Journal of Sound and Vibration (2002) **258**(3), 527–547
doi:10.1006/jsvi.5274, available online at http://www.idealibrary.com on IDEAL®

TEMPORAL AND SPATIAL ACOUSTICAL FACTORS FOR LISTENERS IN THE BOXES OF HISTORICAL OPERA THEATRES

H. Sakai and Y. Ando

Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe 657-8501, *Japan*
E-mail: cizh-ski@asahi-net.or.jp

AND

N. Prodi and R. Pompoli

Department of Engineering, University of Ferrara, Via Saragat 1, 44100 *Ferrara, Italy*

(*Accepted* 30 *May* 2002)

Acoustical measurements were conducted in a horseshoe-shaped opera house to clarify the acoustical quality of a sound field for listeners inside the boxes of an historical opera house. In order to investigate the effects of multiple reflections between the walls inside a box and scattering by the heads of people, the location of the receiver and the number of persons in the box were varied. In each configuration, four orthogonal factors and supplementary factors were derived as temporal and spatial factors by analysis of binaural impulse responses. Each factor is compared to that at a typical location in the stalls of the same theatre. An omni-directional sound source was located on the stage to emulate a singer or in the orchestra pit to reproduce the location of the musicians. Thus, in this paper, temporal and spatial factors in relation to subjective evaluation are characterized against changes in the listening conditions inside a box, and procedures for improvement and design methods for boxes are proposed. The main conclusions reached are as follows. As strong reflections from the lateral walls of a hall are screened by the front or side walls of a box for a receiver in a seat deeper in the box, the maximum listening level (LL) in the boxes was observed at the front of the box, and the maximum range of LL values for each box was found to be 5 dB. Concerning the initial time delay gap (Δt_1), a more uniform listening environment was obtained in boxes further back in the theatre than in one closer to the stage. The subsequent reverberation time (T_{sub}) lengthens for boxes closer to the stage due to the stage house with its huge volume, and a peak is observed at 1 kHz. For the box at the back, T_{sub} monotonically decreases with frequency in the same way as in the stalls, and moreover, its values approach those in the stalls. As the contribution of multiple reflections relatively increases for a receiver deeper in the box, the $IACC$ in such positions decreases in comparison with that seen at the front of the box.

1. INTRODUCTION

The design of box seats within historical opera houses in Europe has traditionally been approached as a way of profitably expanding the capacity of the house and providing private spaces for regular visitors rather than as exercises in providing acoustically (or visually) designed spaces from which to appreciate opera. If the acoustical characteristics of the spaces inside boxes were known, it would be possible to design the boxes properly and to improve the subjective sound quality within them. Relatively few reports on systematic investigations of objective and subjective data in relation to opera houses have appeared (example are cited [1–3]). These reports have not given details of the acoustical characteristics inside the boxes. Recently, in Italy, there has been a trend towards establishing procedures for acoustical measurement inside historical opera houses [4]. Generally, the acoustical characteristics of a box are measured by placing, at its front opening, a head-and-torso simulator with tiny microphones at the ears for use as a receiver. Of course, this procedure is useful for investigating the acoustical characteristics of the box as a whole. However, when a real opera is performed, there are normally several listeners inside each box. The above measurement procedure thus neglects several acoustical effects, including scattering and diffraction by human heads and torsos and the listener's location within the box. This method thus only clarifies the characteristics of the sound field in a box for a single listener in a single location. However, in order to improve the existing conditions of listening and methods for designing boxes, knowledge of how the physical characteristics inside boxes are affected by location and the number of listeners is necessary.

This study is intended to obtain knowledge of the listening conditions for listeners inside the boxes of a historical opera house. Acoustical measurements were thus conducted in an opera house. To investigate the effects of scattering by people's heads and different listening positions, the location of the receiver and the number of people inside the box were varied, and physical factors were calculated from the corresponding binaural impulse responses. A theoretical analysis of such complex phenomena in a box including the coupled-room effects between the hall and each box and multiple reflections between walls inside the small box is difficult with current technology to calculate, so the effects were investigated by acoustical measurement. Four orthogonal factors have been proposed as objective temporal and spatial factors for evaluating subjective preference in concert hall acoustics [5]. These physical factors are practically and theoretically orthogonal, so they may, without problems, be directly adopted for the analysis of opera house acoustics. Each orthogonal factor of the listening conditions was measured in the boxes and, for comparison, at typical locations in the stalls. An omni-directional sound source was located on the stage to emulate a singer or in the orchestra pit to model the locations of musicians.

Cocchi *et al.* [6] investigated the initial time-delay gaps (*ITDG*) inside boxes, and found that, while the first reflection arrives from the walls inside the box, the strongest and most significant reflection was found to arrive later, from the side walls of the hall. They thus selected the *ITDG* at the strongest reflection from the hall. In this study, the same procedure was adopted to determine the initial time-delay gap (Δt_1), which is one of the orthogonal factors.

2. THEATRE AND BOXES

Acoustical measurements were conducted in a typical Italian opera house, the "Teatro Comunale" in Modena, shown in Figures 1 and 2(a, b). The theatre is horseshoe-shaped in plan view and has four tiers of boxes plus a gallery on the walls and a vaulted ceiling with a large chandelier suspended in the centre. There are 900 seats (2/3 in the five tiers of boxes and the gallery). The volume of each box is approximately 6 m^3, and the opening in the front of the box is 1·8 m^2 (height: 1·2 m, width: 1·5 m).

Figure 1. The interior of the "Teatro Comunale" in Modena, Italy.

3. MEASUREMENTS

3.1. CONDITIONS OF MEASUREMENTS

Maximum-length sequence (MLS) measurements were conducted to analyze the binaural impulse responses for two different source locations. The locations of sound sources and receivers are illustrated in Figure 3. Placing the sound source at the center of the stage modelled the typical standing position of a singer (2 m from the front edge of the stage, 1 m from the main axis of the theatre, and 1·5-m high). The other source position used was the orchestra pit (at the position of the first violin, 1.5 m from the pit rail, 1 m from the main axis, and 1·25-m high). The sound source was an omni-directional dodecahedral loudspeaker. Maximum-length sequence signal was reproduced through the loudspeaker. Note that the source employed for the measurements does not feature the directionality of a singer's voice [7]. One position in the stall and two in boxes, one towards the front and the other towards the back of and the third tier, were selected (see Table 1) as positions for receivers. The respective boxes will be referred to as Box A and Box B from now on. The receiver (a real person) and nearby listeners in the boxes were arranged in the four patterns shown in Figure 4. The hatched circle in the figure represents the real person's head with tiny condenser microphones placed at both ears that was used as a receiver. Those in the boxes were asked to face the sound source during the measurement. Pattern 3 was not measured for the stalls. Pattern 2 in the stalls was obtained by placing the receiver in the row behind that for Pattern 1 (i.e., in row 13, no. 7), and Pattern 4 consisted of the receiver in the same position as in Pattern 2 and neighbours in the three seats in front of that (i.e., row 12, nos. 6–8 and row 13, no. 7) were occupied. The equipment set-up used for measurement is shown in Figure 5. During the measurements, the stage was completely empty, and the chairs and music stands for the players in the pit were removed. The theatre was not otherwise occupied.

