Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm

Michael Vorländer
Institut für Technische Akustik, Rheinisch-Westfälische Technische Hochschule Aachen, Templergraben 55, D-5100 Aachen, Federal Republic of Germany

(Received 12 August 1988; accepted for publication 30 January 1989)

A new method for the calculation of room acoustical impulse responses is described, which is based on two well-known computer algorithms, the ray-tracing and the image-source models. With the new method, the procedure of sieving the "visible" image sources out of the enormous quantity of possible sources is carried out by examination of the histories of sound particles. From the obtained list of visible image sources, the impulse response of the enclosure is easily constructed. The new method combines the advantages of the ray-tracing process, namely, the relatively slow increase of computation time with the length of the impulse response, with the accuracy inherent to the image-source model, which is even sufficient to calculate the Fourier transform, i.e., the steady-state transmission function of the room, or to convolve the impulse response with sound signals.

PACS numbers: 43.55.Ka, 43.55.Fw

INTRODUCTION

Nowadays, the exact investigation of the acoustics of an auditorium requires the calculation or measurement of room impulse responses. If they are known, several room acoustical parameters can be evaluated from them, which are correlated to the subjective effects of hearing in rooms.[1] Furthermore, it is often desired to get an impression on how a room will "sound" prior to its completion. In principle, this is possible by convolving sound signals with the impulse responses, provided they have previously been calculated by computer simulation or measured in scale models.

So far, two different algorithms for the calculation of transient sound propagation in rooms on the computer are known. The first one is based on image sources, whereas the second one makes use of sound particles, which are numerically followed up (ray tracing).

Recently, the image source algorithm has often been described in literature, for instance, by Borish,[1] Kirszenstein,[2] and Lee and Lee.[3] All of these authors truncate the length of their impulse responses relatively early because the required computation time increases exponentially with length. The achieved time resolution, however, is impressively high, since the energies and the arrival times of the reflections are obtained with the accuracy of floating point variables. This is important if we intend to perform convolution of impulse responses with reverberation-free signals, since the time resolution must at least correspond to the sampling rate of the signal.

On the other hand, the ray-tracing method[4–6] has also its specific advantages. The computation time increases only proportionally with the length of the impulse response, but the achieved temporal resolution is limited. A further important difference is that ray tracing can take scattering of sound at rough surfaces into account.

In the next sections, the two algorithms are briefly described. Then, a new method will be introduced, which in some ways represents a combination of the ray-tracing and image-source models. It allows the computation of high-resolution impulse responses of great length and thus enables one to carry out Fourier transforms to examine the frequency domain of a room response. In the last section, exactly computed steady-state transfer functions will be discussed.

I. THE CONVENTIONAL IMAGE-SOURCE ALGORITHM

The image-source algorithm is based on geometrical acoustics. If the enclosure consists of plane walls with specular sound reflections, the rays can be traced back from a receiver point to the sound source. For this purpose, virtual image sources are used that, of course, must correspond to the underlying room shape. We start by constructing the first-order mirror images with respect to all walls and proceed in the same way with the generated secondary sources leaving apart only that wall at which the secondary source was mirrored last. This process has to be continued up to a prescribed order of the image sources. Thus an image source of fifth order represents a sound ray that has suffered five wall reflections. A simple way to manage the sources is counting them by the indices of the walls that were involved in the algorithm of mirroring. The number of indices is equal to the order of the image.

Next, we must examine, for each receiver point, which of the obtained image sources are "visible" and which are not, since a meaningful ray path can be assigned only to a visible source. Each source is the last link of a chain of successive image sources and, with its indices, one can read which walls were hit by the associated ray path. A chain related to a source of ith order is

$$S \rightarrow S_{n_1} \rightarrow S_{n_1 n_2} \rightarrow \cdots \rightarrow S_{n_1 n_2 \cdots n_{i-1}} \rightarrow S_{n_1 n_2 \cdots n_i},$$

where $n_k \neq n_{k\pm1}$ can assume values from 1 to the total number of walls. In Fig. 1, the visibility test is illustrated for an image source of second order. The actual receiving point R is connected with the image source (S_{12}), the visibility of which is to be checked. Its last index indicates that wall plane

