AURALIZATION OF ORCHESTRA BASED ON 1/10 SCALE MODEL EXPERIMENT

Takayuki Hidaka, Kazunori Suzuki and Shinichiro Koyanagi   Takenaka R & D Institute, Japan

1   INTRODUCTION

The 1/10 scale acoustic model has the feature of more accurately capturing physical phenomena caused by such as 3D curved walls, surface irregularities, seating areas, and balconies in the acoustic design of a hall. These are pivotal to achieving the fidelity of auralization that can withstand subjective evaluation, while they are still challenging to formulate in CAD simulation within the satisfactory audible quality. In addition, a scale model is more suitable for auralization because, in principle, it can cover the entire audible frequency range: lower than 100 Hz and higher than 5 kHz. On the other hand, scale model experiments remain difficult in measuring apparatus: sound source and microphone, especially at high frequencies due to the sensitivity, directivity, and temporal resolution.

In order to incorporate the merits that CAD simulation does not possess in acoustic design, we have been working on improving the measurement accuracy in scale model experiments for about the past ten years [1-3]. Here, we have confirmed that laser-induced breakdown acoustic waves will solve most of the problems related to sound sources, so we have tried the auralization of an orchestral performance. The outline of the measurement system is presented.

2   FLOW OF AURALIZATION

The configuration of the measurement system is shown in Figure 1. YAG laser-induced breakdown acoustic wave is generated on stage and is captured at the desired audience seats with 1/8" microphones (B&K Type 4138). They are both set on the movable 2D and 3D stages, which are automatically controlled by a PC. The A/D converter is 16 bits (developed in-house), respectively, with sampling frequencies up to 1,000 kHz, which is synchronized by an external clock for synchronous summation processing. The flow of auralization is shown in Fig. 1, in which some of the steps have been previously reported [1-3]. Only the summary is given in Fig. 2.
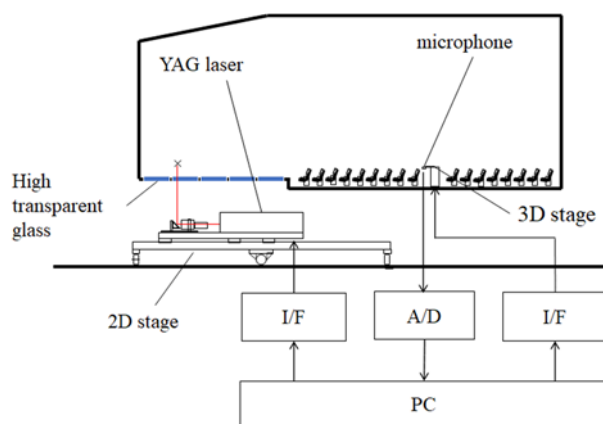

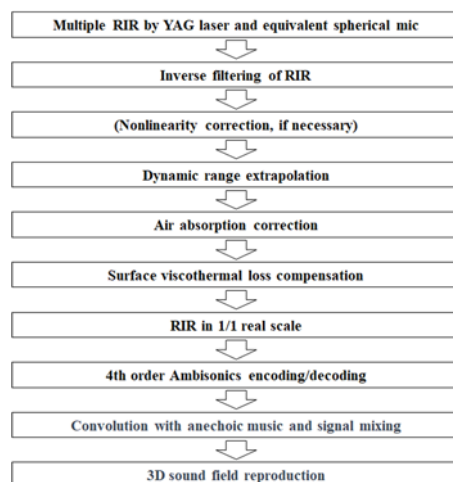Fig. 1 Block diagram of 1/10 scale model measurement.

Fig. 2 Auralization of an orchestra by 1/10 scale model experiment.

2.1 Laser-induced breakdown（LIB）by YAG laser

As a laser-induced breakdown produces an acoustic pulse with superior physical characteristics compared to existing sound sources for 1/10 scale model experiments [4, 5], we can expect to measure more accurate RIRs. A pulse YAG laser (Q-smart 450, Quantel, France) is used in this report. This can vary the pulse energy from 450 to 10 mJ, and it is 450 mJ in this experiment.

