Chapter 9
Auralization

Takatoshi Yokota, Takumi Asakura and Takayuki Masumoto

Abstract This chapter introduces some techniques for the auralization of numerical simulations and provides an example of calculating the head-related transfer function (HRTF) which is a function that is fundamental to reproduce stereophonic sound field. In the first section, we introduce a sound field simulation technique combines a numerical simulation with a multichannel reproduction technique, and we present applications of this method to the problem on room acoustics. In the second section, we discuss an auralization technique that simulates the sound insulation characteristics of the façade of a building. In the final section, we present the results of calculating the HRTF and discuss the applicability and efficiency of the analysis method.

Keywords Room acoustics · Building façades · Insulation characteristics · Vehicle noise · Head-related transfer function (HRTF) · Multichannel reproduction · Finite-difference time-domain (FDTD) method · Boundary element method (BEM) · Fast multipole boundary element method (FMBEM)

T. Yokota (✉)
Kobayasi Institute of Physical Research, 3-20-41, Higashi-motomachi,
Kokubunji, Tokyo 185-0022, Japan
e-mail: t-yokota@kobayasi-riken.or.jp

T. Asakura
Institute of Technology, Shimizu Corporation, 3-4-17, Etchujima,
Koto-ku, Tokyo 135-8530, Japan
e-mail: t_asakura@shimz.co.jp

T. Masumoto
Cybernet Systems Co., Ltd., Fujisoft Bldg. 3 Kanda-neribeicho,
Chiyoda-ku, Tokyo 101-0022, Japan
e-mail: masumoto@cybernet.co.jp

9.1 Room Acoustics

The auralization of simulation results is one of the most important topics in room acoustics. The techniques for the auralization of the results obtained from scale model experiments or simulations based on the geometrical acoustics have been developed and have already been put into practical use in the field of acoustical design. In these techniques, the impulse responses at target positions in the sound field under consideration are obtained by scale model experiments or by geometrical acoustic simulations, and they are then convolved with a dry source. The resultant signals can be listened to through headphones or loudspeakers. In recent years, outstanding progress in computer hardware has made it possible to auralize the results obtained from wave-based numerical simulations. In the finite-difference time-domain (FDTD) calculations, the impulse response at an arbitrary receiver position can be obtained directly, and it can then be listened by simply converting the digital signals to analog signals.

In the following section, we introduce a multichannel sound field simulation system for auralizing the results of the FDTD method. As applications of the auralization technique, we conducted a subjective comparison on the effects of sound diffusers and we also reproduced a particular fluttering echo in a shrine (the world-famous phenomenon known as the "Roaring Dragon").

9.1.1 Multichannel Reproduction System with FDTD Method

Figure 9.1 shows the outline of a six-channel sound field simulation system that combines three-dimensional numerical analysis using the FDTD calculation, with a six-channel sound reproduction system [1]. This system was originally developed to simulate the sound field of actual concert halls in an anechoic room [2] and the technique has been applied to the auralization of numerical simulation results. In this system, the directional impulse responses in the directions of every 90° at a receiving point in three-dimensional sound field are first calculated by the FDTD method. Next, the calculated impulse response signals are reproduced directly through the six loudspeakers arranged at every 90° in an anechoic room. In this way, the acoustical properties at an arbitrary receiving point set in the virtual space assumed in the FDTD calculation can be reproduced at the center of the sound reproduction system, and one can experience them.

In order to calculate the directional impulse responses for the six orthogonal directions, an arbitrary directivity factor is assumed and set to every 90° by rotating its direction at the receiving point. By multiplying the directivity factor and the instantaneous sound pressure at the receiving point, each of the directional impulse responses at the receiving point is calculated. Two kinds of directivity factor have been proposed, as follows:

Type-I (Cardioid):

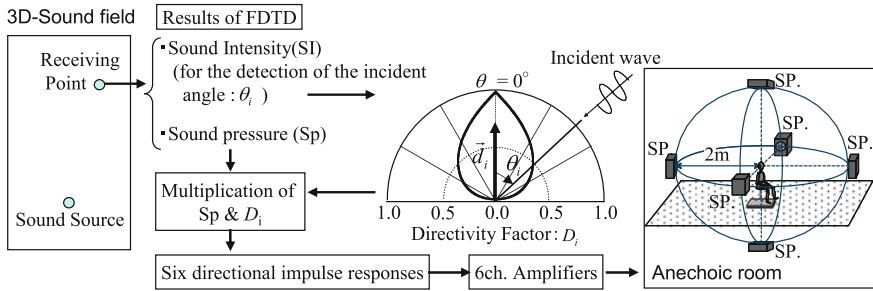$$D_{1,i} = \frac{1 + \cos\theta_i}{2},$$

(9.1)

Fig. 9.1 Outline of a six-channel sound simulation system

Type-II:

$$D_{2,i} = \begin{cases} \frac{\cos\theta_i}{A} & \cos\theta_i \geq 0 \\ 0 & \cos\theta_i < 0 \end{cases}, \tag{9.2}$$

$$A = \frac{1}{2}\sum_{i=1}^{6}|\cos\theta_i|, \tag{9.3}$$

where $D_{1,i}$ or $D_{2,i}$ is the directivity factor ($i = 1$–6), θ_i is the angle between the front direction of each directivity factor and the incident angle of the sound at the receiving point. The incident angle is obtained from the sound intensity vector, which is calculated by the FDTD method as follows:

$$\cos\theta_i = \frac{-d_{ix}\cdot I_x - d_{iy}\cdot I_y - d_{iz}\cdot I_z}{\sqrt{(I_x)^2 + (I_y)^2 + (I_z)^2}}, \tag{9.4}$$

where $\boldsymbol{I} = (I_x, I_y, I_z)$, $\boldsymbol{d}_i = (d_{ix}, d_{iy}, d_{iz})$, $|\boldsymbol{d}_i| = 1$, $\boldsymbol{d}_i$ is the front direction of each directivity factor, and $\boldsymbol{I}$ is the instantaneous sound intensity vector, which is calculated by the FDTD method. By reproducing the resultant six-directional impulse responses, the amplitude of the sound pressure and the direction of the sound intensity vector at the receiving point can be simulated accurately at the center of the reproduction system.

