INTER-NOISE 2007
28-31 AUGUST 2007
ISTANBUL, TURKEY

Virtual Acoustics for Localisation of the Speaker in a Real and Virtual Reverberant Room

Monika Rychtáriková[a] ,Tim van den Bogaert[b],
Gerrit Vermeir[a], Jan Wouters[b]

[a] Katholieke Universiteit Leuven
Laboratorium ATF
Celestijnenlaan 200D, bus 2416
3001 Heverlee
BELGIUM

[b] Katholieke Universiteit Leuven
Exp ORL, Dep. Neurosciences
O.& N2, Herestraat 49, bus 721
3000 Leuven
BELGIUM

ABSTRACT

This study investigates the usage of virtual acoustics in the framework of a research project dealing with the development of the binaural hearing aids and cochlear implants, with emphasis on the localisation of sound in different acoustical scenarios.
Virtual acoustics allows the convolution of a measured head related transfer function (HRTF's) with the impulse response of a room generated by a computer model (using ODEON® software) to realistically simulate binaural sound.
We report on (loud)speaker localization experiments conducted in a reverberant room, where eight listening subjects were asked to localize sound sources in the front horizontal plane. As a stimulus in the experiment, a sound of a ringing telephone was chosen.
The stimulus was played individually (1) as virtually 3D sounds via headphones, based on *in situ* binaural room impulse responses recorded by an artificial head, and (2) via headphones, as virtual sources simulated in ODEON® software (using HRTF receiver properties as measured on the artificial head)
The results from the reverberant room are compared with listening tests performed in anechoic conditions with the same listening subjects.

1 INTRODUCTION

Virtual acoustics combines sound source, listener and room properties to generate audible information (e.g. auralized sound) from impulse responses calculated for an existing or virtual room. The first auralization attempts were monaural and probably date from the time of the first experiments with frequency-transformed music signals played in a scale model by Spandöck in 1934 [1]. Auralization thus dates from before the appearance of personal computers and simulation programs. It seems that the first computer-based auralization experiments were done in the 1960's by Schröder [2], and in the 1970's by Allen and Barkley [2]. Binaural auralization was introduced a decade later in the work of Pösselt [4], then followed a large amount of other researchers. [5],[16]

In the engineering practice, auralization has been proven to be a useful tool to assess the general acoustical comfort by end users during the stage of the room / building designing. In this paper we demonstrate the usage of an auralization method in the framework of a research project dealing with the localization of sound when using noise reduction systems on binaural

[a] Email address: Monika.Rychtarikova@bwk.kuleuven.be

hearing aids in different acoustical scenarios. It is known that noise reduction schemes can alter the binaural cues which are responsible for localizing sound sources [19]. Therefore these schemes should be evaluated on localization performance in several different acoustical environments. Moving research equipment to different acoustical environments is time consuming and hence expensive. Therefore the possibility of using virtual acoustics for localization experiments was brought up. This paper evaluates whether experiments with sound files generated by virtual acoustics software are as good as doing these experiments with binaural head related transfer functions measured *in situ* at the different acoustical locations.

1.1 Localisation of Sound

Localisation of a sound source is a rather complex process that can be studied from several points of view. In general one can distinguish between horizontal and vertical localisation, and distance perception, or focus on externalisation or front-back confusion. Here we deal with the frontal horizontal directional localisation of a sound source.

1.1.1 Localisation in the Frontal Horizontal Plane

Hearing based localisation of a sound source in the frontal horizontal plane is possible mainly thanks to our binaural hearing. Our brain analyzes the two signals coming from the left and right ear with respect to the difference of arrival time between both ears (with respect to the source) and by (spectrally) comparing the amplitude of the sound in our two ears.

