 4th Edition., CRC Press, London, 2000
- [4] Müller G. and Möser M.: Handbook of Engineering Acoustics. Springer, Berlin and Heidelberg, 2014
- [5] Sabine, W. C.: Collected papers on acoustics. Harvard University Press, 1922
- [6] ISO3382-1, Acoustics – Measurement of room acoustic parameters. Part 1: Performance spaces, Geneva (2009)
- [7] Garai M., D’Orazio D., et al.: The acoustics of the Bayreuth Festspielhaus. 22th International Conference of Sound and Vibration (2015)

- [8] Seetharaman P. and Tarzia S. P.: The Hand Clap as an Impulse Source for Measuring Room Acoustics. Audio Engineering Society Convention 132 (2012)
- [9] Rosenkranz A., Burgmayer, R., et al.: Eine mobile App zur Bestimmung raumakustischer Parameter. 43th German Annual Conference on Acoustics (2017)
- [10] Dietrich P., Guski M. et al.: Measurements and Room Acoustic Analysis with the ITA-Toolbox for MATLAB. 40th Italian (AIA) Annual Conference on Acoustics and the 39th German Annual Conference on Acoustics (2013)
- [11] Schroeder, M. R.: New Method of Measuring Reverberation Time. J. Acoust. Soc. Am. 37 (1965), 409–412
- [12] Müller, K.: Die Demokratisierung der Akustik (The democratization of acoustics), in Kiesel M., Das Richard Wagner Festspielhaus Bayreuth. Nettpress, Düsseldorf, (2007) 174-197
- [13] Michaelis T. A. and Kling C.: MEMS-Mikrofone in der Metrologie. 43th German Annual Conference on Acoustics (2017)