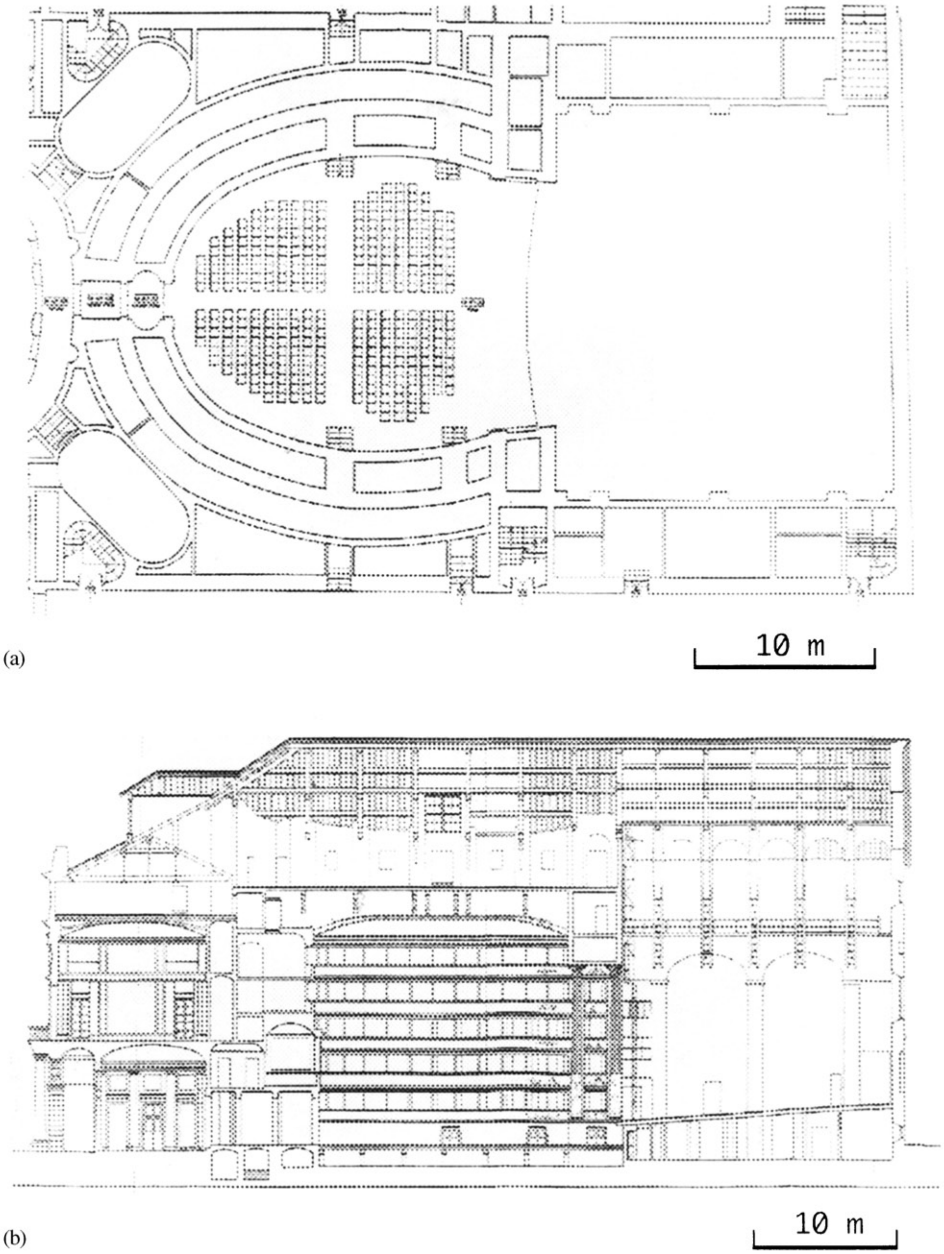

Figure 2. (a) Plan view of the theatre; and (b) cross-sectional view of the theatre.

The test signal consisted of an MLS signal with a duration of 2·97 s (sampling frequency: 44·1 kHz), which was averaged eight times to improve the signal-to-noise ratio. All of the measurement devices were controlled by a laptop PC, and the binaural impulse responses were automatically analyzed. A reference sound pressure in relation to relative listening level was measured at a distance of 1 m from both sources.

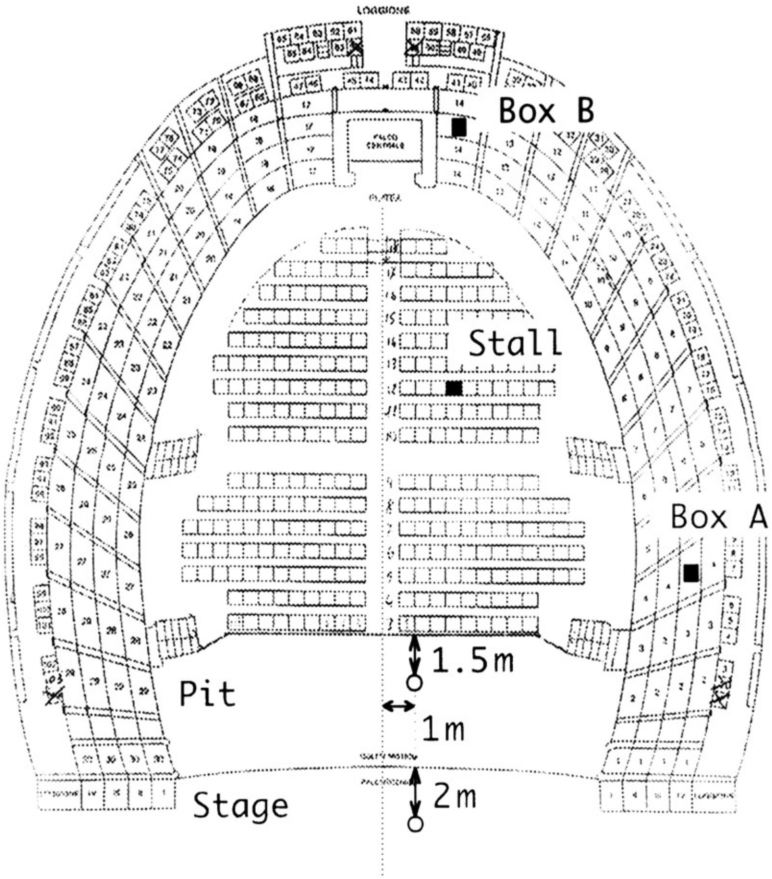

Figure 3. Locations of sound sources and receivers: ■, a receiver location; and ○, omni-directional sound source location on the stage or in the orchestra pit.

TABLE 1

Locations for the sound source and receivers

Source location	Receiver location
On the stage	Stall (row 12; no. 7)
In the orchestra pit	Box A (3rd tier; no. 4)
	Box B (3rd tier; no. 14)

3.2. PHYSICAL FACTORS USED TO EVALUATE THE SOUND FIELDS

In this paper, orthogonal physical factors that have been developed as part of subjective preference theory [5] were analyzed. The four orthogonal factors, which are the listening level (LL), the initial time-delay gap between the direct sound and the maximum reflection (Δt_1), subsequent reverberation time (T_{sub}), and interawal cross-correlation ($IACC$), and other factors, which are the total amplitude of reflections (A), interaural time delay (τ_{IACC}), and width of the interaural cross-correlation function (W_{IACC}), were extracted from the impulse responses at both ear entrances. The definitions of these factors are given

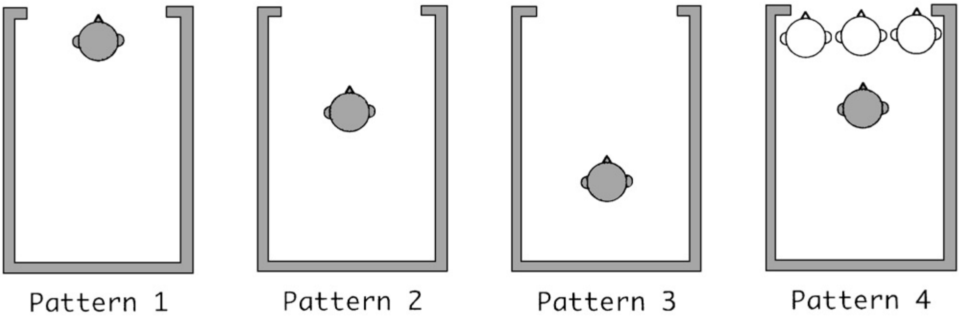

Figure 4. Placement of the listener (Patterns 1–3) and neighbour (Pattern 4). The hatched head in each pattern is that of the receiver, with microphones at the left and right ears.

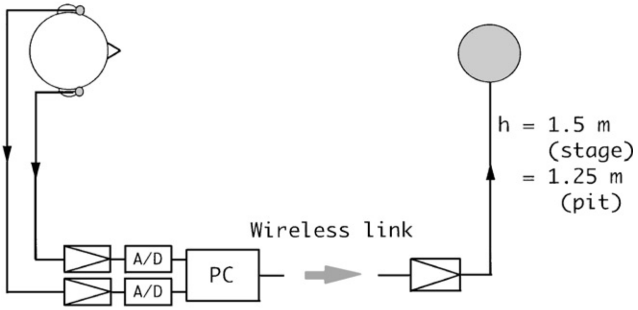

Figure 5. Setup of the equipment used for measurement.

in Appendix A. All of the factors, which reflect the subjective attributes of sound fields for listeners, are included in the binaural impulse responses, which are analyzed after being passed through an A-weighting network.

