PAN
ARCHIVES OF ACOUSTICS
Vol. **40**, No. 2, pp. 205–211 (2015)

Copyright © 2015 by PAN – IPPT

DOI: 10.1515/aoa-2015-0023

Real-Time Acoustic Phenomena Modelling for Computer Games Audio Engine

Bartłomiej MIGA, Bartosz ZIÓŁKO

Faculty of Computer Science, Electronics and Telecommunications
AGH University of Science and Technology
Al. A. Mickiewicza 30, Kraków, Poland; e-mail: bziolko@agh.edu.pl

(*received June 23, 2014; accepted February 19, 2015*)

This article presents an efficient method of modelling acoustic phenomena for real-time applications such as computer games. Simplified models of reflections, transmission, and medium attenuation are described along with assessments conducted by a professional sound designer. The article introduces representation of sound phenomena using digital filters for further digital audio processing.

Keywords: sound reflection, transmission, attenuation, real-time audio processing.

1. Introduction

Designers and producers of computer games were very focused on graphics of games often leaving in this way other multimedia for behind. Graphical engines are typically the most important tools for each company producing games. It is difficult to evaluate whether computer games industry invests so much in graphics, quite ignoring the audio part due to real customers' expectations or because players simply do not yet know how good usage of audio elements can make games more attractive.

There are some games which invested in audio and resulted in huge commercial success, but these were rather based on interesting samples or good direct connection of audio element and the plot, like in the case of the game *Thief*. Our research focuses on creating an audio engine for computer games based on soundtracing (Pałka *et al.*, 2014). Ray-tracing is a technique for generating photorealistic images of three-dimensional scenes, allowing mapping of reflections from objects, generating realistic shadows, and lighting counting. Due to the complexity of computing for many years, its use was limited to the film industry. Soundtracing allows for accurate and powerful sound processing on multi-core architectures, taking into account the reflections from the walls, attenuation on obstacles, a precise calculation of the reverberation from multiple sound sources in an automatic way. At the moment soundtracing is used only in not real-time applications, mainly for virtual evaluations of buildings acoustics, i.e., concert or opera halls (Kamisiński *et al.*, 2009). The most commonly used software are: CATT, DAAD, Odeon (Ozgura *et al.*, 2004; Bellmann *et al.*, 2008). The real-time applications need new numerical methods of simulation, often giving results which are not as close to reality as the known methods working off-line. Our work described in this paper is focused on numerical calculations of sound reflection, transmission, and air absorption.

2. Models of the phenomena

The work described in this paper focuses on modelling of sound reflection, transmission, and air absorption. The suggested methods should be seen not as new models of these phenomena, but as approaches with simplifications aimed at applying known methods to efficient, real-time computations.

As it was mentioned in the introduction, our approach uses soundtracing, which is based on geometrical acoustics methods. Acoustic waves are modelled as beams which propagate within the virtual environment. Each beam originates in a sound source, placed in a scene. When a beam strikes an obstacle, a new one representing, i.e., a reflected or transmitted sound wave is created. Beyond the geometrical properties, each beam carries information about the acoustic phenomenon in whose result it has been created.

Acoustic phenomena affect the frequency response of the sound wave. Analysis and synthesis of an accurate spectrum for each of them requires complex analytical derivations. Such solution is unacceptable in the introduced simulation model. First of all, the authors do not have enough amount of data describing physical properties of materials. Moreover, accurate calculation of full frequency response of each phenomenon would require a lot of computational power, but the authors can not accept such an approach. An audio engine is only a part of really complex software such as video game, which consumes relevant amount of computational and memory resources, i.e., graphics, physics, artificial intelligence, etc. For our needs the model of the acoustic phenomena has to be as simple as possible and capable both to calculate in real-time and provide good experiences for listeners.

In the model applied in the developed audio engine, the acoustic phenomena are represented by data structures called *effects*. An *effect* is an approximation of the full frequency response of a particular phenomenon. It is a vector of eight coefficients – represented as floating point numbers of single precision – describing modification of signal amplitude in octave bands, where centre frequencies are equal to: 125, 250, 500, 1k, 2k, 4k, 8k, and 16k Hz. Each value is in the range from 0 up to 1. Value 1 means that frequencies in the given band are not modified, and for 0 all frequencies are muffed. Such a representation allows the authors to save computational power and memory usage, as one *effect* requires only 32 bytes of memory. There are also more reasons for using this simplified model than the one mentioned above. The crucial thinks is availability of data describing acoustical properties of materials, i.e., the reverberant absorption coefficient.