2.1.1 Frequency characteristics and waveform

The LIB acoustic pulse has a relatively flat amplitude response: the amplitude deviation is about ±12 dB (at 1 m distance) from about 100 to 100k Hz, which meets the frequency bandwidth required for auralization by the 1/10 scale model experiment. This peak-dip response is due to the fact that the original waveform is similar to an N waveform. Still, the frequency dependence is relatively gentle, so inverse filtering makes flat response readily available within the audible frequency range. In the time domain, its duration is about 0.05 ms, meaning the spatial resolution in the model is 10 cm. Since little transient tail is generated, which feature is beneficial for temporal analysis of the RIR.
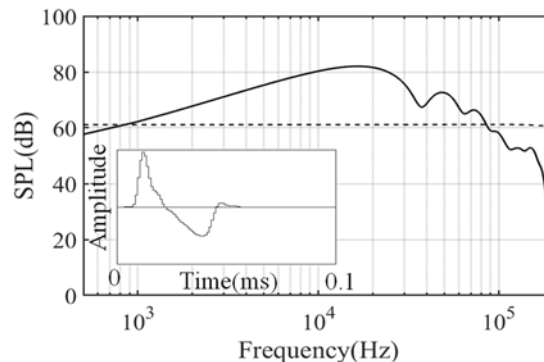

Fig. 3. Amplitude spectrum and waveform measured at 1 m from the laser-induced breakdown. The dashed line is the corrected spectrum by deconvolution.

2.1.2 Nonlinearity: Attenuation by distance

Most research on laser sound sources relating to applications to room acoustics was mainly focused on spherically divergent weak shock waves. The numerical analysis of the Burgers equation by Yuldashev [6] is a representative study. Our interest here is the degree to which this nonlinear wave approaches linear after propagating over a specific distance.

The distance attenuation from a spark source has been reported so far, relating to the positive peak overpressure (for O.A. frequency). The leading edge is a shock wave, and close to the source, the positive peak overpressure is substantially into the nonlinear region. Consequently, the shock propagates faster than the ambient speed of sound. It is known that the positive peak overpressure decreases with distance by more than 6 dB/ dd, but as the positive peak overpressure decreases below a specific value, then it asymptotically approaches a linear behavior [7].

Landau [8] and Peirce [9, p.603-606] published theoretical studies on distance attenuation for weak shock waves.

$$P(r) \sim \frac{1}{r\sqrt{\ln(r / r^*)}} \text{ for } r > r^* \tag{2}$$

Landau also presented an equation of the same type:

$$r^* = 0.62(E / \rho c^2)^{1/3} = 0.012E^{1/3} . \tag{3}$$

Here, E is a finite amount of energy suddenly added to a point. In our experiment, that of the YAG laser is E(J)=450 mJ, when one obtains r^* = 0.01 m. In this connection, Ayrault [10] reported experimental results that follow $r^{-7/6}$, not spherical diffusion r^{-1}.

Figure 4 is a peak sound pressure vs. propagation distance calculated by Yuldashev [6]. This result shows that when the peak pressure decreases around 70-80 Pa, the effect of nonlinearity is almost negligible, and the classical absorption and relaxation become dominant. Therefore, this shock wave approaches linear wave after r = 1-2 m. It should be noted that this calculation is for the overall value of the shock wave.

Figure 5 shows the distance attenuation of the square integral of the sound pressure measured in an anechoic chamber (the peak value also leads to the same result). Equivalent results can be obtained for a 1/3 octave analysis in the 1.25-160 kHz range. Within the scope of our report, the inverse square law holds at least farther than about 1 m (>> r^*, Eq. (2)) when the air absorption is corrected after bandpass filtering.