The difference in the arrival time is called ***Interaural Time Difference (ITD)***, which for steady sound of pure tones can be considered as an interaural phase difference. Psychoacoustical experiments show the considerable ability of a human to localize frequencies around 500-800 Hz with significant accuracy. However, once the wavelength of the incident sound becomes too small in comparison with the distance between the ears, the phase difference increases above 2π, thus becoming ambiguous, resulting in a rapid decrease in ITD interpretation at high frequencies. It has been confirmed by systematic experiments that ITD-based source localization works as a reliable cue up to 1-1.5 kHz.[20] Headphone experiments show that the accuracy to localize sound increases as the source is more frontal, reaching a maximum for the sources directly in front of the listener. In that case, humans are sensitive to differences as small as 1-2°, corresponding with an ITD around 12μs. [6]

Our head acts as a sound barrier for arriving acoustic waves, which, depending on the angle of incidence, puts one of the two ears in an acoustical shadow. This results in a difference between the sound intensities in the left and right ear, often referred to as ***Interaural Level Difference (ILD).*** The smallest detectable ILD measured in psychoacoustical experiments is approximately 0.5 dB. In headphone experiments, the same sensitivity of the neural system was measured for every frequency. [7] However, in practice, for frequencies below 500 Hz the ILD becomes too small due to wave diffraction around the head. ILD-based localization is most efficient for frequencies above 4000 Hz, where the acoustical shadow caused by the head is most pronounced.

1.1.2 Localisation in the Vertical Plane

Localisation of sound in the medial vertical plane is not possible using ITD and ILD since due to geometrical symmetry there is no difference in level and arriving time between the two ears. Nevertheless, it was shown that details on the listeners head, his outer ear, and the whole shape of his upper body, act as an angle dependent acoustical filter for the signal in the ears. [8]

1.1.3 Head-related transfer function (HRTF)

Together with ITD and ILD effects, this filtering (through pinnae, head, and upper body) is included in the ***Head-related transfer function (HRTF)*** and it is defined in frequency domain as the ratio between the Fourier transform of the signal at the ear drum and the Fourier transform that would have been received at the same position without head. The HRTF depends on the direction of the incident sound wave, defined by its azimuth Θ and elevation Φ, on the distance r between the head and the sound source, and on the frequency f. [9],[21] In order to insure well-defined geometrical parameters, the experimental determination of the HRTF has to be performed in anechoic circumstances. Measurements of HRTF on people show quite large differences between different individuals or artificial heads. [10],[11] It is therefore difficult to generalize the common features in localising in median plane.

1.1.4 Anechoic versus Normal Life Circumstances

In anechoic circumstances we deal with behaviour of purely direct sound around the head. In the real life, humans spend most of their time in an environment where the sound of a source is reflected from different objects or room surfaces, resulting in reflected wave arrivals from many directions. It can thus be expected that reverberation decreases our ability for directional localisation. The ILD is e.g. very sensitive to standing waves in the room and the ITD can suffer from non-coherent sound signals resulting from multiple sound reflections. However, as we know from our personal experience, even in very reverberant scenarios, our ability to localise sound remains quite good.

From the acoustical point of view, the absorptive properties of most of the interior materials usually increase with frequency, rendering the direct-to-reverberant sound ratio larger towards higher frequencies. The ILD above 8 kHz can be therefore a potential cue to localize the direction of the (direct sound wave dominated) incoming sound. [6]

From the point of view of perception, the phenomenon called ***precedence effect*** (known also as law of the first wave front) occurs in non-anechoic situations. This effect is based on a ability to map the direction of the reflections to the direction from where the direct sound source is coming from. The time interval in which this effect takes place is also called the fusion zone. The precedence effect is most effective in situations where the spectra of the direct and reflected sound are similar. [12]

The perceptual process of our ear-brain system is much more complex than the results we can obtain from the analysis of binaural recordings. Recent research suggests that the acoustical information is sent to higher centres in the brain, where it is also compared with the information obtained visually. [6] Many aspects of this process are still not understood and several opportunities for useful research exist.

1.2 Virtual Modeling of Rooms

Virtual modeling of binaural sound in a room is in general based on the simulated binaural room impulse response (BRIR), convolved with anechoic sound samples.

The BRIR can be obtained in several ways. In computer simulations, acoustical phenomena can be either described by waves (Finite Element Method or Boundary Element Method), or by particles (Ray-tracing, Cone Tracing or Image Source Methods.)

Wave models are based on solving the wave equations on a 3D grid defined by the user. They can give accurate results at a single frequency but the calculation time is often too long for practical purposes, since min. 6 elements per wavelength are necessary for obtaining reliable results. Moreover, the number of natural modes in the room increases with the third power of the frequency. [13] This is the reason why FEM methods are used in practical applications only for the calculation of small rooms and/or for the lower frequency range.