2.1. Reflection of the sound wave

The first phase of reflection phenomenon modelling takes place during beam calculations process. A reflected beam is always created using the Snell's law known from optics. The reflection edge is equal to the edge of an incident wave. The reflected beam starts on the reflecting plane and its image-source is reflection of a source of its parent beam on the reflection plane (Funkhouser *et al.*, 2002).

The second stage is an assignment of reflection *effect* for the created reflected beam. The material, which a reflection plane is covered with is known, thus, each material has its own reflection and transmission *effects*. They remain constant, therefore they can be created once at the phase of pre-calculations while loading a scene. Frequency response of a given phenomenon depends on the known properties of a specific material – the reverberant absorption coefficient and density, etc.

Although an approximation of the frequency response consists of only eight coefficients, there is still an issue to find appropriate values for each of them. For most utilized materials, there is no sufficient amount of data. The reverberant absorption coefficient α is usually measured for six octave bands in the range from 125 Hz up to 4 kHz (ISO, 2005). Extrapolation of the reverberant absorption coefficient is needed for this aim. For most of materials, the reverberant absorption coefficient α grows with the frequency, but this must not be treated as a rule. Choosing a right method of interpolation remains still a crucial issue, since it depends strongly on the characteristics of a particular material.

Our first approach was to extrapolate them linearly. The values are limited to 0–1. For bands lower than 125 Hz there are additional limits – the absorption coefficient cannot be higher than for a higher band. This assumption gave acceptable results in hearing tests.

When the absorption coefficient is set for all required bands, reflection coefficients can be chosen based on

$$|R|^2 = \frac{I_R}{I_I}, \qquad |R|^2 + \alpha = 1, \tag{1}$$

where I_R is the intensity of the reflected wave, and I_I is the intensity of the incident wave. The actual values for both I_I and I_R are not necessary. All that is needed is the value of proportion between these intensities, therefore it was assumed, that I_I is always equal to 1. Intensity of the plane wave is proportional to the square of the acoustic pressure (p), so the first equation of (1) can be transformed for the plane wave to:

$$|R|^2 \cong \frac{p_R^2}{p_I^2}, \tag{2}$$

where p_R and p_I are magnitudes of the acoustic pressure for the reflected and incident plane waves.

All sound sources in the introduced simulation model emit spherical sound waves, but models of reflections are appropriate only for plane waves. For this reason the authors have to assume that all modelled reflections take place in a far field, where the spherical wave can be considered as a plane wave (Makarewicz, 2009). In most cases this condition is fulfilled, because the distance between each sound source and its nearest obstacle (r) is usually larger than the far field boundary r_b defined by Eq. (3)

$$r_b = \frac{\lambda}{2\pi}. \tag{3}$$

Equations (4) and (5) present the relationship between the reverberant absorption coefficient (α) and reflection coefficient $|R|$ for the acoustic plane wave.

$$\alpha = 1 - |R|^2, \tag{4}$$

$$|R| = \sqrt{1-\alpha}. \tag{5}$$

A variety of approximate absorption coefficients can be found in literature. The most common are the normal and random incidence absorption coefficient and Sabine absorption coefficient (Kuttruff, 2009). The value of $|R|$ coefficient is between 0 and 1. Therefore a full reflection takes place if $|R| = 1$ and there is no reflection wave for $|R| = 0$. The reflection coefficient is calculated for each frequency band in *effect*. The approximations of frequency responses of the reflection phenomenon for three different materials are shown in Fig. 1.

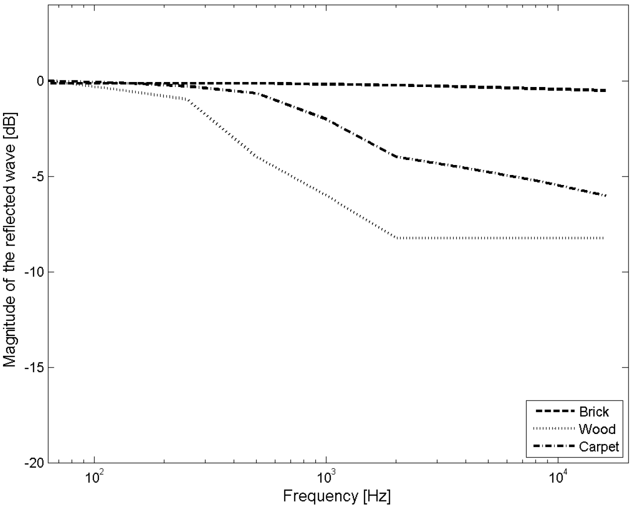

Fig. 1. Acoustic wave reflection from different materials.