In this connection, there are two measurements of LIB peak pressure (OA value) attenuation with distance in a free field when the laser's output energies are similar order to this report. Attenborough and Qin [4] stated that SPL within a source-receiver distance of less than 1.5m was sufficient to result in a nonlinear effect. Bolanos et al. [5] wrote that one might neglect nonlinear effects beyond 1m of propagation, and the equivalent bandwidth of a 1/10 scale model spans from 100 Hz up to 10 kHz in the full scale case.
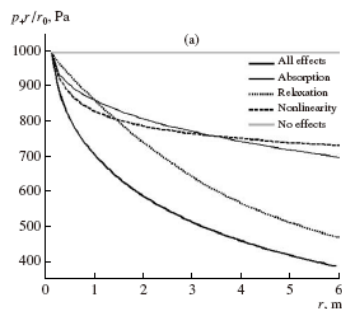

Fig. 4 Dependence of the peak positive pressure of the N-pulse on the propagation distance after Yuldashev [6].
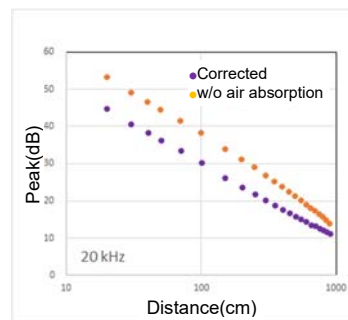

Fig. 5 Distance attenuation of sound pressure after 20 kHz 1/1 octave band pass filtering, with and without air absorption correction.

2.1.3 Directivity

Traditionally, spark discharge has often been used as a sound source in scale model experiments. However, this sound is nearby a dipole and cannot be omnidirectional. Theoretically, when a small region of the fluid is in violent motion, the sources of sound, which feed energy into the acoustic field, are usually of three sorts, monopole, dipole, and quadrupole [11]. The monopole sources correspond to of heat when a modulated laser heats the fluid. Therefore, the LIB sound source can be regarded as omnidirectional. Figure 6 shows the directivity pattern measured at 1 m from the acoustic center in an anechoic chamber. The directional deviation is within 0.1 dB in all directions up to the 160 kHz 1/1 octave band.

Fig. 6 Directional pattern of LIB measured in an anechoic chamber. The peak values are plotted.

2.1.4 Repeatability (Time invariancy)

The repeatability of the waveform generated by LIB is stable [5]; little waveform change is observed even after 1000 repetitions. The sound energy is large enough to measure the precise RIR by several

tens of synchronous summation. For reference, if synchronous summation is performed over unnecessarily large numbers, the late reverberation in the RIR slightly dissipates with time due to the time variance of the medium. It is worth noting that this results in excessive attenuation (cancellation) of the later part, similar to random signal summation.

2.2 Equivalent spherical microphone measurement

To minimize errors at the RIR measurement, a single 1/8-inch microphone (B&K, type 4138) is moved around at a receiving point in the audience area using an automatic stage (Chuo Precision Industrial, ALD-4011-G0M + LV- 4042-1). The spatial accuracy of this automatic stage is within 0.03 mm, which is 1/100 times the wavelength of a 100 kHz sound wave (Fig. 7).

Seven microphone positions are chosen for the 1st order ambisonics. For the second order Ambisonics, they are the desired receiving points (No.1 in Fig. 8, Right) plus six symmetrical points (No.2 to 7) along the x, y, and z axes. In order to obtain the proper S/N ratio of RIR. 11 microphone positions are used by including four additional points (No. 8 to 11), while the requisite number is nine mathematically. With these modifications, numerical errors in the pseudo-inverse matrix estimation can be improved. From the physical limitation of the microphone arrangement, we chose the 4th order Ambisonics for 1/10 scale measurement (Fig. 8, Left).