9.1.2 Simulation of Room Impulse Responses

In this section, we consider the auralization of the impulse responses for rooms with different shapes. Figure 9.2 shows the two-dimensional sound fields for the three different room shapes. The directional impulse responses were calculated by
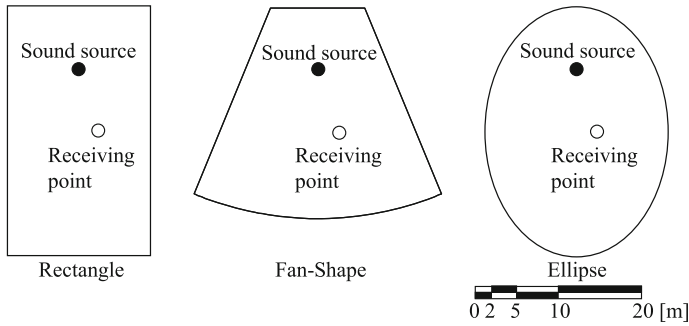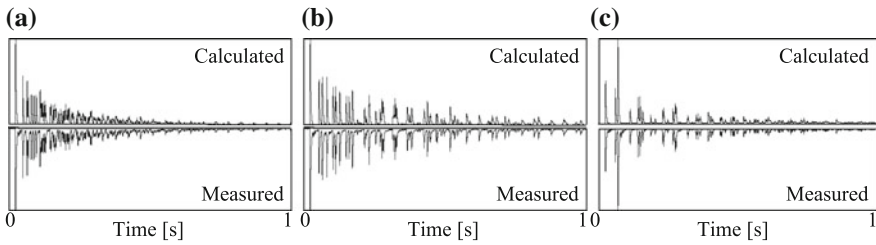
Fig. 9.2 Outline of the sound field

Fig. 9.3 Comparisons of the echo diagrams of the calculations and those of the measurements. **a** Rectangle. **b** Fan-shape. **c** Ellipse

the two-dimensional FDTD method, and the signals were reproduced through four loudspeakers that were set at right angles on an arc of 2 m radius in an anechoic room. In the calculations, the spatial grid size and the time interval were set to be 0.01 m and 0.02 ms, respectively. It was assumed that the normal acoustic impedance of the room boundaries consisted only of the real part, and a constant sound absorption coefficient ($\alpha = 0.2$) was assumed for all boundaries. We used the Type-I directivity factor, defined in Eq. (9.1), to calculate the directional impulse responses.

The calculated directional impulse response signals were reproduced through the four-channel loudspeaker system, and the omnidirectional impulse response was measured at the center of the reproduced sound field. Figure 9.3 compares the echo diagram of the calculation with that of the measurement. These echo diagrams were obtained by passing the omnidirectional impulse response signal through a numerical RMS detector with a 1 ms time constant. As seen in the figure, the calculated results and the measured ones are in very good agreement. The instantaneous sound intensities at the center of the reproduction system were also measured. Figure 9.4 uses radar charts to compare the sound intensity vector of the calculation with that of the measurement. In these figures, it can be seen that in all three cases, the calculated intensity vectors are in fairly good agreement with the measured ones.
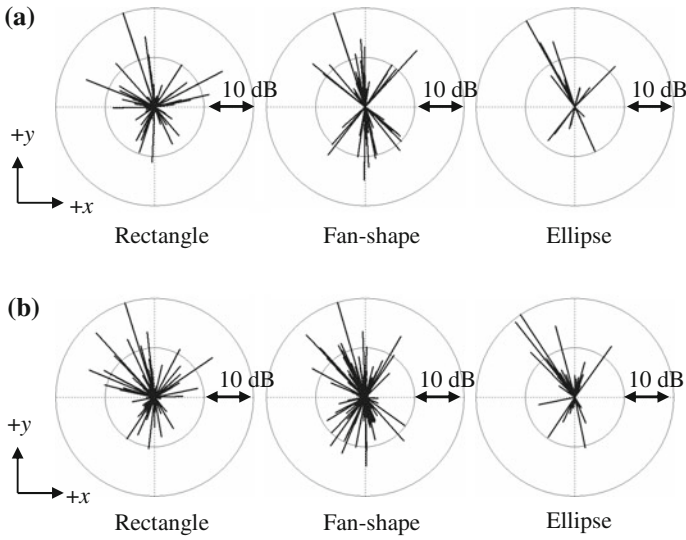
Fig. 9.4 Comparisons between the calculated and measured instantaneous sound intensities. **a** Calculated, **b** measured

9.1.3 Effects of Sound Diffusers in Rooms

As an application to room acoustics of the multichannel sound field simulation method, we conducted an experiment in which we subjectively evaluated the effect of sound diffusers to prevent fluttering echoes. Figure 9.5 shows the rectangular room and two types of sound diffusers attached at the room boundaries (a: Triangular and b: Column). For each type of diffuser, we used four different sets of dimensions, as shown in the figure. For the boundary condition, it was assumed that the normal acoustic impedance at the room boundaries consisted of only the real part, and a constant sound absorption coefficient was assumed for all the boundaries. In the case of a bare room (without diffusers), the absorption coefficient was set to 0.2. For the cases with diffusers, the sound absorption coefficient was assumed to be such that the equivalent sound absorption length (which, in a three-dimensional room, corresponds to sound absorption area) was equal to that of the bare room. In the FDTD calculation, we again used a spatial grid size of 0.01 m and a time step of 0.02 ms. In the experiment, the impulse responses were presented to the subject in a random order through the four-channel loudspeaker system. After hearing each impulse response, the subject judged the strength of the fluttering echo and assigned it to one of five categories as shown in Fig. 9.6. Figure 9.6 shows the results of the subjective experiment. In these figures, each plot shows the arithmetic average of the category number assigned by all of the subjects for each diffuser condition at each receiving point. In the case of the triangular diffusers, it can be seen that the larger the diffusers, the better they prevent the fluttering echoes. On the other hand, in the case of the
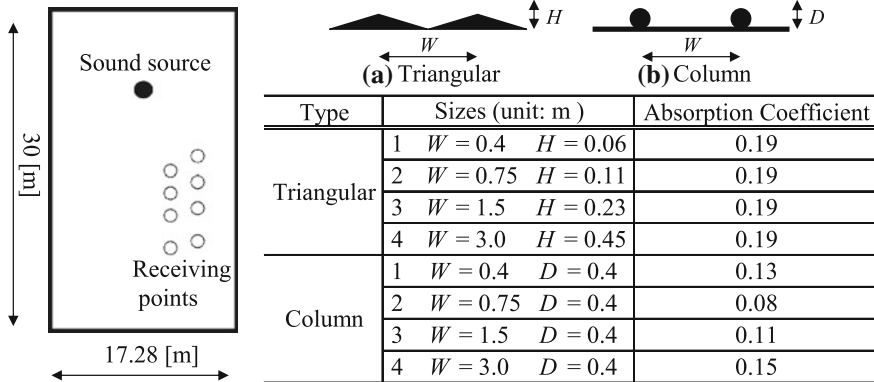
Fig. 9.5 Two-dimensional rectangular room and the two types of sound diffusers that were investigated
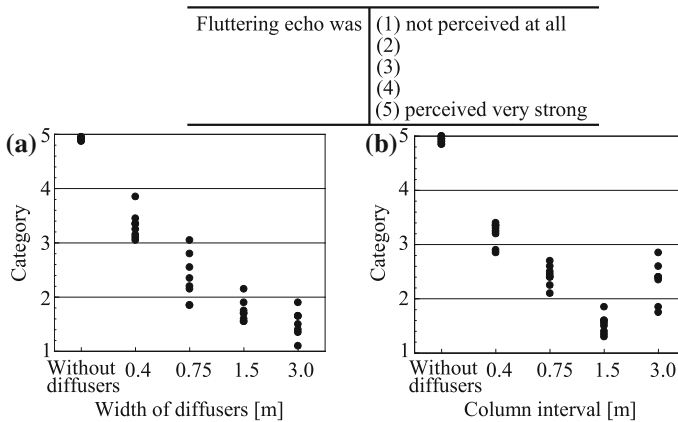

Fig. 9.6 Results of the judgment test on the strength of fluttering echo: **a** triangular diffusers; **b** column diffusers

column diffusers, the effect was greatest when the column interval was 1.5 m. This indicates that an optimum scale exists for the prevention of a fluttering echo by a column diffuser.