In real cases, the value of the reflection coefficient also depends on other factors, for example, the angle of incident wave and porosity of the material (Kuttruff, 2009). These additional factors are difficult to be modelled due to lack of necessary data. Moreover, as mentioned before, in this application a lot of simplifications have to be assumed. Detailed modelling with the additional mentioned factors is impossible because of performance reasons, therefore, for example, dispersion, which occurs with reflection is neglected. In fact only specular reflections without diffusion are currently considered.

2.2. Sound transmission

Transmission of sound through obstacles is another effect which depends on the material features of the obstacle cover. Similarly to implementation of the reflection, transmission effect is read from the material and assigned to a new beam representing the transmitted sound wave. A beam created as a result of transmission through a wall has its source at the same point as the parent beam.

A relevant issue related to the transmission phenomenon is a refraction caused by changes in the properties of a medium. Refraction was skipped for practical reasons. First, the chosen method of modelling the sound via beams is already complicated enough for real-time calculations. It is also difficult to obtain an accurate model of this phenomenon.

The last, most important and really practical reason is the fact, that the listener will never be inside the wall. The coefficients saved in the structure of the effect representing a transmission of the sound on an obstacle are the ratio of the amplitude of the pressure of the acoustic wave behind the obstacle to the amplitude of the pressure of the incident wave. Our approach is based on a method of choosing indexes of acoustic insulation of a wall (Hassan, 2009).

A template method also called Sig Ingemansson's method, based on empirically chosen values is often used for buildings calculations (Hassan, 2009):

1. Calculate the critical frequency
$$f_c = 6 \cdot 10^4 \frac{\sqrt{\varrho}}{h\sqrt{E}}, \tag{6}$$
where E is the Young module, ϱ is the material density, h is the thickness of the wall.
2. Calculate the plateau height, R_p [dB]. In general, for a single wall, this is expressed as
$$R_p = 20 \log M_s + 20 \log f_c - 58.5, \tag{7}$$
where M_s is the surface density calculated as ϱh.
3. Calculate the value of the sound reduction coefficient for frequencies below $0.3 f_c$ using a decrease of 6 dB per octave.
4. Assume $R_p + 3$ dB for frequency $f_c = 0.6$ (point b in Fig. 2).
5. Assume the value of the sound reduction index as $25 + 10 \log \eta$ for the frequency of $2f_c$ (point d in Fig. 2).
6. Calculate the sound reduction index above $2f_c$ (point d in Fig. 2) using an increase of 7.5 dB per octave.

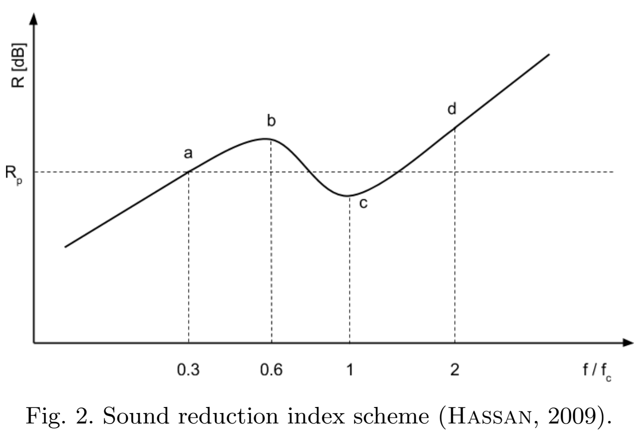

Fig. 2. Sound reduction index scheme (Hassan, 2009).

7. Approximation in the coincidence region. Calculate the value of R index for the critical frequency f_c (point c). Assume that:
   - $R = R_p - 5$ if $f_c \geq 200$ Hz,
   - $R = R_p - 4$ if 160 Hz $\geq f_c \geq 200$ Hz,
   - $R = R_p - 3$ if 125 Hz $\geq f_c \geq 160$ Hz,
   - $R = R_p - 2$ if 100 Hz $\geq f_c \geq 125$ Hz.