 0001-4966/89/070172-07$00.80

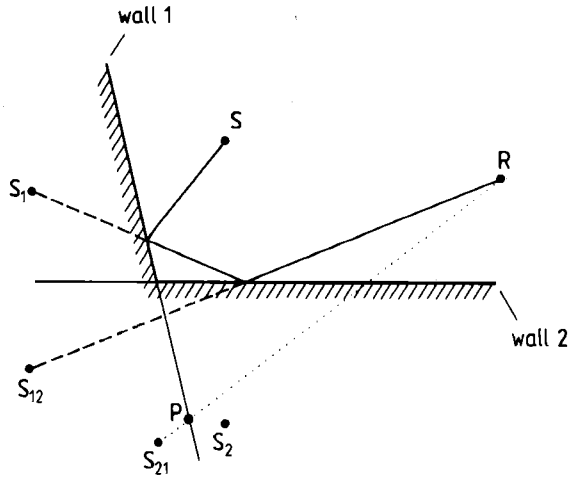

FIG. 1. Geometrical back tracing of sound paths from the receiver R to the original sound source. The image source S_{12} is visible, but S_{21} is not, since the intersection of the straight line RS_{21} (dotted line) with plane 1 is not placed within the boundaries of wall 1.

2 was involved last in the complete sound path and in the mirroring algorithm. If the intersection point of this connecting line is situated within the polygon that defines the actual wall portion on this plane, the image source is considered visible. In this case, we proceed in the same way, considering now the intersection point as receiver point R and aiming at the *predecessor* (S_1) in the chain of image sources. This procedure is repeated until it ends at the original sound source. If at least one of the calculated intersection points is not situated within the real wall boundaries, the image source is not visible and, hence, will be omitted. Otherwise, its energy, which depends on wall absorption and on the directivity of the original source, is stored along the time axis at a position according to its distance. There are no statistical errors whatsoever entering this algorithm, but an important criterion for its efficiency is the achievable maximum order of image sources.

The computation time mainly is dependent on the desired maximum order of image sources. The number of sources up to the order i of a room with n_w walls increases exponentially with i:

$$N_{\mathrm{IS}} = \sum_{k=0}^{i-1} n_w (n_w - 1)^k = \frac{n_w}{n_w - 2} \left[(n_w - 1)^i - 1 \right]. \tag{1}$$

The crucial part of the visibility test is to check whether an intersection point is contained within a certain polygon (see above). This is done by forming the cross products of the vectors pointing from the intersection point to the corners of this polygon.[1] Accordingly, for each corner of the polygon, at least six multiplications have to be carried out. On average, a polygon is defined by four to six corners and, according to experience, the test on "intersection point within a polygon" lasts an elementary time t_{IS} of a few milliseconds on IBM-AT compatible personal computers with a clock frequency of 8 MHz and math coprocessor. Another question is how much of these tests are required to find an invisible source. (As will be shown in Sec. III, almost all sources are invisible.) On average, we have to trace back the chain of successive image sources down to half the length of the chain with i denoting the maximum length. In total, the examination of one image source is consuming a computation time within an interval from t_{IS} to $(i/2)t_{\mathrm{IS}}$ at most. Accordingly, a lower limit for the total computation time per receiver point can be estimated as

$$t_R \approx t_{\mathrm{IS}} N_{\mathrm{IS}}. \tag{2}$$

This may be illustrated by an example: If we want to calculate impulse responses up to a duration of $t_{\max} = 400$ ms for a room consisting of 30 walls with a volume of 15 000 m^3 and an average number of reflections per second of $\bar{n} = 25.5\,\mathrm{s}^{-1}$, image sources of up to tenth order have to be tested. From Eq. (1), their number is $N_{\mathrm{IS}} \approx 4.5 \times 10^{14}$. That immense quantity can be reduced by a constant factor with a more or less complicated preelection of sources,[3] but, in principle, there is no way to escape the exponential law. Thus the total computation time according to Eq. (2) amounts to some 10 000 years using IBM-AT compatible computers. Even with a faster computer, it would not be possible to arrive at an acceptable computation time due to the exponential law.

For this reason, the conventional image-source algorithm is useful only for the following cases: (1) for very short impulse responses with highest time resolution; (2) for small numbers of walls n_w; and (3) for the simple rectangular room, since the complicated visibility test can be omitted. Additionally, there exists a simple law for constructing the image sources.

II. THE RAY-TRACING METHOD

Like the image-source model, the ray-tracing method is based on geometrical room acoustics, but it allows, in addition, the inclusion of sound scattering as it occurs when sound is reflected from rough surfaces. These effects are not negligible if we do not want to restrict ourselves to low frequencies.