4. RESULTS

4.1. IMPULSE RESPONSES

Binaural impulse responses (over the initial 150 ms) measured at Box B when the source was on the stage are shown in Figure 6(a–d). As there were no obstacles in the direct path between the source on the stage and Box B, the maximum amplitude appeared as the direct sound with Pattern 1, as Figure 6(a) shows. The strongest reflection (with a maximum amplitude) came from a side-wall of the hall with a delay time of 26 ms, and multiple reflections between the walls inside the box due to the direct-sound component are visible before the arrival of this strongest reflection. Thus, the reflective surface of the strongest reflection can be detected by the geometrical distance from the plan of the theater. As shown in Figure 6(b, c), the strongest reflection from the sidewall of the hall, which appears in Pattern 1, is not visible at all due to the screening effect of the sidewalls of the box. In addition, the multiple reflections between walls inside the box make a relatively larger contribution than for Pattern 1. A strongest reflection (delay time: 37 ms) which is different from that seen in Pattern 1 is visible in the right-ear impulse response for

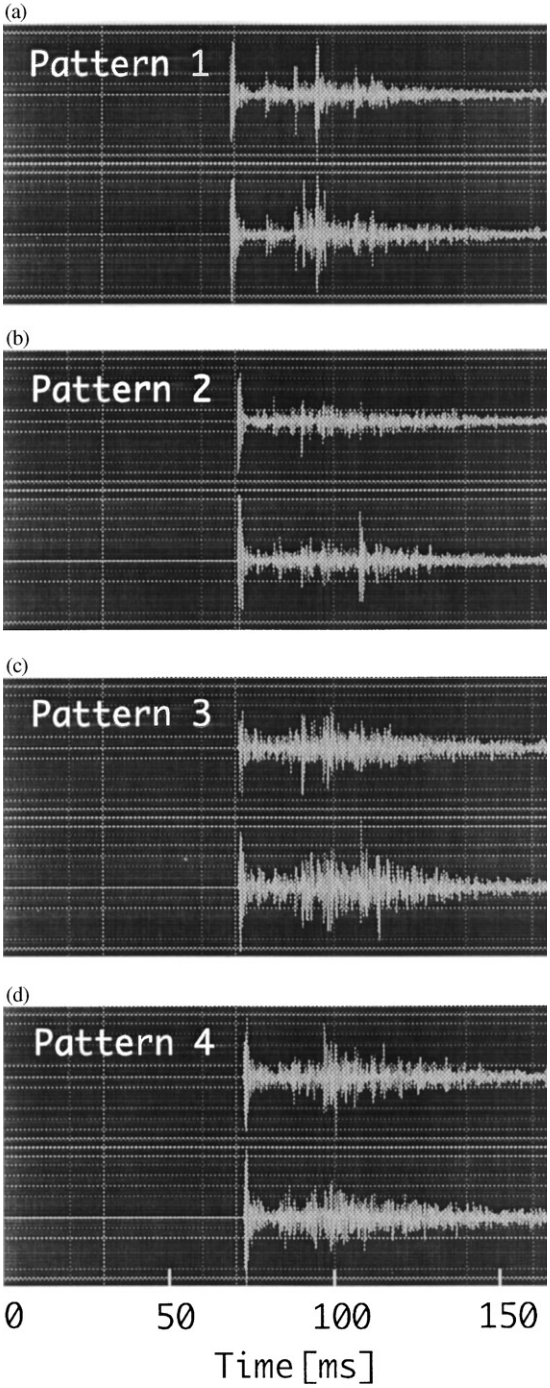

Figure 6. Binaural impulse responses in Box B as measured when the source was on the stage. Top: left-ear signal; and bottom: right-ear signal. (a) Pattern 1; (b) Pattern 2; (c) Pattern 3; and (d) Pattern 4.

Pattern 2. As Figure 6(d) shows, the contribution of initial reflections to the response in Pattern 4 was quite different from that for Pattern 2. Its complexity is due to the head scattering effect of the three people in front of the receiver on top of the effects of the

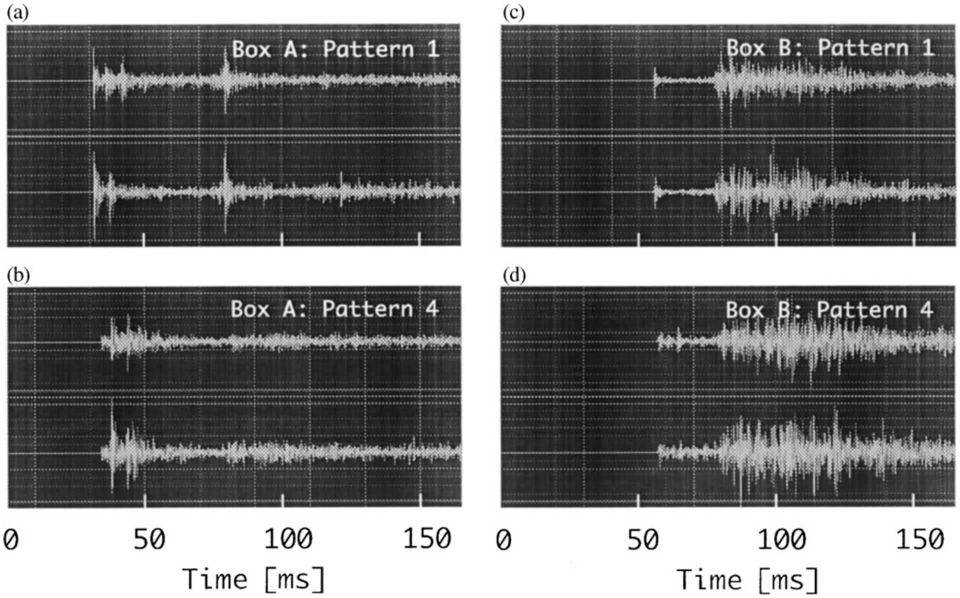

Figure 7. Examples of binaural impulse responses in boxes when the source was in the pit. Top: left-ear signal; and bottom: right-ear signal. (a) Pattern 1 at Box A; (b) Pattern 4 at Box A; (c) Pattern 1 at Box B; and (d) Pattern 4 at Box B. "S" and "R", respectively, indicate the source location and receiver location.

multiple reflections between the walls, although the position of the receiver in Pattern 4 was same as that in Pattern 2.

Figure 7(a–d) gives the initial binaural impulse responses in Boxes A and B when the source was in the pit. Results in Patterns 1 and 4 are shown for both boxes. Except for Box A with Pattern 1 [Figure 7(a)], the front wall of the box or the pit rail obstruct the direct path from the sound source. The strongest initial reflection seen 4 ms after the arrival of the weakened direct sound in Figure 7(b) might come from the ceiling inside the box. As Figure 7(c) clearly shows, the direct sound path for Box B was completely obstructed by the pit rail. In a similar way to the case for the source on the stage, described in the preceding paragraph, multiple reflections between the walls inside the boxes and scattering reflections by human heads are visible in Pattern 4 for both boxes [Figure 7(b, d)].

Thus, the contribution of reflection to the impulse responses varies dramatically according to the placement of the receiver and the number of people in the boxes. That is to say, the physical factors that are derived from the impulse responses must also be expected to vary.