It is possible to calculate all coefficients of the transmission effect $T = p_T/p_I$, based on the values of the reduction index R, where p_T is the amplitude of a wave being the result of transmission and p_I is the amplitude of the incident wave. Similarly to the reflection effect, we do not need to know the values of p_T and p_I, but only the proportion between them, which is equal to:

$$T = \frac{p_T}{p_I} = 10^{-R/20}. \tag{8}$$

Please note that R has now a completely different meaning than the R coefficient in the case of reflection. In the case of transmission *effect*, the value of the reduction index R is given in decibels, therefore it has to be expressed in a linear scale (8). The examples are presented in Fig. 3.

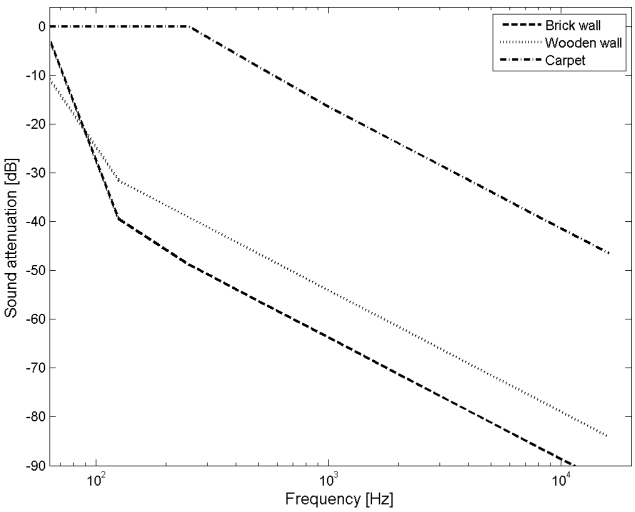

Fig. 3. Sound wave transmission within different materials.

2.3. Air absorption

An air absorption effect simulates changes of amplitude and frequency characteristics of the acoustic wave of passing a given distance. That effect is not related to any of the beams, so it cannot be calculated during the phase of geometrical calculations. It has to be computed dynamically for each path which reaches the listener. The amplitude of the sound wave can be presented as a function of a distance (Kulowski, 2007)

$$I(r) = I(r_0)e^{-m(r-r_0)}, \tag{9}$$

where m is the coefficient of the sound absorption by air. The values of m for chosen frequency bands are presented in Table 1 (Kulowski, 2007).

Table 1. Absorption coefficient m within air with the relative humidity 50% and temperature 20°C (Kulowski, 2007).

Frequency [Hz]	m [m^{-1}]
250	0.00009
500	0.00025
1 k	0.0008
2 k	0.0025
4 k	0.007
8 k	0.02

Table 1 has no air absorption coefficients for octaves lower than 250 Hz and for 16 kHz. Similarly to α, extrapolation is needed to represent the effect. The linear extrapolation does not reflect properties of this phenomenon. The coefficients are calculated according to (Long, 2006), using the equation:

$$m \cong \frac{1.7 \cdot 10^{-8} f^2}{\varphi}, \tag{10}$$

where f is the frequency and φ is the relative humidity expressed in %.

Table 2 presents supplementary coefficients m for the mentioned bands.

Table 2. Approximations of absorption coefficient m extended for other frequency ranges.

Frequency [Hz]	m [m^{-1}]
125	0.000005
16 k	0.087

The authors maintain that it is not necessary to model air absorption for distances lower than 1 m. The ratio of the amplitude of the acoustic pressure of a wave at distance r to the amplitude of the acoustic pressure at a reference distance $r_0 = 1$ m is important for calculating the effect of air absorption. The authors assume that $I(r_0) = 1$. Such a method assumes that the impact of air on frequency characteristics of the sound wave is calculated properly for large distances, where a spherical wave can be treated as a plane wave. In such cases, the intensity of a plane wave is proportional to a square of acoustic pressure (12). In the result of such assumptions, the relation (9) is transformed to

$$A = \frac{p(r)}{p(r_0)} = \sqrt{\frac{I(r)}{I(r_0)}} = \sqrt{e^{-m(r-1)}}, \tag{11}$$

where A stands for the amplitude of an effect of a single band. The value of m depends on frequency (Tables 1 and 2). This is why it is possible to designate a set of coefficients for an *effect* representing a phenomenon of air absorption of the sound.