In architectural acoustics, geometrical methods based on the sound propagation along rays are much more used. These methods usually give reliable results in the high and middle frequencies and their calculation time is much shorter than the one of wave-based programs.

The ***Image Source Method (ISM)*** is based on the principle that a sound ray reflected from a plane can be drawn as if it would originate from an image source that can be evoked as a mirror source considering the reflecting plane in the model as a mirror plane. Also secondary image sources of the initial image sources can be introduced and reflections of the second-order, third-order etc can be calculated. The more surfaces the acoustical model of the room contains, the more image sources has to be constructed. If n is the number of surfaces in the model and i is the reflection order, then the number of possible image sources is approximately [13]:

$$N \approx (n-1)^i \qquad (1)$$

For this reason, the exclusive usage of ISM is suitable only for rooms with a not too complicated geometry and not too much reverberation.

In the ***Ray Tracing Method (RTM),*** a number of rays is sent from a point source, usually in all directions. The trajectory of each ray is determined by the reflection of the ray from the boundary surfaces according to Snell's law. The intensity of a ray is diminished according to the absorption coefficient of the incident surface and to the absorption of the sound in the trespassed air trajectory. Also scattering of the sound is introduced in the computer model, via the scattering coefficient δ [%], which is defined as the ratio between the sound energy reflected in non-specular reflections to the total reflected sound energy. [13]

1.2.1 Binaural Room Impulse Response and Auralisation

For the room acoustical simulations in our research, ODEON® software was chosen. This software uses a hybrid algorithm where two geometrical methods are combined. The simulation of the impulse response (IR) of rooms is therefore performed in two parts: (1) a receiver-dependent part and (2) a receiver-independent part.

The first part, which contains information about early reflections, is calculated combining the ISM and RTM. The second part of the IR, i.e. the late reflections, is calculated by ray-tracing by taking into account also the scattering coefficient of surfaces. Diffuse secondary sources are generated that radiate sound locally from interior surfaces with a directivity according to Lambert's cosine-law.

The duration of the receiver-dependent part can be chosen in the software via the so-called Transition order (TO). This is the maximum number of image sources per initial ray. [14], [15], [17]

The BRIR in Odeon is calculated at a receiver point by filtering the calculated room impulse response with the HRTF. Binaural filtering is achieved by filtering every arriving ray according to its direction of incidence. [13] The advantage of Odeon is the possibility for the user to introduce arbitrary HRTF information. This information can be based on measurement on a real subject or artificial head, or simulated by the boundary element method (BEM). [18]

2 MEASUREMENTS AND SIMULATIONS

The main goal of our experiment is to understand the ability of a person to identify one of the thirteen loudspeakers in the frontal horizontal plane (Figure.2) in the real and virtual acoustical scenario of the reverberant room.

2.1 Measurements in the Anechoic room

Initial HRTF measurements of the artificial head were performed in an anechoic room, in the horizontal plane with the precision of 15°. Single cone loudspeakers with a cone diameter

of 10 cm, were hung on a metal ring with a diameter of 2.2 m and an artificial head was placed in the middle of this ring. The results from these measurements were later used for:

(1) the creation of sound samples for the listening test in anechoic conditions by convolving the dry measured binaural impulse responses with the three chosen stimuli.
(2) the introduction of measured HRTF information in the Odeon software for the calculation of the BRIR and subsequent auralisation

In the anechoic room the directivity of the loudspeaker was measured as well, to be afterwards used in the simulation.

2.2 Measurements in the Reverberant room

Measurements of binaural impulse responses (with the same artificial head as in 2.1) in the reverberant room were done for the same (loudspeakers-receiver) set up as in the anechoic room (Fig.3), i.e. for loudspeakers at 1m distance from the head. Later, to investigate the ability of people to localize the sound source at distances larger than the Hall-Radius [22], loudspeakers were put on stands and placed at 2.4 m distance. (Fig.2) The measured impulse responses from both experiments were used (after convolution with anechoic sound sample) for the listening tests.