4.2. LISTENING LEVEL

The measured relative *LL* in each configuration is shown as a function of 1/1 octave band center frequency in Figure 8(a–f). The left and right columns in the figure, respectively, represent the results for the source on the stage and in the pit. The top row is the case for the receivers in the stalls, the middle row is for Box A, and the bottom row is for Box B. The different symbols represent each pattern shown in Figure 4. Note that it is not possible to compare directly results for the two source locations, although the power level was same for both sources. This is because the results for each source are indicated as

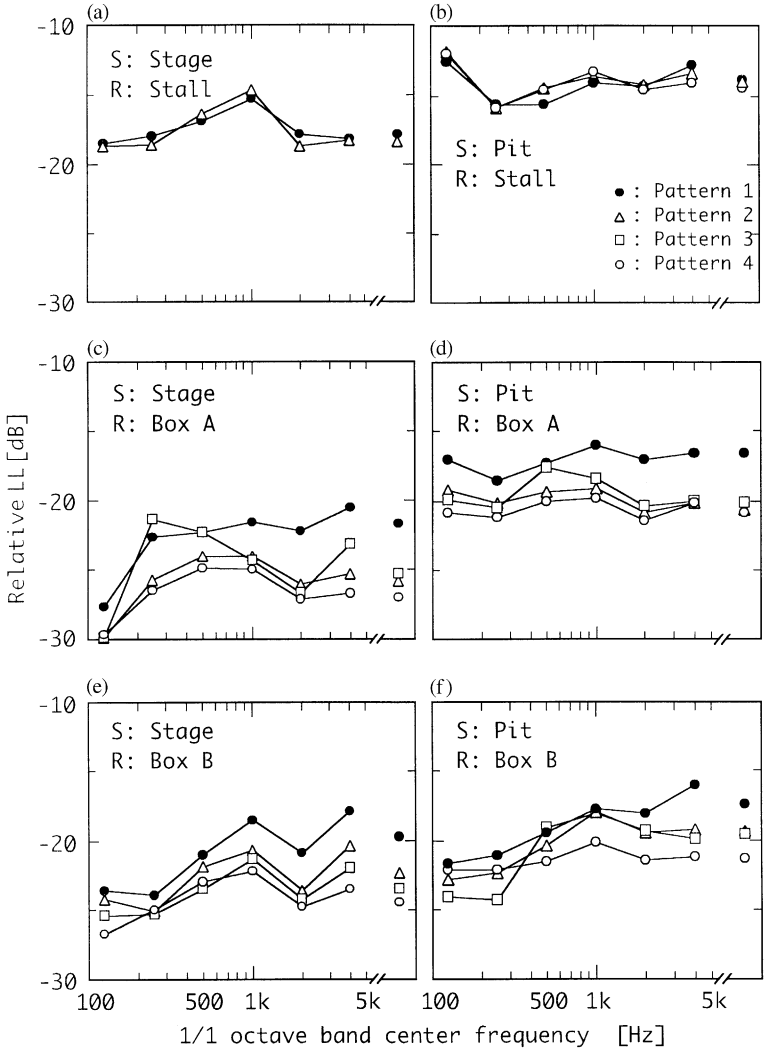

Figure 8. Measured results for the relative listening level (LL) at each 1/1 octave band centre frequency. The points plotted at the extreme right of each graph are the results of allpass-band (A-weighting): ●, Pattern 1; △, Pattern 2; □, Pattern 3; and ○, Pattern 4; S, Source location; and R, receiver location.

relative levels. To compare the results for different sources, it must be assumed that both levels of output sound power are the same in a performance of opera.

In considering the above assumption, the *LL* values for the stage source were lower than those for the pit source due to the following reason. When the omni-directional source is placed on the stage, the main reflective surface that is close to the source is the stage floor. Part of the output signal radiating from the loudspeaker and its reflection from the floor spreads into the absorptive stage house while part spreads into the hall space. On the other hand, for the source in the pit, there are five reflective surfaces near the source, namely the pit floor, the pit rail, the sidewalls, and the rear wall of the pit. Furthermore, the waves reflected by the rear wall of the pit mainly radiate into the hall area, as there is no pit overhang in this theatre.

Values of *LL* at 125 Hz were higher for the pit source, especially for the receiver positions in the stall (up to −14 dB). This boost may be caused by the interference effect, in the low-frequency range, between the direct sound and the initial reflection. For Pattern 3, the *LL* was locally boosted at 250 Hz in Box A when the source was on the stage. The decrease seen in the higher frequency range (around 2 kHz) for the stage source may be affected by the stage house, which is of huge volume and absorbs the higher frequency component.

The maximum *LL* in the allpass band for A-weighting among the configurations was in the frontal position (Pattern 1) for both the stall and in the both boxes. The *LL* values for Patterns 2–4 were spread over such a wide range in comparison with that for Pattern 1 (up to a 5-dB difference) in the boxes, although the *LL* in the stalls was almost constant for the various patterns. This fact is related to the smaller number of strong reflections in the other patterns than in Pattern 1, because the strong reflections are interrupted, within the boxes, by the front walls of the boxes (Patterns 2–4). When Figure 8(b) is compared with the other patterns, it is seen that there was less dispersion of *LL* values for the pit source in the stalls; the standard deviation of these values was less than 0.5 dB across all frequency bands. In this condition, the pit rail screened the direct sound path for all of the patterns. As has already been described, a great part of the strong reflections radiated into the hall was due to the reflections from the flat and hard surface of the rear wall of the pit. These are the probable reasons for the tighter distribution of the *LL*.

4.3. INITIAL TIME-DELAY GAP AND THE TOTAL AMPLITUDE OF REFLECTIONS

Measured results for Δt_1 are presented in Table 2. Values of Δ_1 at each ear were calculated separately, because they were quite different, especially for the receiver positions in the boxes.

Values of Δt_1 in the stall for the stage source were almost constant at around 38–40 ms. In this case, the Δt_1 values at both ears almost coincided, because the paths of reflections via the sidewall of the hall were the same for both ears. On the other hand, the values of Δt_1 in the stall for the pit source differed according to the ear, and the left-ear Δt_1 was

TABLE 2

Measured Δt_1 at the left and right ears for each receiver location, for the sound source on the stage and in the orchestra pit

		Δt_1 (ms)			
	Source	Stage		Pit	
Receiver	pattern	Left	Right	Left	Right
Stall	1	40	40	65	77
	2	38	39	57	77
	4	—	—	101	77
Box A	1	17	42	48	48
	2	7	11	10	4
	3	8	7	10	10
	4	7	11	9	4
Box B	1	26	26	28	43
	2	19	37	33	33
	3	28	26	27	26
	4	19	37	48	30

TABLE 3

Measured A values at the left and right ears for each receiver location, for the sound source on the stage and in the orchestra pit

		a value			
	Source	Stage		Pit	
Receiver	pattern	Left	Right	Left	Right
Stall	1	1·97	2·19	35·57	28·37
	2	2·26	2·10	48·32	36·02
	4	—	—	77·73	37·97
Box A	1	1·49	1·09	3·65	2·54
	2	3·12	2·39	9·72	7·27
	3	5·85	6·56	10·41	8·56
	4	4·14	4·04	8·72	9·42
Box B	1	1·89	1·84	9·41	9·71
	2	1·57	1·45	13·64	12·07
	3	2·28	2·00	19·06	23·15
	4	2·90	2·52	16·90	15·31

shorter than the right-ear Δt_1 for Patterns 1 and 2. Values of Δt_1 in the stall were generally shorter for the stage source than for the pit source.

In the results for Box A with the stage source, Δt_1 values for the two ears were also quite different only for the front position. This may be due to the asymmetrical shape of the box and its closeness to the stage. A similar tendency is not observed in the results for Box A with the pit source. The strong reflections from the flat and hard surfaces of the pit walls arrived as the first reflections, with a delay time of 48 ms for Pattern 1. In the Box B results for the stage source, such a tendency is not visible because this box is almost symmetrical about the main axis of the theatre. As a whole, the first maximum reflection arrived from the sidewalls or ceiling of the hall (Δt_1 > 20 ms), walls inside boxes (7–12 ms), or the floor of the stage or pit (4 ms).

For either source in Box A, the maximum reflection came from the wall inside the box for Patterns 2–4 ($\Delta t_1 < 11$ ms), whereas the first reflection for Pattern 1 came from the sidewall of the hall as described above. On the contrary, in Box B a significant reflection came from the sidewalls in the hall. Thus, different characteristics for this parameter were observed in the two boxes. This can be interpreted meaning that the rear box provides a more uniform listening environment than the front box in terms of Δt_1.