9.1.4 Auralization of "Roaring Dragon"

The "Honji-Do" temple located in the "Nikko Toshogu" area of Nikko City, Japan, is famous for a strange acoustic phenomenon called the "Roaring Dragon." A dragon is painted on the ceiling of this building, and when hands are clapped under the head of the dragon, one can hear a strange fluttering echo. In this section, we present a

challenging study that attempts to reproduce the "Roaring Dragon" phenomenon of the "Honji-Do" temple by combining the FDTD with the multichannel reproduction technique.

9.1.4.1 "Roaring Dragon" Phenomenon in "Honji-Do" Temple

The "Roaring Dragon" phenomenon in the "Honji-Do" temple is a fluttering echo that is caused by repeated reflections between the ceiling, which has very little curvature, and the flat floor. The temple was unfortunately destroyed by an accidental fire in 1961 and was rebuilt in 1969. In the reconstruction work, reproduction of the "Roaring Dragon" phenomenon was one of the most important items, and a one-fourth scale acoustic model experiment was conducted to study the cause of the acoustic phenomenon and to determine how it could be reconstructed. The following is a summary of the study:

1. The duration time of the fluttering echo becomes longer with increasing curvature of the ceiling.
2. The arch rise (the difference in height between the center of the ceiling and the edge) that best reproduces the "Roaring Dragon" phenomenon is 9 cm.
3. When hands are clapped under the head of the dragon, just to the side of directly below the center of the ceiling, the "Roaling Dragon" pulsates.

9.1.4.2 Three-Dimensional FDTD Simulation of "Roaring Dragon" Phenomenon

Three-dimensional FDTD simulation was conducted by modeling the room sound field in the "Honji-Do" temple. Figure 9.7 shows the plan of the "Honji-Do" temple. The source was positioned just below the painted dragon at a height of 1.2 m. The receiving point was positioned just above the source position at a height of 1.5 m. As the arch rise of the ceiling (the difference in height between the center of the ceiling and the edge) was assumed to be 9 cm. The spatial grid size and the time interval were set to be 0.02 m and 0.032 ms, respectively. The Type-II directivity factor, defined in Eq. (9.2), was selected for calculating the directional impulse responses.

9.1.4.3 Reproduction of "Roaring Dragon" Phenomenon by Six-Channel Sound Field Simulation

Figure 9.8 shows the system used to reproduce the "Roaring Dragon", based on a three-dimensional FDTD calculation and the six-channel sound field simulation technique [3]. The basic concept of this system is the same as that of the multi-channel sound field simulation system described above. In this system, the sound of a handclap made at the center of the system was convolved in real time with the directional impulse responses at the receiving point, which was set in the sound
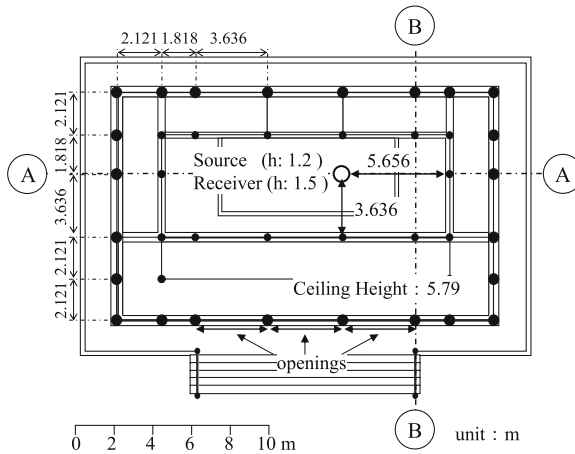
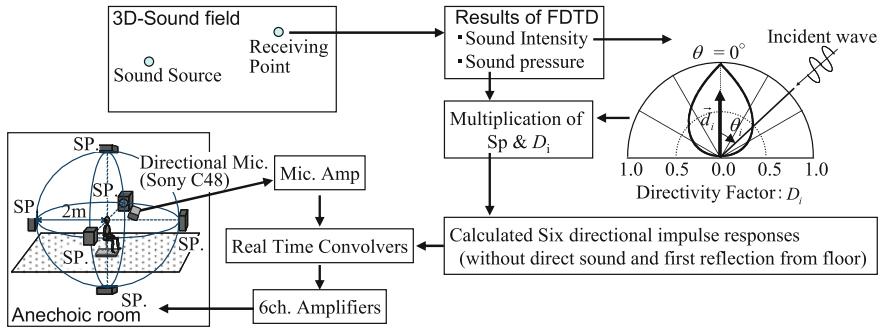**Fig. 9.7** Plan of the "Honji-Do" temple

Fig. 9.8 Outline of the six-channel sound field simulation for reproduction of the "Roaring Dragon"

field of the model of the "Honji-Do" temple. The directional impulse responses were obtained by the FDTD calculation, and the resultant signals were reproduced through the loudspeakers set in an anechoic room. This system was originally developed to simulate the sound field on the stage of actual concert halls in an anechoic room to investigate the acoustic property for music players [4]. The technique has also been applied to the auralization of numerical simulation results. Figure 9.9 shows the directional impulse responses obtained by the FDTD calculations. In the figures, the front direction corresponds to the upward direction in Fig. 9.7. It can be seen that the fluttering echo persists for a long time in the directional impulse responses of Up and Down directions. This indicates that there were repeated strong reflections between the ceiling and the floor. Those who had visited "Honji-Do" temple and experienced the real "Roaring Dragon" within the previous year reported that the simulated fluttering echo could be perceived above their heads and that it pulsated as it decayed. Some subjects also commented that they perceived the duration time of
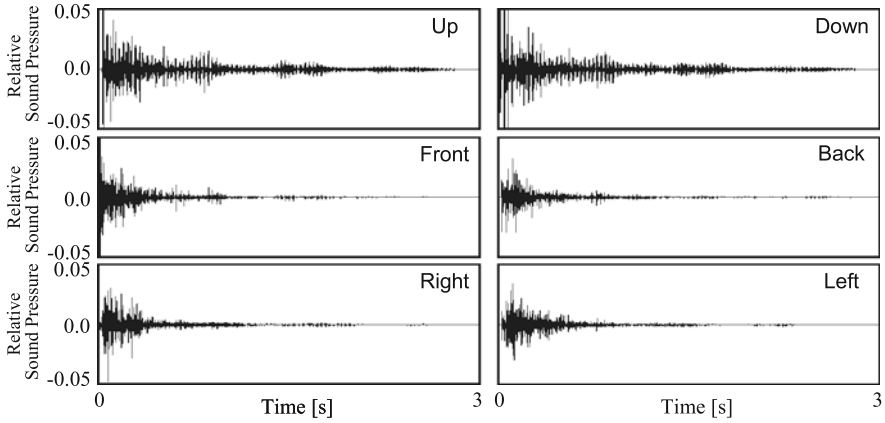
Fig. 9.9 Directional impulse responses at the receiving point (the sound pressure is normalized by the amplitude of the direct sound)

the simulated echo to be longer than that of the real "Roaring Dragon". One reason for this might be that there are many objects, such as ritual articles, which reflect or absorb the sound in the real "Honji-Do" temple. It may also be due to the low level of background noise in the anechoic room.