We imagine the sound to be radiated from a source in the form of particles, sent out omnidirectionally. Each of these particles carries a certain energy, travels with the velocity of sound, and once in a while it will hit the boundary of the room (in the following called "wall" for simplicity). From the wall, it is reflected as shown in Fig. 2; then it will hit another wall, and so on. At each reflection the wall absorption is taken into account. By means of certain detectors, the particle energies and the arrival times with respect to the moment of release can be recorded. After tracing all particles and classifying the obtained data, the impulse response at each detector location can be constructed.

In earlier articles, an efficient ray-tracing program was described, and its accuracy with respect to its computation time was examined.[5,6] There are two alternative algorithms to be distinguished differing in the way in which wall absorption is taken into account. For the present, we assume that absorption is not accounted for by controlled particle annihilation but by multiplicative reduction of the particle energy. The reason for this choice will become obvious below.

The computation time for the ray tracing may be estimated as follows: Provided the direction of one particle is given, one has to find first the wall that it will hit next. This

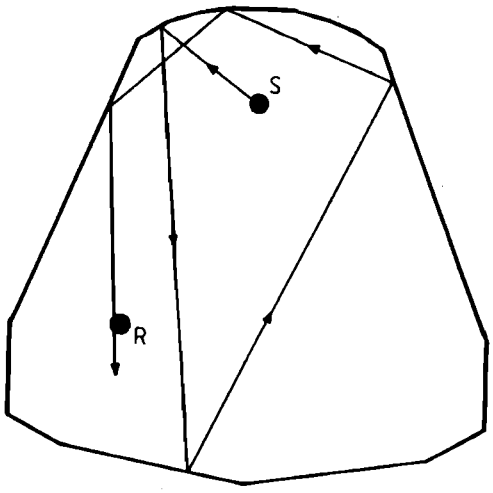

FIG. 2. Tracing a sound particle from the sound source S to the detector R.

involves the investigation of all n_w walls, each of them taking a time t_w. At the same time, one has to determine which of the n_k detectors have been hit by the particle since the last reflection, consuming the time $n_k t_k$. Finally, we must compute the new direction in which the particle will leave the wall after reflection, which takes time t_c. Accordingly, each reflection process requires the "elementary time":

$$\tau \approx n_w t_w + n_k t_k + t_c. \tag{3}$$

This time is multiplied by the number of reflections $\bar{n} t_{\max}$ of a particle during the period $t_{\max}$ and, of course, by the total number of particles N of one burst. Thus the total computation time is

$$t_R \approx N \bar{n} t_{\max} \tau. \tag{4}$$

For our example ($n_w = 30$, $n t_{\max} \approx 10$, $n_k = 1$), Eq. (4) yields $t_R \approx 12$ h with an IBM-AT compatible personal computer.

III. DETERMINATION OF VISIBLE IMAGE SOURCES BY RAY TRACING

In this section, a new method of sound field computation is presented, which combines the advantages of both the sound particles and the image sources. Its application is restricted to purely specular reflections, as is the case for any model involving image sources.

As we have seen in Sec. I, the conventional image-source model has the disadvantage that the relatively few visible sources must be picked out of an immense number of possible sources. To get an idea of the orders of magnitude, we go back to our example of an enclosure made up of 30 walls and an average number of reflections per second of 25.5 s^{-1}. The maximum order of image sources is, as mentioned, $i = \bar{n} t_{\max} \approx 10$, and their total number is $N_{IS} \approx 4.5 \times 10^{14}$. The number of the visible ones can be estimated for a rectangular room with volume V after Cremer[7]:

$$N_{\text{Ref}} \approx \frac{4\pi c^3}{V} \int_0^{t_{\max}} t^2 \, dt = \frac{4\pi (c t_{\max})^3}{3V}. \tag{5}$$

If we insert $t_{\max} = 400$ ms and $V = 15\,000\ \text{m}^3$, Eq. (5) yields $N_{\text{Ref}} \approx 700$. (Originally, this formula was derived for rectangular shapes only; it has been shown, however, that it may be applied as well to enclosures of any shape.[8,9])

The ratio of the total number of sources to the number of visible ones is as high as 10^{12}! Now it is easy to see that the tedious and time consuming process of separating the visible image sources from the invisible ones can be done by ray tracing. The only thing we have to do is to note the index of each wall from which a detected particle has been reflected during its history. From this list, the position of the image source can be calculated. Furthermore, from the absorption coefficients of the walls involved in the particle's history, its relative strength can be determined. Once the locations and strengths of all relevant image sources are known, we can forget our ray-tracing program and just add steady-state or transient contributions of these sources to construct the impulse response. This procedure explains, by the way, why it was the energy multiplication variant of the ray tracing that was to be used here and not the annihilation scheme.