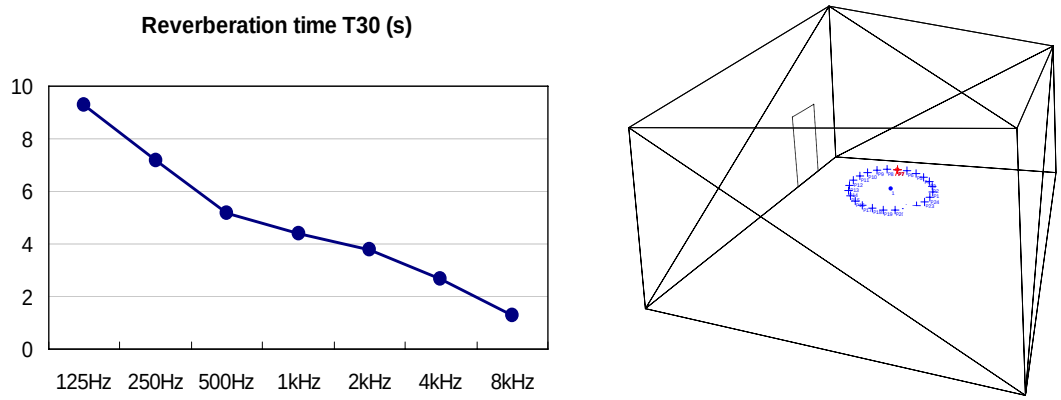

Figure 1: Measured reverberation time of the reverberant room under described conditions (left) Model of the virtual reverberant room from Odeon (right)

2.3 Simulation of the Reverberant Room

The geometrical computer model of the reverberant room was based on the measured dimensions of the room. Then, the reverberation time of the room was experimentally determined, by using an omni-directional point source and an omni-directional microphone. During the measurements, equipment (such the loudspeakers) and one person sitting on the chair was present in the room, to approach as much as possible the absorption properties of the room conditions as during the listening tests. The acoustical model of the reverberant room in Odeon was calibrated according to this *in situ* measurement. (Fig.1)

The loudspeakers were simulated with the proper directivity and spectrum, as measured in anechoic room. The receiver properties were defined by the measured HRTF of the artificial head. The simulation was performed with 6000 rays, a maximum reflection order of 2000 and transition order TO = 2.

3 LISTENING TESTS

Listening tests were performed in the (1) anechoic and (2) reverberant room, with the same listening subjects. In all listening tests, the task was to identify, the loudspeaker from which the sound was heard, out of the 13 possibilities. Samples were played randomly, with three repetitions of each speaker in every test. Every person has participated on the same test

two times (test and retest) on two different days. All listening persons who have participated in the experiment had normal hearing, which was confirmed by their audiograms.

3.1 Listening Scenarios

3.1.1 Listening Scenarios in the Anechoic Room

In the ***anechoic room,*** the listening person was sitting in the middle of the ring with 13 loudspeakers located in the frontal horizontal plane between the azimuth angles -90° and 90° (every 15°), at a distance of 1m, with his or her ears at the height of the loudspeakers, around 1.2m. Sound samples based on measured anechoic binaural signals were played from headphones using an RME Hamerfall DSP II soundcard and the listening subject was asked to identify the virtual loudspeaker (HPM)

3.1.2 Listening Scenarios in the Reverberant Room

Listening scenarios in the ***reverberant room*** were created for the real and the simulated reverberant room. The listening subjects were sitting in the real reverberant room, identifying the loudspeaker:

(1) played via headphones based on measurement done on artificial head in the reverberant room (HPM)

(2) played via headphones based on the simulation with the task to identify the virtual loudspeakers in the simulated reverberant room (HPS)