In order to get a proper S/N ratio for the RIR measurement, the total frequency range was divided into three. Accordingly, the distance d (2 times of radius r of the equivalent spherical microphone) between the opposing microphones was varied to avoid spatial aliasing at each frequency range. Here, we chose 2, 4, and 8 mm as the radius r that covered each of the frequency ranges.
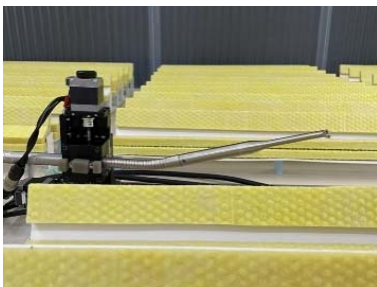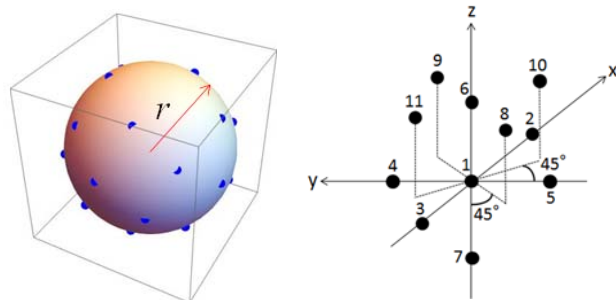

Fig. 7 3D automatic stage.

Fig.8 Left: microphone positions of equivalent spherical microphone (N = 25 for 4th order ambisonics); Right: that for 2nd order ambisonics.

2.3 RIR compensation steps

2.3.1 Inverse filtering

The amplitude spectrum of a sound generated by LIB is smoother comparing other sound sources for scale model measurements. Hence, a simple deconvolution is applicable to obtain a desirable amplitude response within the audible frequency range. Figure 3 shows that the flat response from 500 to 200,000 Hz is available when the waveform at 1 m from the LIB is used as a reference.

2.3.2 Nonlinearity

RIR should be a measurement in the linear domain for auralization. As mentioned above, the nonlinearity disappears around the source-receiver distance r exceeds 1 m in the model. In comparison, if r is less than 1 m, RIR needs to be corrected according to Eq (2).

2.3.3 Dynamic range

Auralization requires a broader dynamic range than objective parameter measurements, say 60 dB or more. However, such a value is too wide to fulfill at the RIR measurements, so post-signal processing is necessary. If the background noise of the measurement system is larger than the late part of RIR, the acoustic quality of the resulting sound after being convoluted with a musical signal would deteriorate.

We decomposed the measured RIRs into 1/1 octave bands and expanded the dynamic range at each band by linear extrapolation, if necessary, as shown in Fig. 9. Then, we employed the average slope of the initial part from -5 to -25 dB range. This procedure is based on the truth that the reverberant component has the Gaussian noise statistics and the phase information of the later reflections does not substantially influence on listener's subjective impression [12].

2.3.4 Air absorption

In an actual concert hall, variation in air humidity is more sensitive to sound attenuation than that in temperature (Bass et al., 1995). However, in a scale model experiment, the temperature has often more effect on sound attenuation.

Figure 10 shows the sound pressure attenuation for a propagation distance of 34 m (duration in RIR is 1 s in 1/1 scale). At ordinary temperature and humidity, the attenuation gradient with temperature change is most steep at 20-40 kHz. This frequency range has a notable effect on the acoustical quality of the synthesized sound, considering the ears' sensitivity. For example, at 20 kHz, under a standard condition ($20°C$, 50%), a 1% difference in humidity results in a distance attenuation of -11.6 dB →1.8 dB, whereas a $1°C$ temperature difference results in that of -11.6 dB → -12.6 dB. The latter is four times larger than the former (0.25 dB vs. 1.03 dB). Accordingly, accurate measurement of the temperature and humidity in a model and the maintenance during the experiment are vital.