There is a definite relationship between the number of counts at a given detector, on the one hand, and the total traveling path of a particle connecting the sound source and the detector, on the other, as shown in Fig. 3. Here and in the following, we assume spheres as particle detectors, since their sensitivity is omnidirectional. Whenever a particle crosses such a sphere, its arrival time, energy, and, eventually, its direction is stored.

Let N be the number of sound particles released in one burst of our ray-tracing source. They are radiated into the whole solid angle 4π. A sphere at distance r appearing under the solid angle Ω is hit by

$$k = N \, \Omega / 4\pi, \tag{6}$$

of them. Since (see Fig. 3)

$$\Omega r^2 = \pi r_k^2, \tag{7}$$

we obtain from Eq. (6)

$$k = N \, r_k^2 / 4 r^2. \tag{8}$$

Now, we identify $r = ct$ with the total distance a particle has traveled since its release at $t = 0$ (c = sound velocity). We require that, after the time $t_{\max}$, there is still one particle arriving at the detector. According to Eq. (8) with $k = 1$, this will be true if the sound burst contains at least

$$N_{\min} = 4 (c t_{\max})^2 / r_k^2 \tag{9}$$

particles. Our example with $t_{\max} = 400$ ms yields

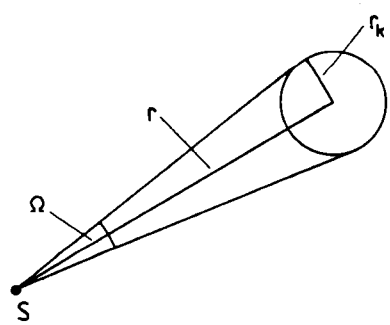

FIG. 3. Geometry of sound source and detector sphere with diameter $2r_k$ in distance $r = ct$.

$N_{min} \approx 74\,000$ if we assume $r_k = 1$ m. The computation will last less than 12 h from Eq. (4).

There remains one little problem to be mentioned. The image-source model assumes a point receiver. However, because of the finite extension of the detector sphere, some particles are counted and related to image sources that are not visible from the center of the sphere. This means that in the source list there are some "wrong" sources. Either one neglects these effects, which depend on the sphere diameter, or one uses the conventional image-source algorithm, namely, the back tracing of the paths from the receiver point to the original source in order to make sure that all found image sources are really visible from the center of the sphere. The latter procedure is accelerated by a factor N_{IS}/N_{Ref} with respect to the conventional image-source process, which in our example was 10^{12}! Thus the computation time of that added evaluation can be neglected relative to the procedure of ray tracing.

Figure 4 shows an impulse response of the example mentioned above, which contains all image sources up to the tenth order on the average with a time resolution that is limited only by the thickness of the plotter pen.

The enormous gain of computing speed for the room with 30 walls and $\bar{n} = 25.5\ s^{-1}$ is illustrated by Fig. 5. For this diagram, it has been assumed that the test of "point of intersection within an area" takes nearly the same time in both the particle and the image-source model ($t_w \approx t_{IS}$), according to Eqs. (2) and (4). Furthermore, the total number of emitted particles is N_{min} after Eq. (9); thus the total computation time increases only with the third power of t_{max}

$$t_R \approx (4c^2 \bar{n}\, t_{max}^3 / r_k^2)\ (n_w t_w + t_k + t_c), \qquad (10)$$

which is to be compared with the computation time of the conventional image-source algorithm after Eq. (2) in Fig. 5. For the calculation of these curves, t_k and t_c are neglected relative to $30\,t_w$; furthermore, r_k is assumed to be 1 m. Below 150 ms, the conventional image-source algorithm would require less computation time, but, for longer impulse responses, the computation time of the new method becomes rapidly shorter by orders of magnitude.

If several receiving points are to be examined, the relative gain in computing speed is still greater. With the conventional image-source model, the computation time is multiplied by the number n_k of receivers, but with the new method the computation time t_R given by Eq. (10) has to be modified by replacing t_k with $n_k t_k$ and becomes only slightly longer since t_R normally is dominated by $n_w t_w$.