9.2 Noise Propagation

The frequency and time-transient characteristics of the leak sound transmitted into the residential buildings are strongly influenced by the sound insulation characteristics of the building façade. In this section, an auralization technique [5] in which the sound insulation characteristics of the façade can be realized through numerical simulation is described.
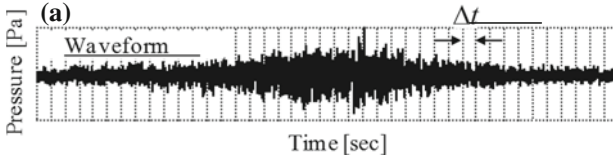
9.2.1 Auralization Method

Detailed scheme of the proposed auralization system shown in Fig. 9.10 is described in this section.
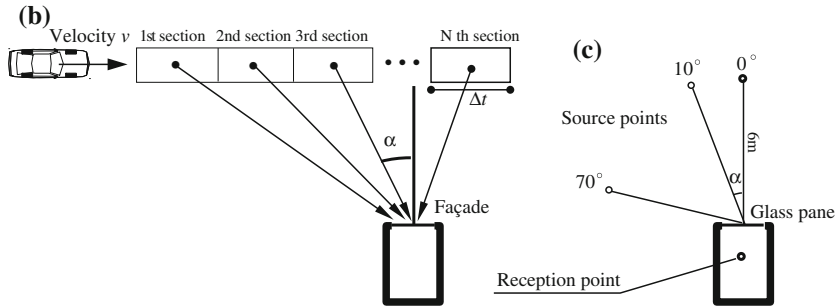
9.2.1.1 Recording of Vehicle Noise

A waveform of a pass-by noise of a vehicle was recorded by an omnidirectional microphone which was set at a point 7.5 m away from the running lane (waveform (a), shown in Fig. 9.10), and the waveform was divided into N sections with every
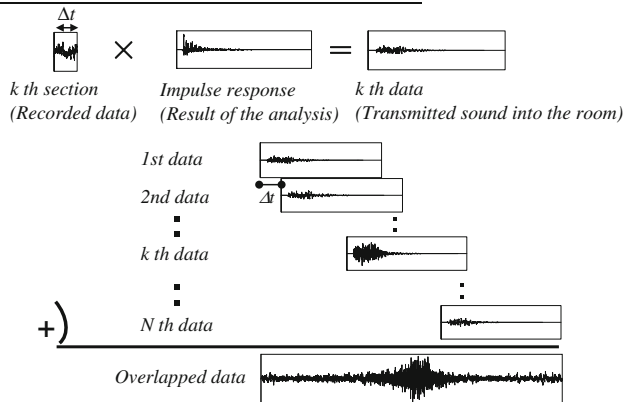
Step 1 Recording of the vehicle noise

(a)

Pressure [Pa]

Waveform

Δt

Time [sec]

Step 2 Simulation of the sound insulation characteristics of the building facade

(b)

Velocity v 1st section   2nd section   3rd section   N th section

Δt

α

Façade

(c)

10°   0°

Source points

70°

6m

α

Glass pane

Reception point

Step 3 Synthesis of the vehicle noise transmitted to a room

Δt

×   =

k th section
(Recorded data)

Impulse response
(Result of the analysis)

k th data
(Transmitted sound into the room)

1st data

2nd data Δt

k th data

+)   N th data

Overlapped data

Step 4 Simulation of the traffic flow (optional process)
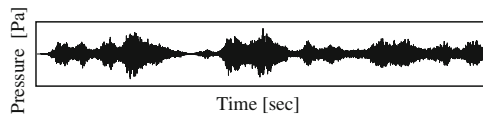
Pressure [Pa]

Time [sec]

Fig. 9.10 Flow chart of the auralization scheme. **a** Waveform, **b** incident angle of the vehicle noise into the room, **c** arrangement of the numerical analysis (sound transmission through façade)

time interval, Δt. Signal processing on these divided N sections was carried out in the following steps.

9.2.1.2 Simulation of Sound Insulation Characteristics of Building Façade

As shown in Step 2 of Fig. 9.10, the incident angle, α, of the sound which propagates from the running vehicle to the building façade, is varying every moment and the sound insulation characteristics of building façade, especially glass pane, also vary in correspondence with the angle of sound incidence.

In order to simulate the directional characteristics of sound insulation of the building façade, vibroacoustic numerical analysis using FDTD method was applied. The plan of the sound field for the three-dimensional analysis is shown in Fig. 9.10c. Each of the three-dimensional sound field was analyzed, and the impulse response at the receiving point was calculated.

9.2.1.3 Synthesis of Vehicle Noise Transmitted into a Room

The incident angle, α, of the vehicle noise to the façade was estimated by the geometrical relationship between the positions of the vehicle and the building façade. The convolution of the kth vehicle noise whose incident angle to the building façade is α and the impulse response of the same incident angle, α, obtained by numerical analysis was performed for the data of N sections. The transmitted noise into the room was made by overlapping the convoluted data by shifting every data with interval, Δt as shown in Step 3 of Fig. 9.10.

9.2.1.4 Simulation of Traffic Flow

In order to simulate a road traffic noise with an arbitrary traffic volume, a pass-by noises of multiple automobiles were overlapped. An example of the simulated road traffic noise with a traffic volume, 1,500 vehicles/h, is shown in Step 4 of Fig. 9.10.

9.2.2 Simulation Results

In this study, a simulation of the vehicle noise at the point of reception in indoor spaces was performed using the proposed method. The assumed condition is shown in Fig. 9.11. The geometrical relationship between the running lane of the vehicle and the points of reception and the details of the reception room are shown in Fig. 9.11. It was assumed that a glass plate with dimensions of 1.8 m (W) × 1.8 m (H) is set in the opening of a room whose dimensions are 2.7 m (W) × 2.2 m (H) × 3.6 m (D), as shown in the figure.
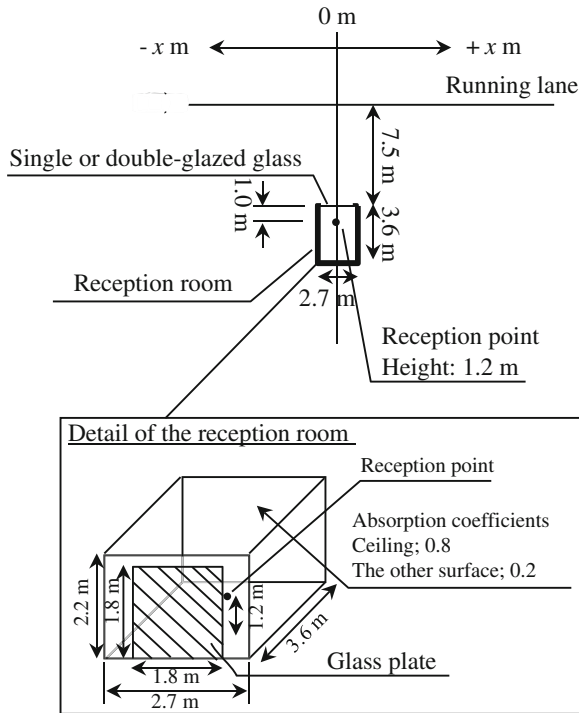
Fig. 9.11 Investigated conditions in the case study

The values of the parameters for the physical properties of the glass plate were set as follows: density, $2{,}500\,\mathrm{kg/m^3}$; Young's modulus, $7.16 \times 10^{10}\,\mathrm{N/m^2}$; and Poisson's ratio, 0.22. In addition, an elastically supported condition as described in Sect. 8.3.2 was applied in the vibration analysis. The absorption coefficient of the ceiling was set at 0.8, assuming a ceiling with absorption treatment, and that of the surfaces of the other walls were set at 0.2.