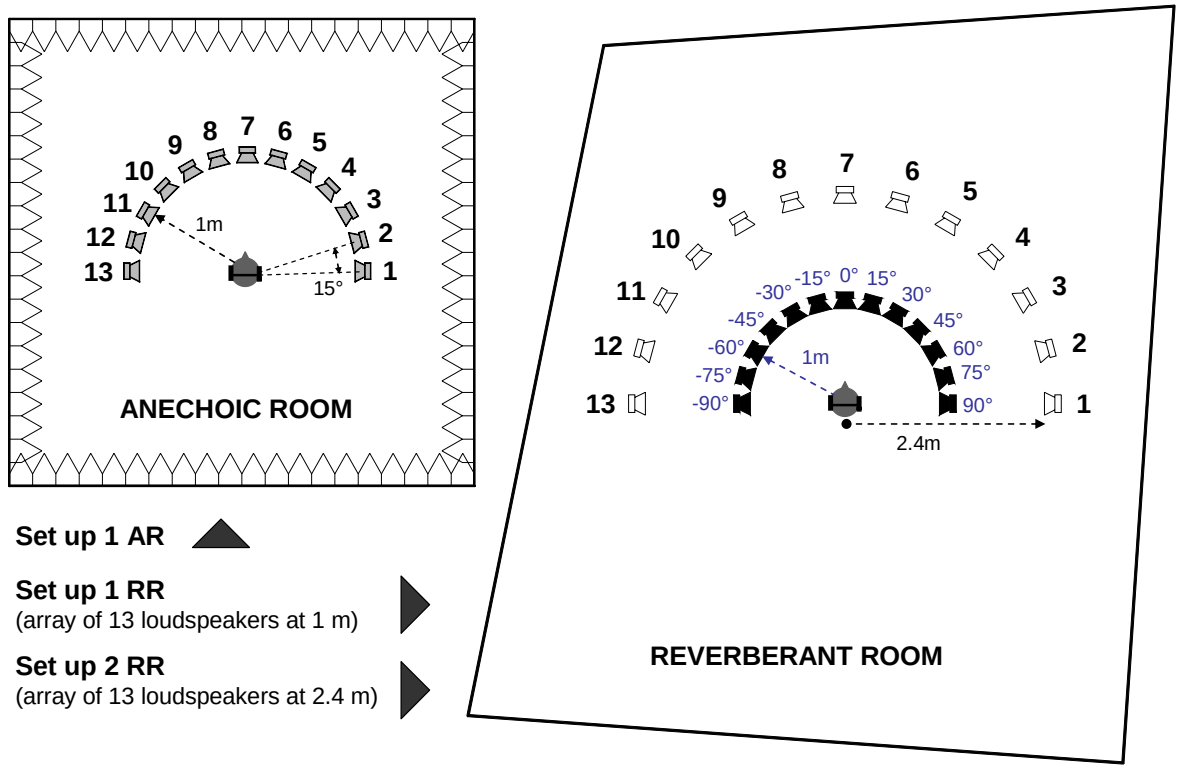

Figure 2: Top view of the set-up in the Anechoic and Reverberant room.

In the headphone tests (HPM and HPS), the subjects were sitting in the same position in the room in the presence of thirteen inactive real loudspeakers, for the sake of supplying visual association.

3.2 Stimuli

For our experiment, a broadband telephone ring sound (duration 1 second) which contains both, low and high frequency components, was chosen. Sound in all the listening tests were presented at 68 dB SPL and a roving level [19] of 4 dB was used.

4 ERROR CALCULATION

To analyze the results, the root mean squared error (RMS) in degrees was used:

$$\mathrm{RMS}(^\circ) = \sqrt{\frac{\sum_{i=1}^{n} (\text{stimulus-response})^2}{n}} \qquad \text{(Eq.1)}$$

where n is the number of played stimuli.

5 ANALYSIS AND RESULTS

In the comparisons between the tests and the retests, no significant differences were found. This has allowed us to use the average RMS error of the two tests in the following analysis. The following trends were observed:

In the reverberant room the two set-ups (based on the distance between the listener and loudspeakers 1m and 2.4 m) were first analyzed separately, and afterwards compared with each other. In setup1RR (Tab.1), people were localizing the source with very high probability around 10° RMS, in spite of the very long reverberation time. No significant difference between the different listening scenarios (HPM and HSM) was found.

Table 1: Comparison of average RMS errors per listening subject for 3 listening test-ups and listening scenarios (HPM – measurement based stimuli played via headphones, HPS – simulation based stimuli, played via headphones)

RMS error (°)	**Anechoic**	**reverberant Room (1m)**		**reverberant room (2,4m)**	
	HPM	HPM	HPS	HPM	HPS
LS1	6,8	10,8	11,6	8,4	11,7
LS2	6,7	5,6	6,8	9,6	13,2
LS3	9,9	9,6	7,6	10,6	14,0
LS4	7,0	7,0	9,3	10,2	11,7
LS5	14,7	17,6	20,3	21,7	18,0
Average (°)	**9,0**	**10,1**	**11,1**	**12,1**	**13,7**
Stdev (°)	**3,4**	**4,7**	**5,5**	**5,4**	**2,6**

In general, it seems that the RMS errors slightly increase in the presence of reverberation and with increasing loudspeakers-listener distance in the reverberant room.