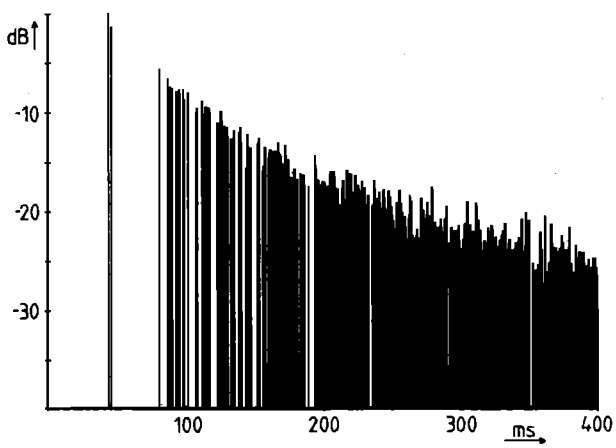

FIG. 4. Example of a highly resolved energy impulse response obtained with the new sound particle–image-source method.

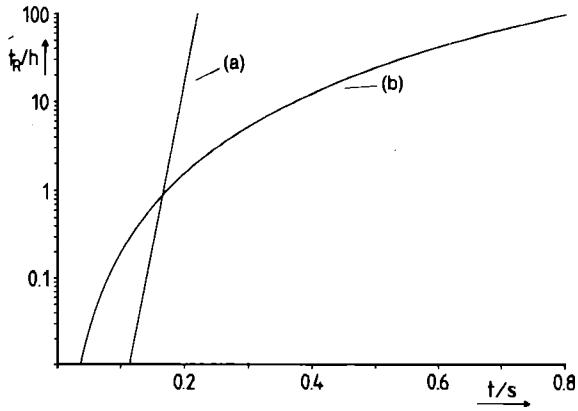

FIG. 5. Computation time in hours on a personal computer for a particular case as a function of the length of the impulse response. (a) Conventional image-source algorithm. (b) New sound particle–image-source algorithm.

If the frequency dependence of wall absorption is to be accounted for, one just has to attribute different power outputs to the image sources according to the wall reflections they represent, but the set of image sources is the same for all frequencies. This is a big advantage of the image-source model both in the conventional and the modified forms; in contrast, with the ray-tracing method including scattering, the whole process must be completely repeated for each frequency, i.e., for each set of absorption and scattering coefficients.

IV. CALCULATION AND INTERPRETATION OF ROOM TRANSFER FUNCTIONS

The possibility of calculating long and highly resolved impulse responses opens interesting possibilities for precise room acoustical analysis. The convolution of impulse responses with reverberation-free signals and examination by listening to the result was already mentioned. Furthermore, it is possible now to determine steady-state transfer functions.

For this purpose, we postulate that the walls have real impedance and that all diffraction effects on edges can be neglected. Then, the pressure impulse response $p(t)$ is composed of Dirac pulses and, hence, is just the square root of the energy impulse response. The room transfer function $H(f)$ is obtained as the Fourier transform of $p(t)$.

Figure 6 shows the geometry of the room that served as an example before. Two receiver positions are marked on the plan for which the impulse responses are computed in the described way. Some well-known room acoustical parameters characterizing the sound field are evaluated from the obtained energy impulse responses. The results are shown in Table I, which lists the definition D, the clarity C, the early decay time T_{10}, Barron's "lateral energy ratio" S, and the steady-state sound-pressure level L. (For explanation of

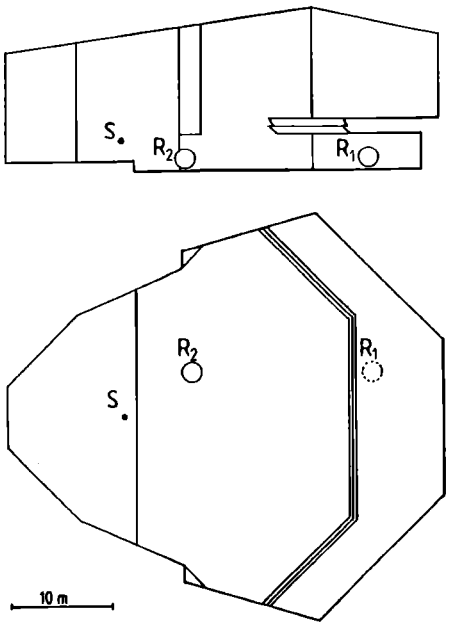

FIG. 6. Longitudinal section and ground plan of an auditorium with sound source and two receivers.

these quantities, see Ref. 5.) The latter is related to the corresponding free-field level at a distance of 10 m.

Figure 7(a) shows the energy impulse response calculated for receiver 1, (b) represents the octave-filtered sound-pressure impulse response obtained from it, and (c) shows the corresponding steady-state transmission curve $|H(f)|$. Finally, in Fig. 7(d), we see an enlarged section of the transmission curve. Figure 8 shows similar results for a receiver in the stalls (receiver 2).