So far, frequency band segmentation [13] or temporal segmentation [14] have been proposed to correct the air absorption attenuation in RIRs. As both methods use a single value of the air absorption coefficient in each segmented band or time, the resultant RIRs sometimes include numerical errors. In contrast, we have proposed a mathematical method for continuous frequencies based on Fourier transform pairs to obtain more reliable RIR measurements. This method does not introduce any numerical errors caused by the approximations by segmentation. The details of this method are omitted here due to space limitations, and further information is given in [3, 15].

To be sure, let p_m(t) be the RIR observed in the 1/10 model, then the real-scale RIR is given by

$$p_r(t) = p_m(t/10)\exp\{(m_m/10 - m_r)ct/2\} \qquad (4)$$

where m_m and m_r are the air attenuation coefficients in the scale model and at 1/1 scale (for example, $20°C$, 50%), respectively, and c is the speed of sound. From this, the relationship between the band passed reverberation time RT_m and RT_r in the model and at the real scale is given by

$$RT_r = \frac{10RT_m}{1-(c/6\ln10)\Delta_m \cdot 10RT_m} \qquad (5)$$
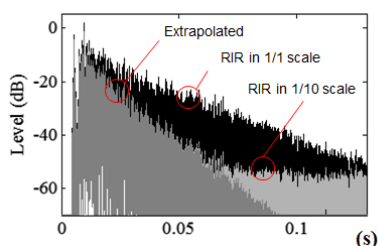
$$\Delta_m = m_m/10 - m_r$$

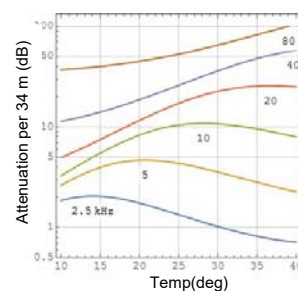Fig. 9 Extrapolation and air absorption compensation of RIR.

Fig. 10 Attenuation by air absorption for propagation distance of 34 m at humidity 50 %.

2.3.5 Effect of boundary layer

Most of the interior walls of a hall can be regarded as rigid, except for the seating areas and intentionally absorbing surfaces for sound control. A boundary layer is generated on the rigid wall due to viscosity and heat conduction, resulting in an admittance component given by [9, p. 508-534].

$$\beta = \frac{1-i}{2^{1/2}}\sqrt{\frac{\omega\mu}{\rho c^2}}\left(\sin^2\theta + \frac{\gamma-1}{\Pr^{1/2}}\right)$$ (6)

where μ is the viscosity coefficient of air, Pr is the Prandtl number, ω is the angular frequency, ρ is the density of air, and γ is the specific heat ratio. This term is an unavoidable loss that occurs independently of the scale dimension and causes a significant numerical impact in scaled model experiments. Using β in Eq. (6), the random-incidence sound absorption coefficient is given by the following equation (Fig. 11):

$$\alpha_{stat} = \int_0^{\pi/2}\left(1 - \left|\frac{\cos\theta - \beta}{\cos\theta + \beta}\right|^2\right)\sin 2\theta d\theta$$ (7)

This term should be added to the residual absorption coefficient used to calculate the reverberation time. Here, RTs' in the scale model are given by (V and S are those for the scale model),

$$RT_m = 0.161\frac{V}{S\bar{\alpha} + 4m_m V + S_{res}\alpha_{stat}} \quad : \text{with the loss.}$$ (8)

$$RT'_m = 0.161\frac{V}{S\bar{\alpha} + 4m_m V} \quad : \text{w/o the loss.}$$ (9)

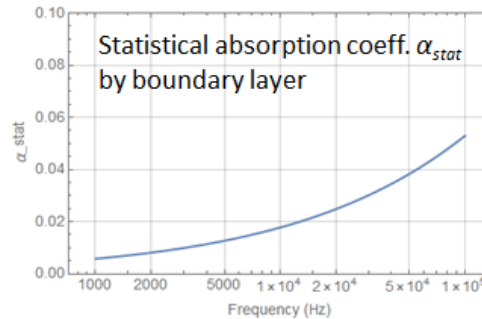$$RT'_m = \frac{1}{\dfrac{1}{RT_m} - \dfrac{\alpha_{stat}S_{res}}{0.161V}}$$ (10)

Fig. 11 Residual absorption coefficient caused by visco-thermal loss.