The investigated conditions are as follows: Case 0 (outdoor space), Case 1 (inside a room with a single glass plate with a thickness of 6 mm in the opening), Case 2 (inside a room with single glass plate of thickness 10 mm), and Case 3 (inside a room with double-glazed glass composed of two glass plates, each of thickness of 6 mm, separated by a 6-mm-thick layer of air).

Simulation results of the time-transient characteristics of sound pressure levels at the reception point caused by one pass-by vehicle is shown in each one-third Octave band in Fig. 9.12. In this graph, the timing at which the running vehicle reaches the 0 m position in Fig. 9.11 is set at 0 s. The sound pressure levels are calculated so that the sound exposure level of the pass-by sound is 75 dB. The speed of the vehicle is assumed to be 60 km/h, and the horizontal axis of the figure describes a relative time from the timing of the vehicle passing in front of the room.
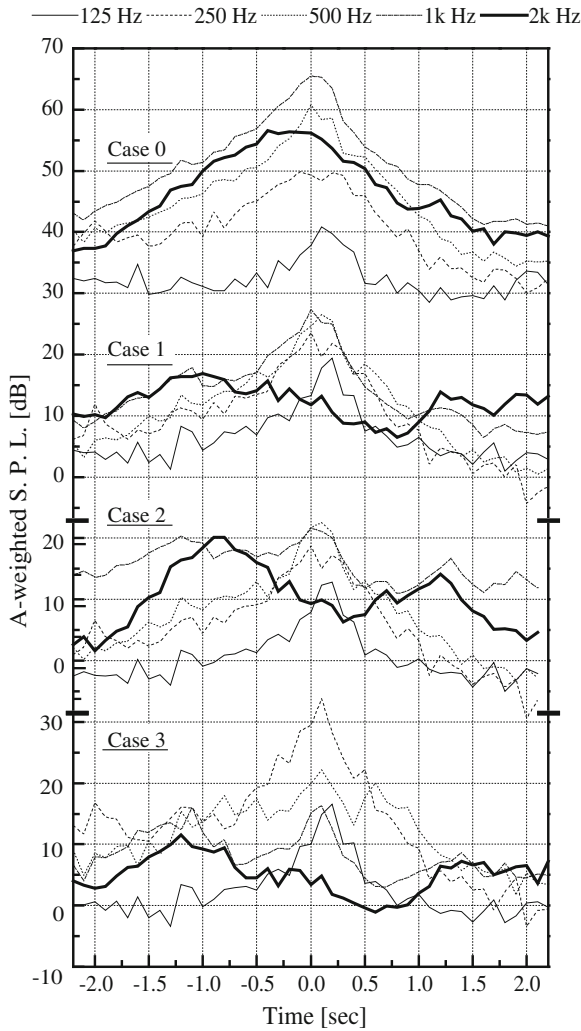
—— 125 Hz ········ 250 Hz ········ 500 Hz ------- 1k Hz —— 2k Hz

Fig. 9.12 Time-transient characteristics of the pass-by sound in 1/3. Oct. band frequency

In Case 0, the sound pressure level decreases as the position of the running vehicle moves further from the 0 m position in all frequency bands. However, in Case 1, the time-variant characteristics for 2 kHz have peaks at −1.2 s and +1.2 s, and the sound pressure level at 0 m is less than those at the peak positions. It is considered that a large quantity of sound energy is transmitted through the glass plate when the incident angle of sound to the plate is larger, in the 2 kHz band, which includes the coincidence cut-off frequency of a glass plate with a thickness of 6 mm.

In Case 2, the time-variant characteristics for 1 and 2 kHz have the same peak characteristics as those for 2 kHz in Case 1. It is also considered that this result is due

Fig. 9.13 Frequency characteristics of the transmitted sound into the room
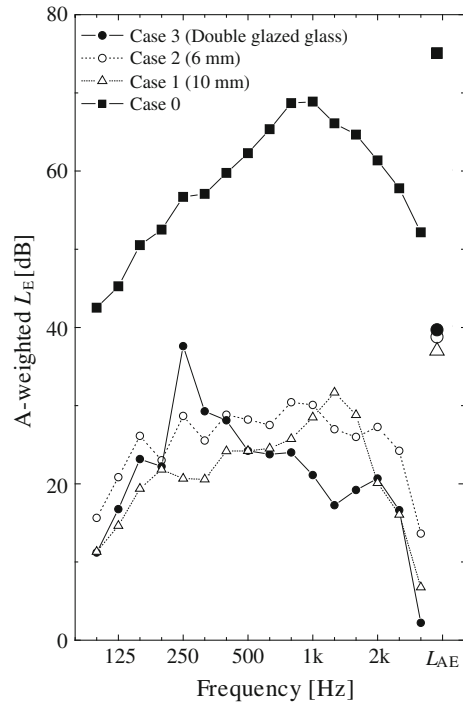
to the coincidence phenomenon. In Case 3, the time-variant characteristics in 250 Hz have a large value, especially at around 0 s. The reason for this is that a double-glazed glass plate has a resonant frequency in the low frequency range, which is caused by the resonant phenomenon that characterizes its composition of mass (glass)-spring (air layer)-mass (glass).

Based on the obtained time-transient characteristics, the single event sound exposure levels, L_E, of the pass-by sounds are calculated, and the results are shown in Fig. 9.13. Comparing the conditions of Case 1 and 2, the sound energy level of Case 2 in 250 Hz has larger value than that of Case 1, and it is caused by the mass-spring-mass effect as described above.

9.3 Head-Related Transfer Functions

A wide variety of research fields are discussing the need to reproduce stereophonic sound fields. To do this, it is necessary to obtain head-related transfer functions (HRTFs), which are acoustic transfer functions between sound sources located around the human head and ear. Based on this, several measurements have been performed [6]. However, when the physical load of the subject is taken into account, it is extremely difficult to conduct the necessary procedures for obtaining HRTFs.

Attempts have therefore been made to obtain HRTFs by numerical analysis, resulting in a number of published reports.

Yet the numerical analysis also has some drawbacks. Huge calculation costs are required since data for the entire audible frequency range are required for HRTFs. For example, if analysis at 20 kHz is conducted using the boundary element method (BEM), the number of degrees of freedom (DOF) to represent a human head is around 100,000, and the required memory will exceed 160 GB. For this reason, upper limit frequencies have been limited to several kHz if the entire head is taken into account, or several assumptions have been made to conduct analysis up to 20 kHz (e.g., pinna is connected to the infinite baffle plane) [7].