A remarkable feature is the regular structure of the transmission curve obtained for receiver 2. It is caused by a strong reflection from the floor arriving immediately after the direct sound with a short delay. Its energy dominates the other reflections and produces the comb filter structure of the transmission curve. With broadband sound signals, we would probably perceive a characteristic tonal coloration at receiver point 2. Furthermore, for receiver 1, the average spacing between two maxima yields $\Delta f_{\max} = 6.6$ Hz, which agrees well with the theoretical value $3.91/T = 6.5$ Hz (Ref. 10) obtained with the reverberation time T_{10} (see Table I). For receiver point 2, however, we determine an average maxima spacing of 5.1 Hz, which does not agree with the expected value of 3.3 Hz.

Furthermore, the statistical distribution functions of the real part $\mathrm{Re}\{H(f)\}$, of the imaginary part $\mathrm{Im}\{H(f)\}$, and of the magnitude $|H(f)|$ have been evaluated. In a diffuse

TABLE I. Objective parameters for both receiver places.

Receiver position	D/%	C/dB	T_{10}/s	S/%	L/dB
1	60	7.2	0.6	41	−0.1
2	75	8.7	1.2	12	5.5

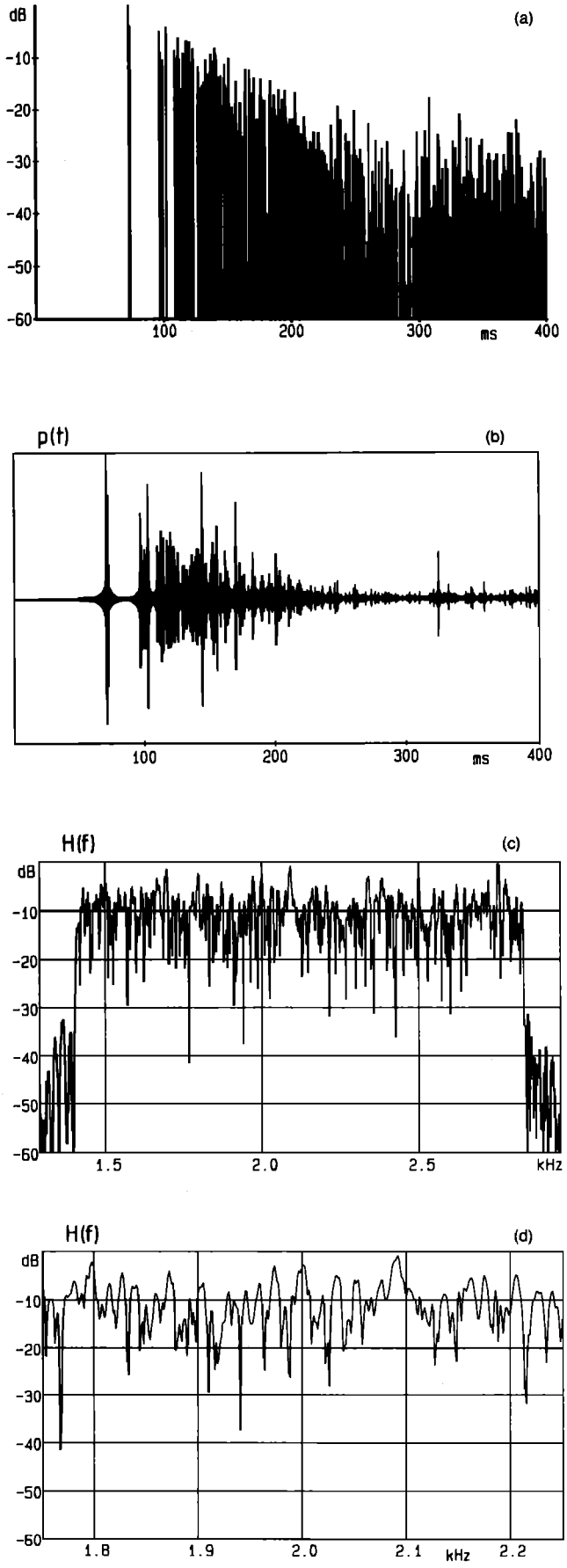

FIG. 7. Receiver 1: (a) energy impulse response; (b) 2-kHz-oct filtered pressure impulse response $p(t)$: (c) transmission curve $|H(f)|$; and (d) enlarged section of $|H(f)|$.