2.3.6 Ambisonic signal processing

In the 1/10 scale model, the microphone positions ($N = 25$) for 4th-order ambisonics are the physical limitation, which is determined by the setting ability of the 3D microphone stage. This step consists of the usual signal processing in ambisonics: the encoding, low pass filtering for the spatial aliasing, high pass filtering against the noise-boosting, and the decoding.

3   ACOUSTIC SCALE MODEL

Figure 12 shows the 1/10 scale model to be measured, a hall for 800 seats with $V = 8,000$ m^3 and S = 2,500 m^2 in 1/1 scale (dimension: 37.5m L × 20m W × 13.6m H). Since this model does not simulate any actual hall, but a simplified model for only the experiment, no balconies or other large-scale irregularities are provided. Therefore, several spherical diffusers are attached to improve the diffusivity in the hall. In addition, to suppress the generation of standing waves, the lower side walls are corrugated. Also, resonant absorbing materials are installed to suppress standing waves in the

upper stage space. The audience chairs are installed on the main floor, which is the same model used in our previous experiments. The reverberation time of the model is shown in Fig. 13.

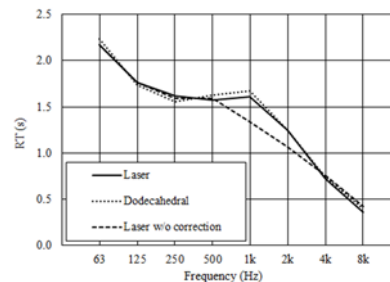Fig. 12 1/10 scale model for the measurement.

Fig.13 RT measurement at the receiving point in 1/10 scale model.

4   ANECHOIC RECORDING OF ORCHESTRA

Each orchestral instrument was individually recorded by a member of the Tokyo Philharmonic Orchestra in the anechoic chamber of the Takenaka Research & Development Institute. The volume of the anechoic room was 642 m^3 (8.4 m L × 7.8 m W × 9.8 m H) with a cutoff frequency – at which the normal incident sound absorption coefficient fell off to 0.99 – of 60 Hz. This recording procedure was basically the same as that stated in the literature [16,17]. First, to maintain synchronization of the performance, a video consisting of an audio signal of the piano performance with a complete score and an image of a conductor conducting it was prepared in advance. The orchestra members played in the anechoic chamber, following the conductor on a video monitor while simultaneously listening to the piano accompaniment through headphones. The total number of players was 31. For the first violin, three concertmasters and one associate principal (Vorspieler) participated. For the other instruments, the principal players participated. This recording was performed in just temperament.

From classical, romantic, and modern categories, the following pieces were chosen for the sake of diversity in dynamic range, register, and tempo; dramatic orchestration; and general popularity.

・ Symphony No. 41 in C major, K. 551 by W.A. Mozart, 4th movement Allegro molto, bar 272－423 (end), duration：138 sec

・ March of the Swiss Soldiers from Guillaume Tell Overture, by G. Rossini, Allegro vivace, bar 226－387, duration：130 sec

・Symphony No. 9 in D minor by A. Bruckner, 1st movement Moderato, bar 518－567 (end), duration：102 sec

・Infernal Dance from Ballet Suite (1945) The Firebird by I. Stravinsky Vivo, bar 1－98, duration：107 sec

5   MULTIPLE   SOURCE   MEASUREMENT   AND   3D REPRODUCTION

Figure 14 shows an instrument layout the Tokyo Philharmonic Orchestra selects at regular concerts, on which sound source positions are overlaid. This configuration is a German-style layout in which the first and second violins are positioned on both wings in front of the stage. For the string sections, corresponding source positions were distributed so that the distance between adjacent sources was shorter than the minimum audible angle as seen from the receiving points. For wind instruments, one sound source corresponded to one instrument.