Incidentally, in recent years, research has also been conducted on the implementation of sound field analysis using fast multipole algorithms with the three-dimensional BEM (see Sect. 4.3). Using this fast multipole BEM (FMBEM) it has been possible to calculate HRTFs for the entire audible frequency range within feasible memory sizes and calculation times [8]. This study outlines the results of these calculations.

9.3.1 Basic Examination

9.3.1.1 Checking Uniqueness of Solution

When analyzing the external field using the BEM, the unique solution cannot be obtained at the eigenfrequencies of an internal field whose boundary is the same as the target geometry. To avoid this, the proper formulation should be selected by comparing theoretical solutions with the solution from the BEM. Checked formulations are basic form (BF), normal derivative form (NDF) (see Sect. 4.1.1 for these formulations), Burton–Miller form (BMF) and dual form (DF) (see Sect. 4.2.1 for these formulations). A sphere 0.25 m in diameter is used as a simplified shape of a human head (Fig. 9.14). The angle θ is introduced to represent locations on the sphere, starting from 0° to 180°. The piston oscillation of 1 mm/s is defined in the range from $\theta = 0°$ to $\theta = 20°$. All other parts of the surface are taken to be rigid. In this case, the theoretical solution can be calculated with Eq. (9.5) [9]:

$$p(r,\theta) = \frac{j\rho c V_0}{2} \sum_{n=0}^{N} (P_{n-1}\cos\alpha - P_{n+1}\cos\alpha) \frac{h_n^{(1)}(kr)}{h_n^{(1)'}(ka)} P_n(\cos\theta), \quad (9.5)$$

where j is the imaginary unit, ρ is the medium density, c is the sound speed, k is the wave number, V_0 is the vibration velocity, a is the sphere's radius, r is the distance between the sphere's center and the field point ($r > a$), P_n are the Legendre polynomials, $h_n^{(1)}$ are the spherical Hankel functions of the first kind, and $h_n^{(1)'}(x) = h_{n-1}^{(1)}(x) - ((n+1)/x)h_n^{(1)}(x)$. In this study, $\rho = 1.225\,\text{kg/m}^3$, $c = 340\,\text{m/s}$, $V_0 = 0.001\,\text{m/s}$, $a = 0.125\,\text{m}$, $\alpha = 20°$ are used and response at 20 kHz is calculated.

Fig. 9.14 Model used for
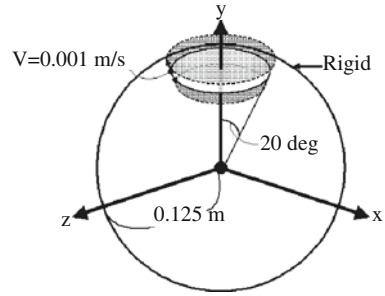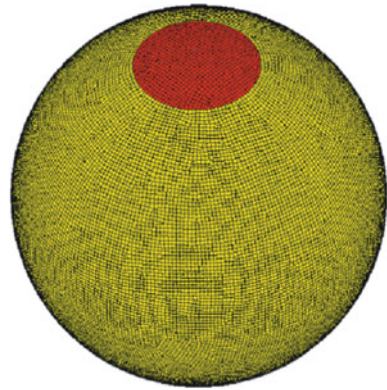checking uniqueness of the
solution

Fig. 9.15 Boundary element
mesh (num. of elements:
31,200)

The mesh used for the boundary element model is shown in Fig. 9.15. The sphere
is discretized with elements whose size is about one-sixth of a wavelength, so analy-
sis can be performed up to 20 kHz and the number of elements is 31,200. The velocity
B.C. ($V_0 = 0.001$) is defined as the part indicated with red in the figure. Only in
the DF case, to avoid the fictitious eigenfrequency problem, the specific impedance
of air is defined as impedance B.C.s to the negative side of all the elements. The
ILUT(1^{-6},100) [10] was used as a preconditioning technique. The generalized min-
imal residual (GMRes) method was adopted as an iterative solver with the restart
number being set to 2,528 for BF, NDF, and BMF cases, and 325 for the DF case, so
3 GB of memory is required. The machine used in this work is an IBM IntelliStation
with the following specifications; CPU: AMD OpteronTM 2.79 GHz; OS: Windows
XP Professional ×64 edition; memory: 8 GB.

Figure 9.16 shows the sound pressure distributions on the sphere as a function of
the θ at 20 kHz. The results obtained by BF or NDF differ greatly from the theoretical
solution, which is caused by the deteriorated uniqueness of the solution. With BMF,
this problem does not occur, but analysis precision is poor, especially in the area
of $\theta > 90°$. With DF, small differences from the theoretical solution were seen in
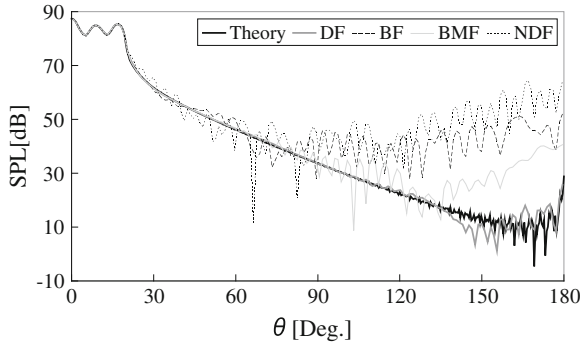the area of $\theta > 140°$, so we regard it as sufficiently accurate. Table 9.1 shows the

Fig. 9.16 Distribution of sound pressure on sphere

Table 9.1 Calculation cost

Formulation	BF	NDF	BMF	DF
Required memory (GB)	1.64	1.64	1.64	2.68
CPU time (min)	51.8	132.4	1432.7	23.6
Number of matrix-vector multiplications	65	166	1836	18

amount of memory, CPU time, and number of iterations required for analysis. A larger amount of memory is required since the number of DOFs of the DF is twice that of other formulations. However, due to the smaller number of iterations, the analysis time is the shortest. DF is used for HRTF calculation because of its accuracy and short calculation time.

9.3.1.2 Checking Reciprocity

The reciprocity between the point sound sources and the field points is introduced to reduce the number of calculation cases. However, in the case of HRTF calculation using reciprocity, the distance between the point sound source and the boundary elements tends to be short, so these analysis precision goes down. Caution should therefore be paid to this point.