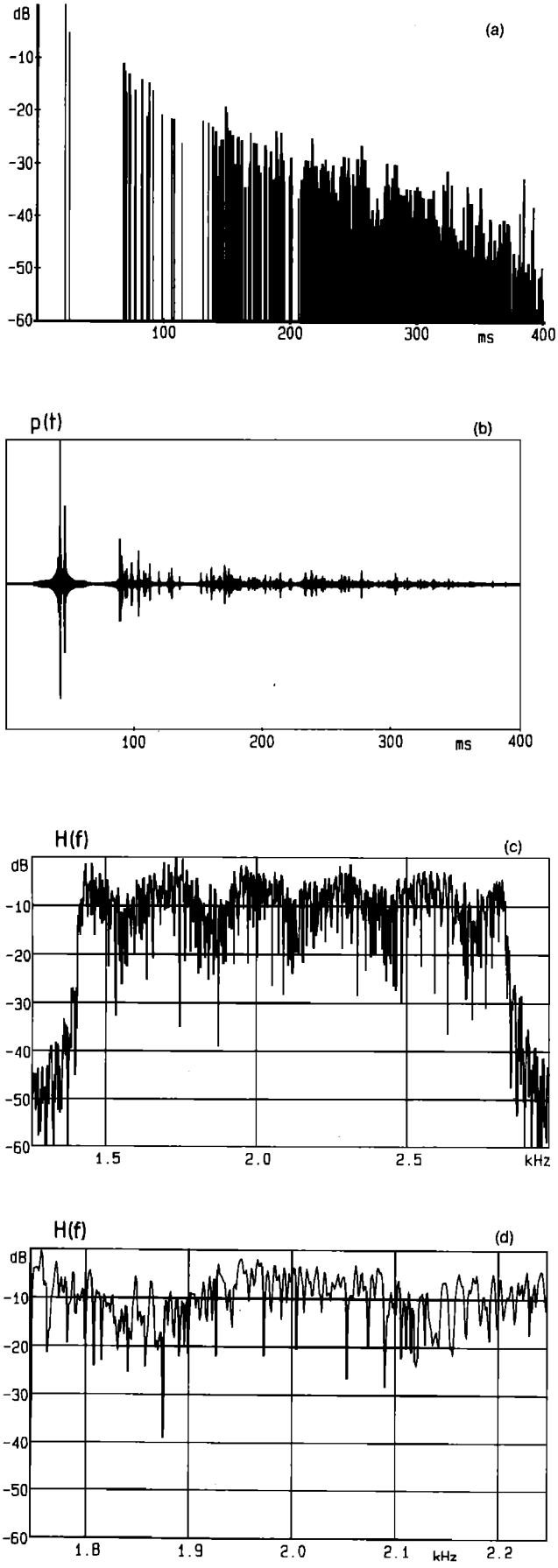

FIG. 8. Receiver 2: (a) energy impulse response; (b) 2-kHz-oct filtered pressure impulse response $p(t)$; (c) transmission curve $|H(f)|$; and (d) enlarged section of $|H(f)|$.

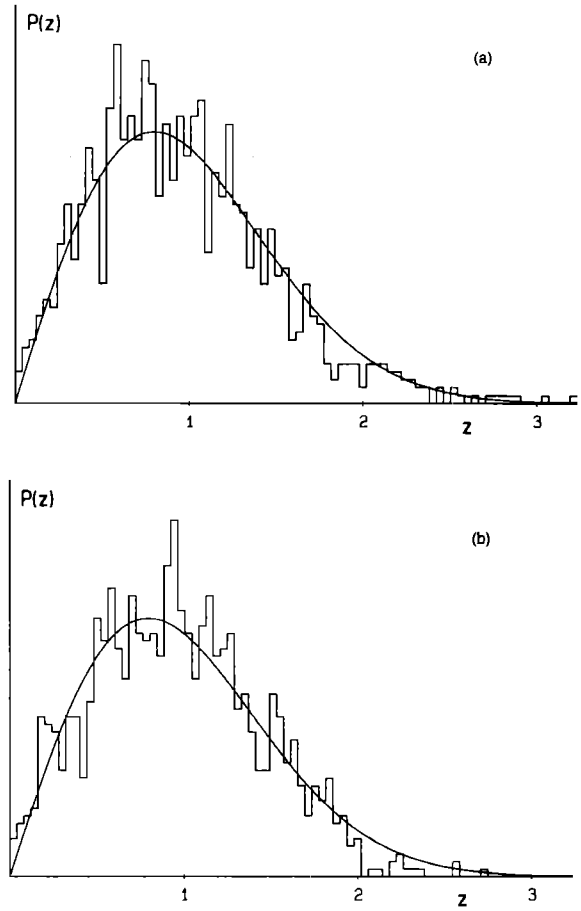

FIG. 9. Distribution density functions $P(z)$ for (a) receiver 1 and (b) receiver 2. Here, z is the absolute value of $H(f)$ divided by its frequency average; the smooth curves represent the Rayleigh distribution.