Table 3 shows that values of A for the stage source were lower than for the pit source, although there are some exceptions. The differences between the average values for each source location were 4.0 in Box A and 13.1 in Box B. In Box A, A values for Pattern 2–4 were greater than those for Pattern 1. This is true in all cases for both source locations, except for the case of Box B with the stage source. This indicates the increased contribution of reflection because of the multiple reflections between walls or the scattering reflection by the heads of the other listeners at positions deeper within the box.

4.4. SUBSEQUENT REVERBERATION TIME

The parameter T_{sub} did not vary widely according to receivers patterns. An average of the results for all patterns is thus given in Figure 9(a, b). When further averaged across

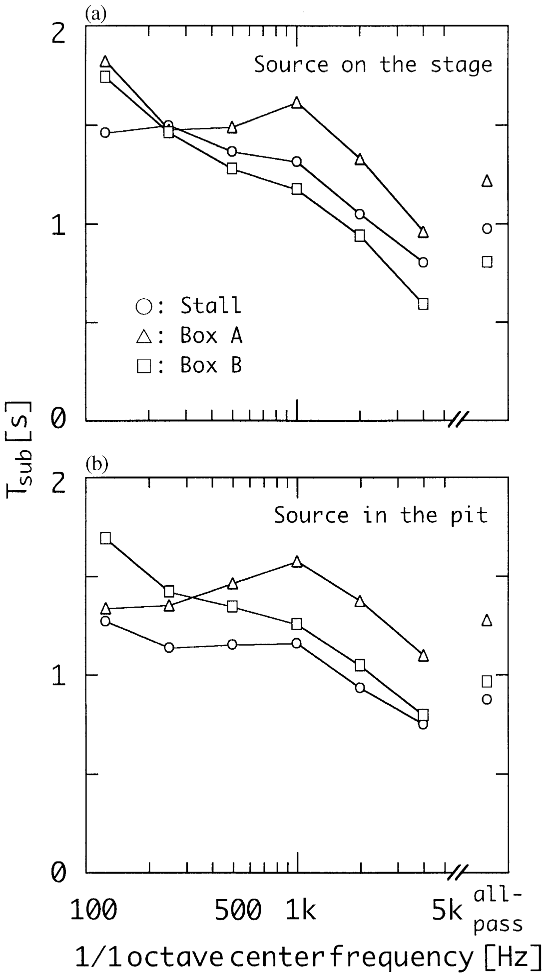

Figure 9. Measured results of subsequent reverberation time (T_{sub}) at each 1/1 octave band centre frequency. The points plotted at the extreme right of each graph are the results of allpass-band (A-weighting). (a) Source on the stage; and (b) source in the pit; ○, stall; △, Box A; and □, Box B.

location, the values for T_{sub} at 500 Hz and 1 kHz were 1·4 and 1·3 s for the stage source, and 1·3 and 1·3 s for the pit source respectively.

For both source locations, T_{sub} in the stall and in Box B monotonically decreased with frequency. In Box A, T_{sub} had a peak at 1 kHz and the T_{sub} values at frequencies above 500 Hz were greater than those in the stalls and in Box B. One of the reasons for the peak is Box A's proximity to the stage house, since this lengthens reverberation times. On the other hand, the influence of the hall was stronger than that of the stage house in results for the stalls and for Box B.

The T_{sub} values in Box B for the pit source were greater than those in the stalls below 500 Hz, although the values in the stalls and in Box B for the stage source were almost the same. More precisely, T_{sub} values at the receivers in the stalls for the pit source were uniformly shorter than T_{sub} values for the stage source. This can be considered to be because the initial strong reflected component, which arrived in the stall from the orchestra pit, was more weakened by the pit rail than in cases in Boxes A and B at their higher locations.

4.5. FACTORS FROM INTERAURAL CROSS-CORRELATION FUNCTION

Measured values of *IACC* as a function of 1/1 octave band centre frequency are shown in Figure 10(a–f). The symbols are the same as for Figure 8 for *LL*. In the stalls, ranges of *IACC* across the patterns were rather small for both sources, and the effects of the scattering by heads which decreases the *IACC* are not visible [see Figure 10(a, b)]. On the other hand, *IACC* values above 500 Hz in the stalls were smaller for the pit source than for the stage source, as the direct sound was weakened by the pit rail.

The *IACC* at frequencies above 500 Hz increased more strongly at the front of the boxes (Pattern 1) than for the other arrangements, except for the case of Box B with the pit source. This may be because the ratio of the contribution of direct and reflected sound mainly from the walls inside the boxes is greater for Pattern 1.

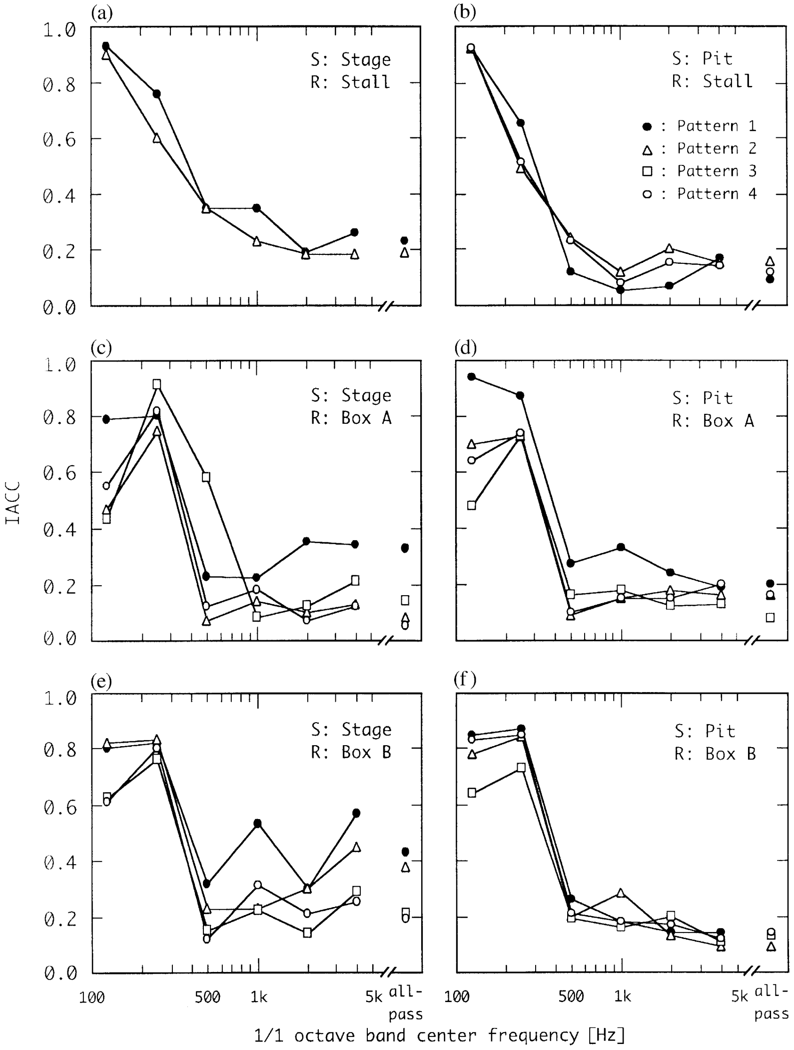

Figure 10. Measured results of IACC at each 1/1 octave band centre frequency. The points plotted at the extreme right of each graph are the results of allpass-band (A-weighting): ●, Pattern 1; △, Pattern 2; □ Pattern 3; and ○, Pattern 4; S, Source location; and R, Receiver location.