This study uses the same geometry as in the previous study. The field point position is defined 2 mm outside in the $\theta = 0°$ direction. The point sound source is defined at 1 m from the spherical center in the $\theta = 90°$ direction (Fig. 9.17). In analysis using reciprocity, the source and field point position are set reversed. The boundary element model is checked in two cases. Case 1 is where the surface is discretized uniformly with an element size enabling analysis of each frequency. Case 2 is where the portion of the surface that is assumed to be the range of the pinna (in this case $\theta < 20°$) is discretized with 1 mm (half the distance between the point sound source and the sphere) elements, the lower hemisphere ($90° < \theta < 180°$) is discretized
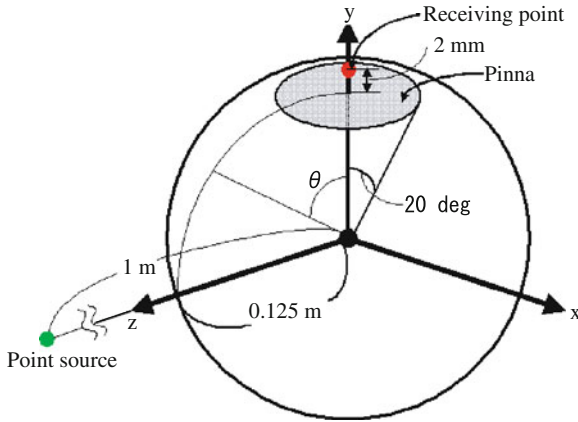
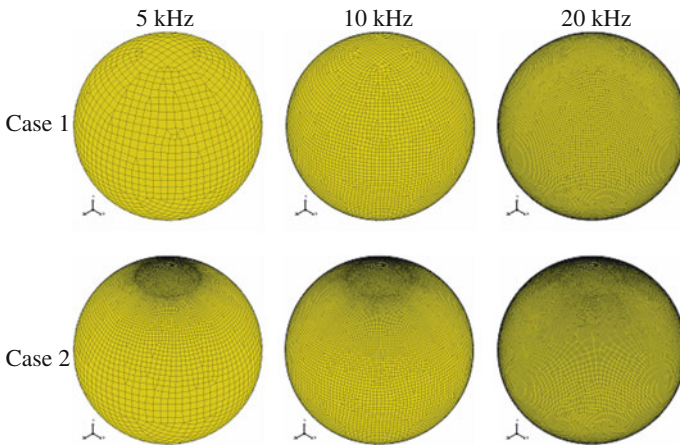**Fig. 9.17** Model used for reciprocity check

Fig. 9.18 Boundary element meshes used in checks of reciprocity

with the same condition as in Case 1 and the other part discretized to maintain the continuity of the elements. Figure 9.18 shows the meshes used in this check.

Table 9.2 shows sound pressure amplitude at the field point at 5, 10, and 20 kHz. The value indicated as an error is $(p_r - p_n)/p_r$. p_n is the pressure obtained by the model that does not consider reciprocity, and p_r is the pressure obtained by the model that does consider reciprocity. The reciprocity does not hold because errors in Case 1 are 24 %.

Table 9.2 Result of reciprocity checking

| Frequency (kHz) | Case | Pressure in amplitude (Pa) Taking reciprocity into considerration | | Relative error (%) |
		No	Yes	
5	Case 1	1.160	0.874	24.66
	Case 2	1.180	1.169	0.932
10	Case 1	1.161	1.057	8.958
	Case 2	1.167	1.161	0.514
20	Case 1	1.139	1.127	1.053
	Case 2	1.138	1.138	0.000

Fig. 9.19 Picture obtained by MRI

9.3.2 Simulation Results

9.3.2.1 Generation of Boundary Element Mesh

Individual geometry data is required to calculate the individual HRTF. To this end, the geometry used in the analysis is created from image data using magnetic resonance imaging (MRI). Using this geometry (Fig. 9.19), the surface is discretized.

By MRI, 108 pictures are taken with 1 mm clearance in the sagittal direction (parallel to the body axis z and separating the body into right and left). The format of these pictures in Digital Imaging and Communications in Medicine (DICOM), used as a format for medical pictures such as computed tomography (CT) and MRI, and those are read by a special program. Using this, the image information inside the eardrum is removed. The new image information is extracted to Standard Triangulated Language (STL), which is a data format to represent three-dimensional geometry as a cluster of small triangles. Finally, these data are used for the boundary element mesh. The generated boundary element mesh is shown in Fig. 9.20. Because the maximum element length is 2 mm (about one-sixth of a wavelength at 27 kHz when the sound speed is 340 m/s), the number of elements is 186,380.
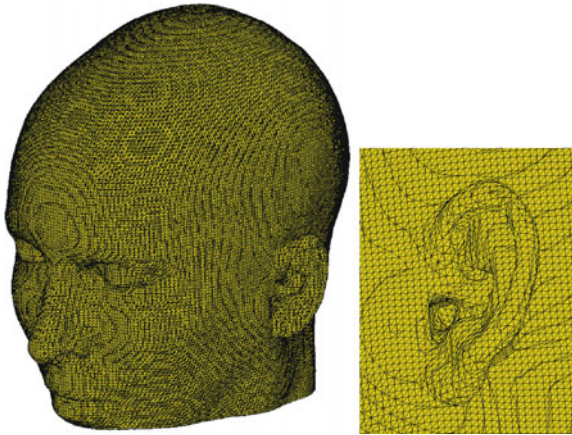
Fig. 9.20 Boundary element mesh used for HRTF calculation (num. of elements: 186,380)

9.3.2.2 Settings for Solver

When DF is applied for the calculation using the boundary element mesh obtained by
the described procedure, more than 1 TB of memory is required if conventional BEM
is employed. FMBEM is thereby employed. Moreover, in order to obtain HRTF for
full audible frequency range with a single boundary element mesh, both FMBEM for
low frequency (LF-FMBEM) and high frequency (HF-FMBEM) (see Sect. 4.3) are
used. Detailed settings for LF-FMBEM and HF-FMBEM can be found in [11, 12].
The deepest cell level in the hierarchical cell structure L and the appropriate solver
are selected so as to avoid exceeding the maximum memory requirement (in this case
16 GB). In this study, about 11.0 GB is required by LF-FMBEM below 6.2 kHz, and
about 13.6 GB is required by HF-FMBEM for higher frequencies.

9.3.2.3 Approximation Using Rational Function

An approximated HRTF is evaluated from the response at the limited frequency by an
interpolation with the aid of a rational function. The response at other frequencies can
be obtained by directly evaluating this function. The calculation time was reduced
because the approximated HRTF can be obtained from the limited frequency. The
function is approximated using the following rational function:

$$f(x) = \frac{ax + b}{1 + cx}.$$

(9.6)

Coefficients a, b and c are calculated by the responses of the BEM model at three
frequency steps: $x = x_0$ the center frequency, $x = x_{+1} = x_0 + \Delta x$ (where Δx
is the difference in frequency between x_0 and the next calculated frequency to the

center frequency), and $x = x_{-1} = x_0 - \Delta x$. The frequency x_0 is selected through the following procedure.

1. Two frequencies (the next to lowest frequency [x_L] among the total frequency range and the frequency next to the highest one [x_H]) are selected as x_0.
2. For each x_0, calculate the response at three frequencies ($x_0 - \Delta x$, x_0, $x_0 + \Delta x$). Then with coefficients a, b, and c, the approximate functions Eq. (9.6) are defined. Those are called $f_{n,L}(x)$ and $f_{n,H}(x)$ later.
3. To define the switching frequency between $f_{n,L}(x)$ which is the approximate function for the lower frequency and $f_{n,H}(x)$, which is the approximate function for the higher frequency, explore the frequency x_k where the error defined below shows the minimum between x_L and x_H:

$$E = \frac{1}{p} \sum_{n=1}^{p} \left(\min \left[\frac{f_{n,H}(x_k) - f_{n,L}(x_k)}{f_{n,L}(x_k)}, \frac{f_{n,H}(x_k) - f_{n,L}(x_k)}{f_{n,H}(x_k)} \right] \right), \quad (9.7)$$

where p is the number of field points.
4. If the error at x_k is lower than the specified value, $f_{n,L}(x)$ is selected as the approximation function below x_k and $f_{n,H}(x)$ is selected as the approximation function above x_k. In the case of greater than the specified value, first the center frequency x_M between x_L and x_H is calculated, next x_L and x_M are again set as a new variable group of x_L and x_H, and finally go back to 1. Similarly, x_M and x_H are set again as a new variable group and go to 1.