sound field, these functions follow well-known laws, so the real parts and the imaginary parts should be normally distributed, whereas the amplitudes should follow a Rayleigh distribution.[11,12] Other criteria for "reasonable" transmission curves are the autocorrelation function of both the real part and the imaginary part and their cross-correlation function (see Ref. 13). As an example, Fig. 9 shows the distribution function $P(z)$ for z as the absolute value of $H(f)$ divided by its frequency average for both receiver points. Again, receiver 2 shows an abnormal distribution, since the maximum value is shifted to the right. Hence, we conclude that, in this case, the sound field would not be diffuse, because the receiver is too close to the sound source. In contrast, the distribution obtained for receiver 1 appears quite normal.

V. DISCUSSION AND CONCLUSION

Most of the objective parameters used for the prediction and characterization of the listening conditions in a room are based upon its impulse response, obtained either by measurement or by computation. In contrast, characterization in the frequency domain has been neglected during the past decades and almost no research has been conducted in this field. This is not in the least due to the time-consuming measuring process by which transmission functions are experimentally

obtained and to the impossibility of computing them from room data.

The new method of a combined sound particle–image-source algorithm, which results in high-accuracy impulse responses and subsequent Fourier transform, solves the latter problem and, hence, reopens the frequency domain for room acoustics. It seems worthwhile to investigate the practical aspects of this kind of evaluation.

As discussed above, the new method is restricted to rooms that can be treated on the basis of geometrical acoustics. This is to say all diffraction effects and also scattering from rough walls are neglected. The attenuation due to grazing sound propagation above audience, however, can be included in our model. To which extent the neglect of scattering and diffraction will affect the results of convolution experiments with reverberation-free speech or music must be studied by suitable psychoacoustic investigations.

ACKNOWLEDGMENTS

The author is grateful to Professor H. Kuttruff for his helpful support of this work and for reading the manuscript. S. Müller is acknowledged for developing the excellent FFT program.

[1]S. Borish, "Extension of an image model to arbitrary polyhedra," J. Acoust. Soc. Am. **75**, 1827 (1985).
[2]J. Kirszenstein, "An image source computer model for room acoustics analysis and electroacoustic simulation," Appl. Acoust. **17**, 275 (1978).
[3]H. Lee and B.-H. Lee, "An efficient algorithm for the image model technique," Appl. Acoust. **24**, 87 (1988).
[4]A. Krokstadt, S. Strøm, and S. Sørsdal, "Calculating the acoustical room response by the use of a ray tracing technique," J. Sound Vib. **8**, 118 (1968).
[5]M. Vorländer, "Ein Strahlverfolgungsverfahren zur Berechnung von Schallfeldern in Räumen," Acustica **65**, 138 (1988).
[6]M. Vorländer, "Die Genauigkeit von Berechnungen mit dem raumakustischen Schallteilchenmodell und ihre Abhängigkeit von der Rechenzeit," Acustica **66**, 90 (1988).
[7]L. Cremer, *Geometrische Raumakustik* (Hirzel Verlag, Stuttgart, 1948), p. 27.
[8]H. Kuttruff, *Room Acoustics* (Applied Science, London, 1979), 2nd ed., p. 87.
[9]M. Vorländer, "Untersuchungen zur Leistungsfähigkeit des raumakustischen Schallteilchenmodells," Ph.D. thesis, RWTH Aachen, Federal Republic of Germany (1988).
[10]M. R. Schroeder and K. H. Kuttruff, "On frequency response curves in rooms. Comparison of experimental, theoretical, and Monte Carlo results for the average frequency spacing between maxima," J. Acoust. Soc. Am. **34**, 76 (1962).
[11]M. Schröder, "Die statistischen Parameter der Frequenzkurven in grossen Räumen," Acustica **4**, 594 (1954).
[12]H. Kuttruff and R. Thiele, "Uber die Frequenzabhängigkeit des Schalldrucks in Räumen," Acustica **4**, 614 (1954).
[13]M. R. Schroeder, "Frequency-correlation functions of frequency responses in rooms," J. Acoust. Soc. Am. **34**, 1819 (1962).