TABLE 4

Measured τ_{IACC} for the sound source on the stage and in the orchestra pit

		τ_{IACC} (ms)	
Receiver	Pattern	Stage	Pit
Stall	1	0·02	0·02
	2	0·09	0·02
	4	—	0·05
Average		0·06	0·03
Box A	1	0·02	0·05
	2	−0·02	0·11
	3	0·61	0·95
	4	−0·02	0·07
Average		0·15	0·30
Box B	1	−0·05	−0·07
	2	0·00	0·00
	3	−0·02	−0·05
	4	0·00	−0·07
Average		−0·02	−0·05

The average *IACC* values above 500 Hz for the stage source were greatest in the stalls with Box B second, and Box A least. One of the main reasons for this is the extent to which there is a direct sound path. In fact, the arrival of strong direct sound at receivers results in the *IACC* large values. However, for Pattern 1 with the stage source, the *IACC* was smaller in Box A, which is symmetrical and is close to the source, than in Box B. For the pit source, there was only a direct sound path to Box A, and the highest *IACC* values at frequencies above 500 Hz were for Box A. For both source locations, *IACC* fell more dramatically between 250 and 500 Hz in the boxes than in the stalls. In general, the *IACC* values tended to be lower for the pit source than for the stage source.

For all patterns in both boxes, there is a dip in the *IACC* at 125 Hz. As shown in Figure 10(e), the *IACC* for Pattern 1 was boosted at 1 and 4 kHz for Box B with the stage source. In general, the *IACC* monotonically decreased with frequency, as shown in the results for the stalls. However, the *IACC* decreased at 125 Hz in both boxes. This may be caused by phase differences at the ears, and the difference is especially evident in Patterns 2, 3, and 4.

In terms of the results for the allpass band, the *IACC* fell below 0·2 at those locations at which the direct sound path was screened, i.e., at locations deeper within the boxes. For example, for the source on the stage, the *IACC* was smaller in the stalls (0·23 for Pattern 1) than in Boxes A (0·33 for Pattern 1) and B (0·43 for Pattern 1).

Table 4 presents the measured results of τ_{IACC} obtained by calculating the interaural cross-correlation function (IACF) at each receiver. If τ_{IACC} has a narrow range, i.e. within $\pm 0{\cdot}1$ ms, the frontal source direction from the receiver should be clearly perceptible and such low τ_{IACC} values were obtained for almost all patterns in these measurements. Larger τ_{IACC} values were, however, obtained in some cases (for example, for Pattern 3 in Box A, with 0·61 ms for the stage source and 0·95 ms for the pit source). All of these cases have some minor IACF peaks below 0·15 within its delay ± 1 ms and there is a peak (but not the maximum) near the origin, although a subjectively diffuse sound field is perceived by humans when the *IACC* is small, i.e., less than 0·15 [5]. The W_{IACC} was almost constant around 0.4 under all conditions. This parameter is not very significant in relation to white noise or to MLS signals, because it is closely related to the sound source itself.

5. DISCUSSION

The results for *LL* indicate that if the output power level of a singer on the stage and the musical instruments in the pit are the same (although this is not realistic), the resulting balance in the hall will favour the instruments. However, according to a study using a professional tenor singer, a "singing formant" has been discovered at around 3 kHz [8]. This singing formant may allow a singing voice to reach the positions of listeners with sound of the instruments, even though the *LL* is lower for a stage source than for a pit source.

With the boxes, *LL* inside (Patterns 2–4) was lower in the dream position than at the front of the box (Pattern 1). This is, of course, not the case in the stalls. In order to reduce this attenuation of *LL*, careful design must ensure that strong reflections from all possible source locations come into the positions deeper within each box.

It is interesting to note that the left-ear and right-ear Δt_1 values in Patterns 2 and 4 were quite similar in both boxes for both source positions, with the only exception being Box B for the stage source. This means that Δt_1 is at almost the same value for a person in the rear seats of a box almost independently of the box's position, whether or not there are listeners in the front row or the listener is in the second or third row. The difference between left and right Δt_1 values in Box B for the stage source may be because the sidewalls of the box screen the strong reflections from the hall.

The Δt_1 values in the stalls were almost all shorter for the stage source than for the pit source. Thus, in the stalls, the relationship between the effective duration τ_e of a source signal and the optimum Δt_1 for listeners is relatively consistent. That is, the τ_e value should generally be lower for a singing voice on the stage than for music from the orchestra pit. On the other hand, this tendency was not obtained for any of the patterns in the boxes. In terms of Δt_1, this tendency can be interpreted as meaning that listening conditions in Box B are more uniform than those for Box A. The sidewalls of boxes should be properly designed to improve the unevenness of conditions in frontal boxes.

Thus, Δt_1 values were clearly found to vary according to the number of people and their locations. In general, however, the binaural analysis showed that the presence and location of the listener and of "dummy" listeners creates changes in Δt_1 that may alter listening conditions.

According to preference theory, the maximum amplitude is dominant in determining Δt_1 even if there are other earlier reflections [9]. However, in the previous investigation cited, this condition for amplitude of reflection was found to apply less strictly when they are close. In the present investigation, reflections arrived in the boxes from the walls of the boxes as well as from the sidewalls of the hall. A more extensive study of this effect of the dominance of reflection on the basis of psychological activity is yet to appear. In particular, the earlier first reflections have been neglected in a concert hall acoustics and treated as reflections from the stage floor though there are many other early reflections such as reflections between the inner walls of boxes. Psychological experiments are necessary to establish which reflection is dominant in subjective preference, for such complicated cases as the difference at the left and right ears. Early reflections within 20 ms can consequently be considered as not being perceived by the listener, and as merely enhancing loudness. However, it is not yet known how such a case (screening of the strong initial reflection by the front wall of a box as occurs in Patterns 2–4, with multiple weak reflections remaining) affects subjective evaluation. In the acoustical design of boxes, many strong reflections from the hall must be made to arrive inside the depth of the boxes. In addition, this causes a clear peak in the IACF determining *IACC*, and affects the localization of sound sources.

As stated in the Introduction, Cocchi *et al.*, [6], described that the first significant reflection, i.e., the reflection with maximum amplitude, arrives from the sidewalls of the hall. However, in this investigation in which locations deep within boxes were included, a different characteristic of the first reflection was obtained. This includes the screening effect of the sidewalls of the boxes and multiple reflections between the walls of the boxes.

In the present measurements, T_{sub} was about 1·3 s in the middle frequency range. According to subjective preference theory, the optimum T_{sub} is related to the source signal [5]. For a vocal source, a T_{sub} of 1·5 s is considered too long. However, for operatic performances, an experiential listening condition may be included that make such longer reverberation times match the performance. In any case, the inconsistency of optimum T_{sub} in relation to different source types in the same theatre remains an unresolved topic, and psychological data for sources, which consist of an orchestra and a vocal element, is currently lacking.

T_{sub} (as well as RT, EDT, etc.) has been evaluated as an average of values for several positions in the theatre. However, as indicated in the results of measurement, values of T_{sub} are distributed over a wide range according to whether they are for stalls or boxes and the positions of boxes positions. For instance, T_{sub} values at 1 kHz for the stage source were 1·6 and 1·2 s in Boxes A and B respectively. In addition, frequency characteristics varied according to positions. Values of T_{sub} are distributed over a wide range according to the listening location, including the positions of box seats, although they do not vary according to seating position in a normal concert hall. When considered as subjective evaluation that takes preference into account, this means that the preferred listening position in relation to T_{sub} is more important in an opera house than in a concert hall. Note that the variation in T_{sub} in a given box is negligible.

In the stalls, *IACC* was not very small even with three people in front of the receiver. Considering that *IACC* values become low in boxes at frequencies above 500 Hz, the contribution of the strong reflections between the walls inside a box is more significant in decreasing the *IACC* than the effect of scattering by listeners' head. This is explained by the fact that the multiple reflections inside a deep box increased relatively more than at the front of the box, as was described in the passages on the measured results for impulse responses.

For a source on the stage and a receiver in a stall, sharp peaks in the IACF, which determines *IACC* values, were clearly obtained, just as they are in a concert hall where there are always direct sound paths. In Box A, the *IACC* for Pattern 1 has a major peak for the stage source, but the peaks for the other patterns become vague as is shown in Figure 11(a–d). A clear and sharp peak was obtained in Box B for the same source when there is no pit rail to screen the direct path, as is shown in Figure 12(c, d). Conversely, a vague and minor peak was obtained when the direct sound path is screened by the pit rail, as is shown in Figure 12(a, b, e, f). In such cases, visual information may be more significant information than acoustic for the localization of sound.