The evaluation of the performance for HPS was done related to the HPM case since here the same HRTF was used. It seems that the simulation method is a good tool for directional localisation tests in frontal horizontal plane. More research is being done at the moment on testing low and high frequency sounds separately to validate the simulation software on the two main different localization methods separately (ITD processing by using simulations with lower frequencies and ILD processing by using simulations with higher frequencies). Also we are testing the same subjects using their own ears to localize sounds in these setups to validate the localization performance when using headphones with pre-recorded or simulated HRTF's compared to localizing the same sounds when the subject is using its own ears.

6 ACKNOWLEDGEMENTS

This research was financed through the FWO-Vlaanderen project G.0334.06. The authors would like to express special thanks to Tinne Boons and to Michal Jelínek for their help during the listening tests and to people who have participated at the tests as listening subjects.

7 REFERENCES

[1] F. Spandöck, *Ann. Physik* V, **20**, p.345 (1934)
[2] M. R. Schröder, B. S. Atal, C. Bird, in *Proc. 4th Int. Cong. Acoustics* (Copenhagen, Denmark, 1962, paper M21
[3] J. B. Allen and D.A. Berkley, J. Acoust. Soc. Am., **65**, p.943 (1962)
[4] C. Pösselt, in Forschrifte der Akustik, DAGA '87, DPG-GmbH, Bad Honnef, Germany, p.725 (1987)
[5] K. H. Kutruff, "Auralisation of Impulse Responses Modeled on the Basis of Ray-Tracing Results, J.Audio Eng. Soc. **41** (11), 876-880 (1993)
[6] W. M. Hartmann, "How We Localize Sound", Physics Today, 24-29, November (1999)
[7] W.A. Yost, J.Acoust.Soc.Am. **62**, 157 (1997)
[8] W. M. Hartmann, T.L.McCaskey, "Identification and Localisation of Sound Sources in the Median Sagittal Plane", Acoust.Soc.Am, **106** (5), 2812-2820 (1999).
[9] D. N. Zotkin et all: "Fast head-related transfer function measurement via reciprocity" J.Acoust.Soc.Am., **120** (4). 2202-2214 (2006).
[10] Moller et al, "Head-Related Transfer Functions of Humans subjects" J.Audio Eng.Soc. **43**, 300-321 (1995)
[11] P. Minnar et al, "Localisation with Binaural Recordings from Artificial and Human Heads", J.Audio Eng.Soc. **49** (5), 323-335 (2001)
[12] E. F. Toole, "Loudspeakers and Room for Sound Reproduction – A Scientific review", J.Audio Eng.Soc., **54** (6), 451-476 (2006)
[13] J. H. Rindel, "The Use of Computer Modeling in Room Acoustics", Journal of Vibroengineering, **2** (4), 219-224 (2000)
[14] M. Lisa, J. H. Rindel, C. L. Christensen, "Predicting the Acoustics of Ancient Open-Air Theatres: The Importance of Calculation Methods and Geometrical Detailes", *Joint Baltic-Nordic Acoustics Meeting,* (2004)
[15] Odeon Manual
[16] M. Kleiner, B.-I. Dalenbäck & P.Svensson, "Auralisation – An Overwiew", J.Audio Eng. Soc. **41**, 861-875 (1993)
[17] M. Rychtáriková, G. Vermeir, "Comparison Between Image Source Method and Ray tracing method in the Binaural Room Impulse Responses, *Proceedings of the 2nd International Symposium "MAP 2006"*, Zvolen, Slovakia, 2006
[18] J. Fels, P. Buthmann, M. Vorländer: "HRTF's of Children, Acta Acustica", **90**, 918-927 (2004)
[19] T.vd. Bogaert et al, "Horizontal Localisation with bilateral hearing aids: Without is better then with", J.Acoust.Soc.Am. **119**, 515-526 (2006)
[20] B.Ross et al, "Physiological detection of interaural phase differences", Acoust.Soc.Am. **121**, (2) 1017-1027 (2007)
[21] C. I. Cheng, G.H.Wakefield, "Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequencey and Space, J.Audio Eng. Soc. **49** (4), 231-249 (2001)
[22] L. Cremer et al,.: "Principles and Applications of Room Acoustics. Vol 1." ISBN 0-85334-113-3 (1982)