As shown in Fig. 14, string sections were simulated with fewer sound sources than the actual number of players. Therefore, the playback levels were calibrated to reciprocate 16, 14, 12, 10, and 8 players for the first violin, the second violin, the viola, the cello, and the double bass.

According to the signal processing above, up to 32 anechoic recordings were assigned to the 46 RIRs, which were mixed down to make final acoustic signals. The synthesized sound was presented to the listener in a 4th-order Ambisonics sound field in a semi-anechoic chamber (7.5 m L × 5.5 m W × 5.7 m H). Thirty-one loudspeakers (B&W-685S2) were placed regularly on an upper hemispherical plane with a radius of 2.5 m in the chamber, and they were directed to the listener's head (Fig. 15). Throughout this study, recording, playback, and digital editing were controlled by using Nuendo (Steinberg Co., Germany).
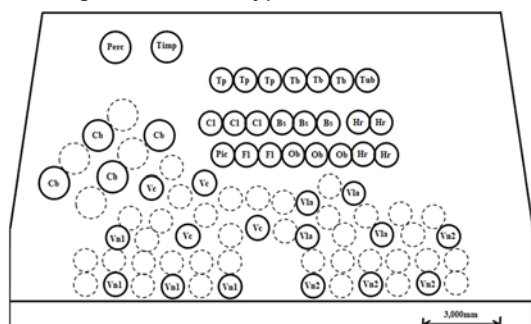

Fig. 14 Arrangement of the sound sources on the stage for Bruckner. Dotted circles are the positions of string players, which are not simulated.

Fig. 15 Semi-anechoic chamber for presentation.

6   CONCLUSION

The auralization of orchestral music is one of the ultimate challenges in room acoustic design. In this study, we might have got closer to this goal by employing a laser as the sound source for 1/10 scale model experiments. However, further new issues have become apparent through this research. Among them, the effect of sound absorption and scattering by the musicians on stage and the sound source directivity (i.e., omnidirectionality causes unnatural sound under some conditions) have vital implications for realizing a more natural auralization. These will be discussed in the presentation.

7   REFERENCES

1.   T. Hidaka et al. (2010), Proc. ICA, Sydney
2.   K. Suzuki et al. (2018), J.A.S.Jpn. 74, 244-253 (in Japanese)
3.   K. Suzuki et al. (2019), Proc. ISRA, Amsterdam
4.   K. Attenborough et al. (2002/03), 159-161 Central Laser Facility Annual Report.
5.   J. G Bolanos et al. (2013), J.A.S.A. Express Letters 133, EL221
6.   P. V. Yuldashev et al. (2008), Acoustical Physics 54, 32-41.
7.   R. D. Ford et al. (1993) JASA 94, 408-417.
8.   L. Landau (1945), J. Phys. U.S.S.R. 9, 196
9.   A. Peirce (1981), Acoustics (McGraw-Hill, N.Y.)
10.   C. Ayrault et al., (2012), Proc. Acoustics 2012, 23-27.
11.   P. M. Morse and U. Ingard (1968), Theoretical Acoustics (McGraw-Hill, N.Y.) p.322
12.   H. Kuttruff (1991), Acta Acustica United Ac. 74, 3-7
13.   X. Meynial et al. (1992), Proc. Inst. Acoust., 14, 171–177 (1992).
14.   J. D. Polack et al. (1993), J. Audio Eng. Soc., 41, 939–945.
15.   S. Koyanagi et al. (2013), Proc J.A. S. Jpn, September, 1021-1022 (in Japanese).
16.   J. Pätynen et al. (2008), Acta Acustica united Ac. 94, 856 – 865.
17.   M. C. Vigeant et al. (2008), Acta Acustica united Ac. 94, 866 – 882.