In concert hall acoustics, a hall is evaluated by using the orthogonal factors at each seat to analyze preference scale values. However, it is quite difficult to evaluate sound fields inside boxes. This is because two temporal factors (Δt_1 and T_{sub}) have greatly different characteristics from those they have in concert halls. For Δt_1, some initial reflections come from the walls inside a box, and identifying the characteristics of the source signals is quite complicated. That is why, in this paper, the scale values were not calculated to evaluate the sound fields inside boxes.

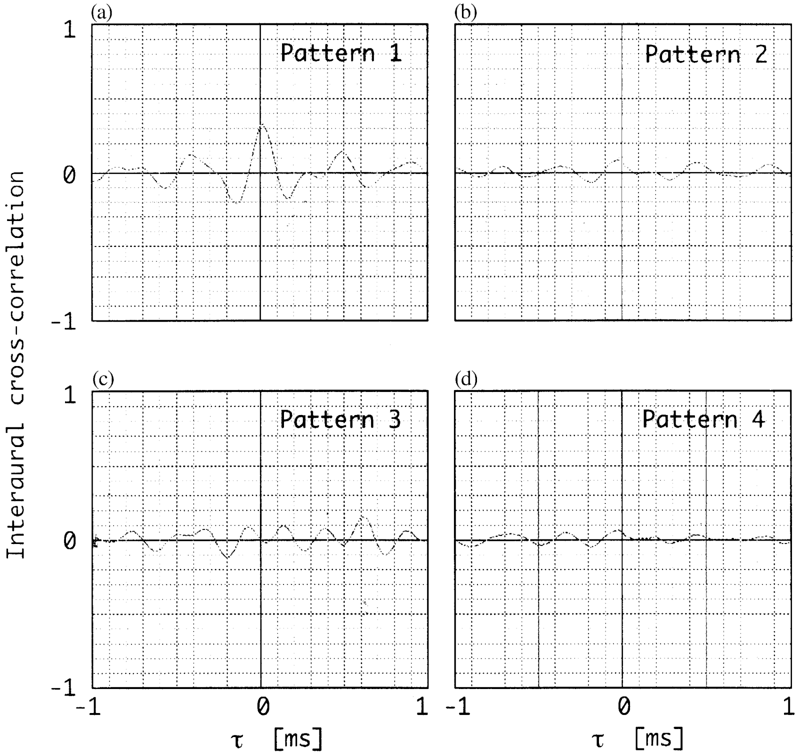

Figure 11. Interaural cross-correlation functions (IACF) in Box A when the source was on the stage: (a) Pattern 1; (b) Pattern 2; (c) Pattern 3; and (d) Pattern 4.

6. REMARKS

In this paper, the listening conditions in the boxes of a historical Italian opera house have been partially characterized by using the orthogonal factors of a sound field as derived from the results of measurements which were conducted with variation of the locations of sound sources and configurations of receivers. Temporal and spatial factors in relation to subjective evaluations and the shapes of the impulse responses were found to vary listening conditions. These factors included the sound source location, the receiver location, and configurations of people in a box. This is due to the multiple reflections between the walls inside a box and the effect of scattering by the heads of people close to the receiver, in addition to the known effects of the interruption of the direct sound path. Consequently, psychological response can also be assumed to vary. In addition, proposals for the improvement of the sound fields of boxes and some thoughts on design procedures were presented.

Concluding remarks on the measurements follow. (1) In the impulse responses measured in boxes, many reflections arrive between arrival of the direct sound and of the first reflection that comes from the sidewalls of the hall. These multiple reflections are between the walls inside the box and because of scattering by the heads of the people close to the listener. These reflections are effective in decreasing the *IACC* at listener positions in boxes. (2) At receivers deeper within a box, as the strong reflections from the sidewalls of the hall are screened by the front and sidewalls of the box, the maximum *LL* for the A-weighting allpass band in the box is seen at its opening. The range of *LL* values in a

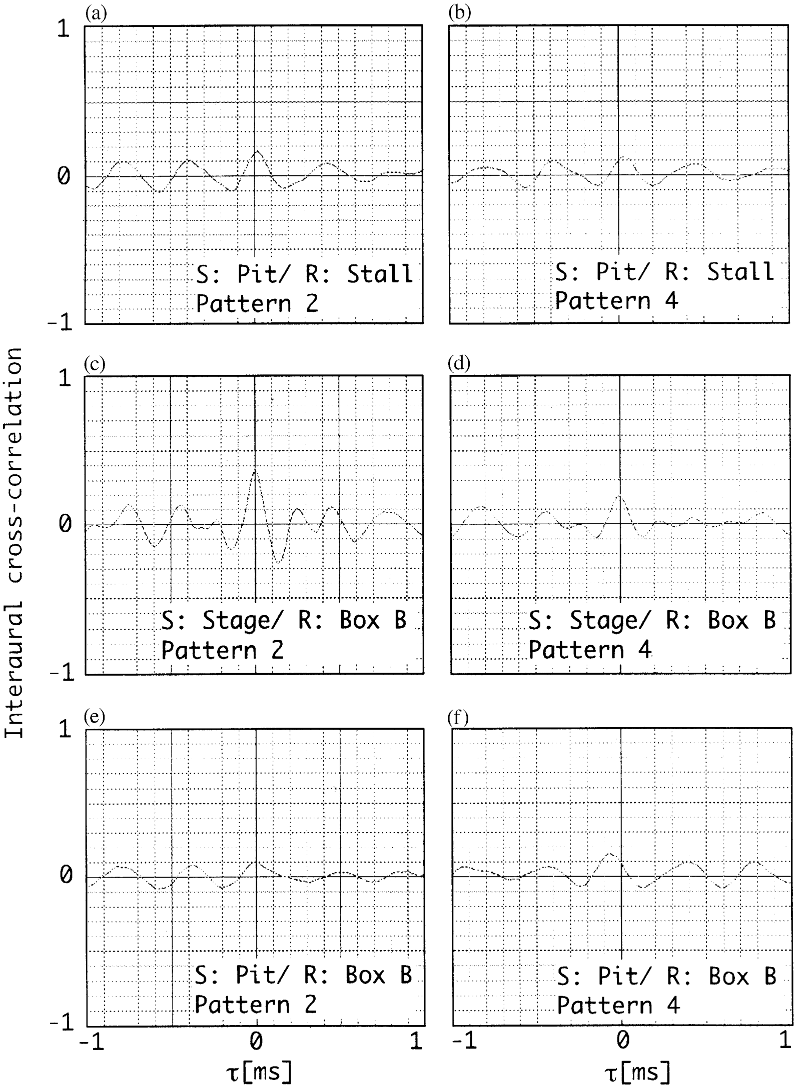

Figure 12. Typical results for interaural cross-correlation functions (IACC). (a) Pattern 2 in the stall (source in the pit); (b): Pattern 4 in the stall (source in the pit); (c) Pattern 2 in Box B (source on the stage); (d): Pattern 4 in Box B (source on the stage); (e) Pattern 2 in Box B (source in the pit); and (f) Pattern 4 in Box B (source in the pit).

given box is up to 5 dB. (3) Values of LL at 125 Hz are boosted for the pit source, especially at receivers in the stall (by up to -14 dB), due to the effects of interference between the direct sound and the initial reflections between walls inside the orchestra pit. (4) Regarding the factor Δt_1, the more uniform listening environment is provided in the rear box. The Δt_1 value is almost the same for people in rear seats almost independently, whether or not there are listeners in the front row or the listener is in second row. (5) At the receiver positions deeper inside the box, the reflection with the maximum amplitude does not always arrive from the sidewalls of the hall due to the screening effects of the walls of the box and the multiple reflections inside the box. (6) The T_{sub} values for the various receiver arrangements were distributed over a wide range, but were not affected by the listening position or the number of people in the box. In the box close to the stage, T_{sub}

was lengthened by the effect of the stage house, and had a peak at 1 kHz. In the box further back, a monotonic decrease of T_{sub} with frequency was observed and this characteristic is similar to that of T_{sub} in the stalls. Such a variation of T_{sub} results suggests the importance of the selection of a preferred listening position in an opera house, as compared with the easier situation for a concert hall. (7) At the receiver positions deeper inside a box, *IACC* is normally smaller (i.e., better) than at the opening. This is because the contribution of multiple reflections is relatively greater in the deeper position. For both source locations, the *IACC* dramatically decreased between 250 and 500 Hz in the boxes than in the stalls.