The approximation function for all frequency ranges is obtained by repeatedly executing the above procedure.

9.3.2.4 Other Settings

The actual calculation was conducted taking into account the reciprocity between the evaluation point and the sound source locations. Therefore the evaluation point and the source locations were switched, enabling the response calculation to be performed all at once.

The sound source is located 2 mm from the left eardrum (Fig. 9.22), and three field points are located as indicated by Fig. 9.21 in the horizontal plane, which includes the ear canal entrance. The sound speed is 340 m/s and medium density is 1.225 kg/m^3. Response is calculated by 50 Hz up to 22.05 kHz. The error indicated in Eq. (9.7) is 1.0. The Generalized Product Bi-Conjugate Gradient (GPBiCG) was adopted as an iterative solver. The machine used is a Dell Precision 690 with the following specifications; CPU: Intel R XeonR, X5355 processor, 2.66 GHz; OS: Windows XP Professional x64 edition; memory: 32.0 GB. The other settings are the same as in Sect. 9.3.1.
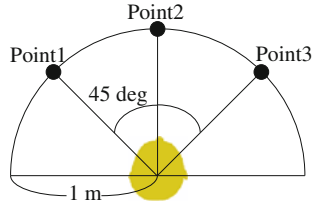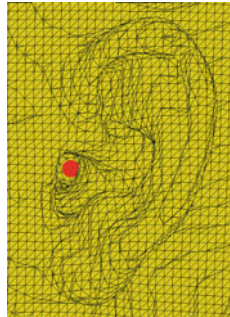
Fig. 9.21 Field point locations

Fig. 9.22 Point source location (indicated by *circle* near entrance of ear canal)

— Point1 — Point2 — Point3
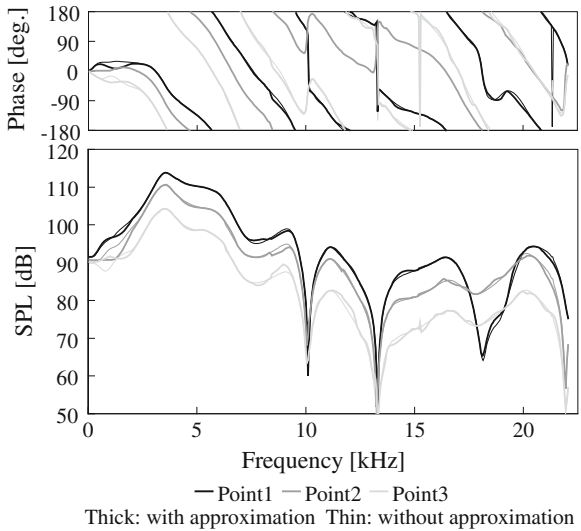Thick: with approximation Thin: without approximation

Fig. 9.23 HRTF: (*Top*) phase shift (value minus the change within distance between sound source and evaluation points), (*bottom*) sound pressure level (normalization of sound pressure at a point 1 m from sound source)
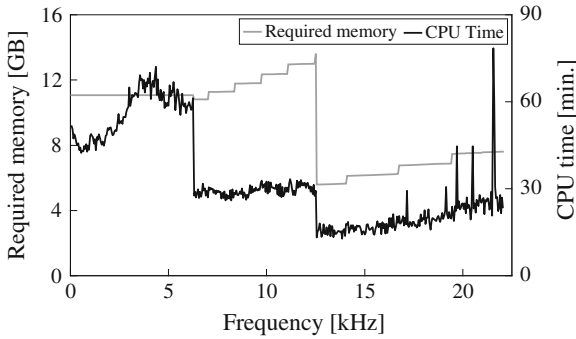
Fig. 9.24 Required amount of memory and calculation time

9.3.2.5 Results

The calculated HRTFs are shown in Fig. 9.23. In this figure, both HRTFs with and without rational function approximation are shown.

Figure 9.24 shows the required amount of memory and the calculation time of the case without rational function approximation. In this case it takes about 10.1 days to acquire the HRTFs for all frequency range, whereas by the case of using rational function approximation the total calculation time is reduced to 1.6 days because the number of calculation steps is reduced from 441 to 72.

References

1. T. Yokota, S. Sakamoto, H. Tachibana, Sound field simulation method by combining finite difference time domain calculation and multi-channel reproduction technique. Acoust. Sci. Tech. **25**, 15–23 (2004)
2. S. Yokoyama, K. Ueno, S. Sakamoto, H. Tachibana, 6-channel recording/reproduction system for 3-dimensional auralization of sound fields. Acoust. Sci. Tech. **23**, 97–102 (2002)
3. T. Yokota, S. Sakamoto, H. Tachibana, K. Ishii, Numerical analysis and auralization of the "ROARING DRAGON" phenomenon by the FDTD method. J. Environ. Eng. (Trans. AIJ) **629**, 849–856 (2008)
4. K. Ueno, H. Tachibana, Experimental study on the evaluation of stage acoustics by musicians using a 6-channel sound simulation system. Acoust. Sci. Tech. **24**, 130–138 (2003)
5. T. Asakura, T. Miyajima, S. Sakamoto, Prediction method for sound from passing vehicle transmitted through building facade. Appl. Acoust. **74**(5), 758–769 (2013)
6. B. Gardner, K. Martin. HRTF measurements of a KEMAR dummy-head microphone. Technical report, MIT Media Lab Perceptual Computing, 1994
7. Y. Kahana, P.A. Nelson, Boundary element simulations of the transfer function of human heads and baffled pinnae using accurate geometric models. J. Sound Vib. **300**, 552–579 (2007)
8. T. Masumoto, T. Oshima, Y. Yasuda, T. Sakuma, M. Kabuto, M. Akiyama. HRTF calculation in the full audible frequency range using FMBEM, in *Proceedings of the 19th International Congress on Acoustics*, Madrid, number COM-06-012, 2007 (In CD-ROM)
9. E.G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Chap.6.7 (Academic Press,1999)

10. Y. Saad, ILUT: a dual threshold incomplete LU factorization. Numer. Linear Algebra **1**, 387–402 (1994)
11. Y. Yasuda, T. Oshima, T. Sakuma, A. Gunawan, T. Masumoto, Fast multipole boundary element method for low-frequency acoustic problems based on a variety of formulations. J. Comput. Acoust. **18**, 363–395 (2010)
12. Y. Yasuda, T. Sakuma, Fast multipole boundary element method for large-scale steady-state sound field analysis. Part II: examination of numerical items. Acta Acust. United Ac. **89**(4), 28–38 (2003)