ACKNOWLEDGMENTS

The authors are deeply indebted to the stuff of the "Teatro Comunale" in Modena for providing the opportunity to make the measurements. The authors also wish to thank Francesco Pompoli for his cooperation and the considerable assistance he gave us with our measurements.

REFERENCES

1. M. Barron 1993 *Auditorium Acoustics and Architectural Design*. New York: E&FN Spon chapter 9.
2. L. L. Beranek 1996 *Concert and Opera Halls: How They Sound*. Woodbury, NY: Acoustical Society of America.
3. T. Hidaka and L. L. Beranek 2000 *Journal of the Acoustical Society of America* **107**, 368–383. Objective and subjective evaluations of twenty-three opera houses in Europe, Japan, and the Americas.
4. R. Pompoli and N. Prodi 2000 *Journal of Sound and Vibration* **232**, 281–301. Guidelines for acoustical measurements inside historical opera houses: procedures and validation.
5. Y. Ando 1998 *Architectural Acoustics—Blending Sound Sources, Sound Fields, and Listeners*. AIP Press/ Springer-Verlag. New York: AIP Press/ Springer-Verlag.
6. A. Cocchi, M. Garai and C. Tavernelli 2000 *Journal of Sound and Vibration* **232**, 171–191. Boxes and sound quality in an Italian opera house.
7. A. H. Marshall and J. Mayer 1985 *Acustica* **58**, 130–140. The directivity and auditory impressions of singers.
8. J. Sundberg 1977 *Scientific American* **236**, 82–91. The acoustics of the singing voice.
9. Y. Ando and D. Gottlob 1979 *Journal of the Acoustical Society of America* **65**, 524–527. Effects of early multiple reflections on subjective preference judgments of music sound fields.
10. M.R. Schroeder 1965 *Journal of the Acoustical Society of America* **37**, 409–412. New method of measuring reverberation time.

APPENDIX A: DEFINITIONS AND PROCEDURE FOR CULCULATING THE PHYSICAL FACTORS

All of the physical factors used in this paper were calculated from the binaural impulse response $h_{jl,r}$. Index j indicates samples of an MLS taken at a constant time interval σ ($j = 0, 1, 2, \ldots, L—1$). Indices l and r represent the left and right ears respectively.

A.1. LISTENING LEVEL *LL* AND TOTAL AMPLITUDE OF REFLECTIONS *A*

The listening level, *LL*, is defined in dB as the sound pressure level SPL at each ear of the receiver relative to the SPL at a reference position. For each receiver location, the values of LL in each 1/1 octave band (six bands from 125 Hz to 4 kHz) and the A-weighting allpass band were obtained. Usually, the LL is given as a geometric mean of the left-ear and right-ear LL values. The LL at each ear is calculated as auto-correlation function $\Phi_{ll,rr}(\tau)$ at $\tau = 0$ of the impulse responses $h_{jl,r}$

$$\Phi_{ll,rr}(0) = \sum_{j=0}^{L-1} h_{jl,r}^2. \tag{A1}$$

Relative LL is obtained by

$$LL = 10\log_{10}\frac{\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)}}{\Phi^{(ref)}(0)} \quad \text{if } h_{jl,r} \neq 0, \tag{A2}$$

where

$$\Phi^{(ref)}(0) = \sqrt{\Phi_{ll}^{(ref)}(0)\Phi_{rr}^{(ref)}(0)}. \tag{A3}$$

Here, $\Phi^{(ref)}(0)$ is the geometric mean of the auto-correlation functions of the binaural impulse responses for $\tau = 0$ at the reference position.

The value of the total amplitude of reflections A is calculated from $h_{jl,r}$ as the ratio between the energy in the direct sound and the early plus-subsequent reverberation.

$$A = \sqrt{\sum\nolimits_{\varepsilon+1}^{L-1} h_j^2 / \sum\nolimits_0^{\varepsilon} h_j^2}, \tag{A4}$$

where ε is the short delay time that covers the duration of the direct sound. The value of A is represented as an arithmetic mean across both ears. Note that the value of A is not physically an orthogonal factor of a sound field. In fact, the value of A is strongly related to the value of Δt_1 [5].

A.2. INITIAL TIME DELAY GAP BETWEEN THE DIRECT SOUND AND THE FIRST REFLECTION, Δt_1

The initial time delay gap between the direct sound and the first reflection, Δt_1 (ms), is defined as the time interval between the direct sound and the reflection with the maximum amplitude arriving at the ears. The initial reflections Δt_2, Δt_3, ... are strongly related to the reflection with the largest amplitude, Δt_1, so factors Δt_2, Δt_3, ... are not orthogonal factors.

A.3. SUBSEQUENT REVERBERATION TIME T_{SUB}

Subsequent reverberation time T_{sub} (s) is defined as the time required for a sound to decrease by 60 dB after the arrival of the first reflection for an attenuation curve of reverberation. The logarithmic transformation of the attenuation curve is done by linear regression for the initial 10–15 dB attenuation, and the interval of 60 dB attenuation is calculated as T_{sub}. The value of T_{sub} is calculated by squaring and integrating impulse responses [10]. The values of T_{sub} for each position are given by obtaining the arithmetic mean of the left and right T_{sub} values. The results should be represented as centre frequencies of 125 Hz–4 kHz (6 bands) of the 1/1 octave band.

A.4. FACTORS OF INTERAURAL CROSS-CORRELATION FUNCTION ($IACC$, τ_{IACC}, AND W_{IACC})

The definitions of $IACC$, τ_{IACC}, and W_{IACC} as factors of the interaural cross-correlation function (IACF) are shown in Figure A1. The normalized IACF is

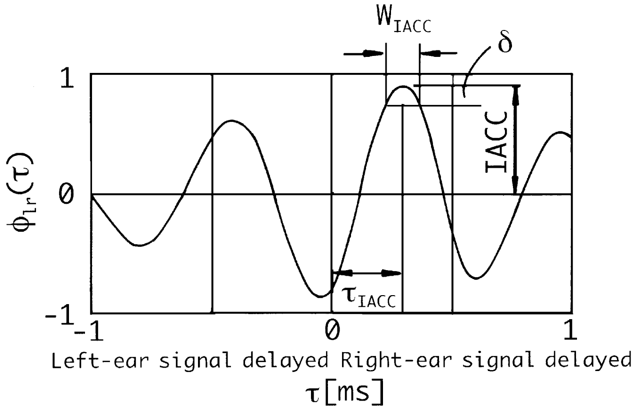

Figure A1. *IACC*, τ_{IACC}, and W_{IACC} as representative factors of the interaural cross-correlation function.

given by

$$\phi_{lr}(\mathrm{j}\sigma) = \frac{\Phi_{lr}(\mathrm{j}\sigma)}{\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)}}, \tag{A5}$$

where the values of $\Phi_{ll,rr}(0)$ represent the auto-correlation functions ($\tau = 0$) of the impulse responses at both ears. The denominator is the geometrical mean of the sound energies arriving at the two ears, and $\Phi_{lr}(\mathrm{j}\sigma)$ is the IACF of the impulse responses at both ears. Time duration $\mathrm{j}\sigma$ corresponds to τ in Eq. (A6). The magnitude of the interaural cross-correlation function *IACC* is defined as

$$IACC = \left|\phi_{l,r}(\tau)\right|_{max}, \quad |\tau| \leqslant 1\ (\mathrm{ms}). \tag{A6}$$

This is a significant factor in determining the degree of subjective diffuseness in a sound field as well as subjective preference. It represents the degree of similarity in sound waves incident on the two ears.

The interaural delay time, at which the *IACC* is determined as shown in Figure A1, is denoted by τ_{IACC}. When τ_{IACC} is zero, which is one of the preferred conditions, sounds from a frontal sound source can be perceived and a well-balanced sound field can usually be obtained.

The width of the interaural cross-correlation function, W_{IACC}, is defined as the interval of the delay time at 10% below the orthogonalized *IACC*, which means Δt_1, as shown in the figure. The W_{IACC} factor is significant and is related to the apparent source width (ASW).