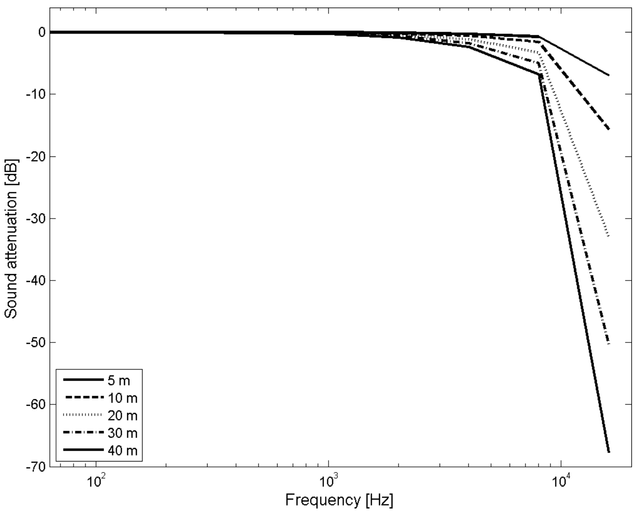

Fig. 4. Examples of frequency characteristics of air absorption phenomena.

The authors are aware that several important issues were neglected, for example, variability of humidity and temperature of the air. They have important impact not only on m but also on the speed of acoustic wave propagation. However, it should be stressed that the presented solution is dedicated for computer games audio engine. It is very unlikely that a player will follow changes in spectral characteristics of sound waves as a result of weather changes. However, if this feature is needed by sound designers, it will be implemented according to (10) in a future release.

2.4. Attenuation with distance

The amplitude of the acoustic pressure for a spherical wave depends on the inverse distance from the source. The intensity is inversely proportional to r^2, where r is the distance between the source and the receiver

$$I(r) = I(r_0)\frac{{r_0}^2}{r^2}. \tag{12}$$

The relation of changes of the amplitude of the acoustic pressure of a spherical wave is inversely proportional to the distance from the emitting source.

$$p(r) = p(r_0)\frac{r_0}{r}. \tag{13}$$

It was assumed, that the drop of intensity in a function of a distance is modelled by (9). Its confrontation with (13) shows differences between a theoretical mathematical notation and results obtained in the experiments. Equation (9) proved our engine to be much more successful. However, it was necessary to introduce an additional scaling factor. The factor depends on the distance between the sound source and the listener, characteristics of the sound, its significance, and the desired effect. It is caused by a need of adapting a method of sound decreasing to a particular scene. $A(r)$ is the amplitude of a sound sample signal at a distance r expressed in meters. The above factors are neglected if $r \leq 1$ m. The following solutions were implemented:

- lack of an additional scaling factor (always equal to 1) – for quiet, ambient sounds;
- $A(r) = 1/r$ – applies in most cases, used by default;
- $A(r) = 1/r^2$ – for impulse sounds, which should be audible only close to the sound source.

Obviously, apart from the second one, they are not proper with respect to the physics laws describing changes of wave pressure. However, film and computer games industries use them frequently. Following the suggestions of a sound designer from a cooperating company, the authors decided to implement them as well.

3. Representing the phenomena as digital filters

As it was mentioned in the previous section, the introduced models of the acoustic phenomena are represented by vectors of eight floating point numbers. To apply the obtained models to audio signals assigned to particular sound sources, a few digital signal processing operations have to be performed. The most trivial operation is scaling the signal by a factor responsible for sound attenuation for a given distance between the source and the receiver. Modification of the frequency characteristics of the audio signal is obtained by filtering using digital filters. The authors have considered two most commonly used approaches – the finite impulse response (FIR) and the infinite impulse response (IIR) filters. Here the first approach is presented.

The main reason is the fact that the FIR filters are easier to design than the IIR ones. Moreover, their stability is always guaranteed by their mathematical model because of lack of the poles. We decided to use the FIR filters of the order 512. Such an approach allows to reach a 93 Hz frequency resolution between spectral bars for the sampling frequency equal to 48000 Hz. The design process of filters is performed in the frequency domain where the amplitude of each frequency between 0 and 24000 Hz is calculated by a linear interpolation of eight coefficients. Only the amplitude spectrum is modelled, the phase for each frequency is equal to 0. According to the properties of the spectrum of the digital signal, the second part of the amplitude spectrum is symmetrical to the first obtained part.

A transmittance of each sound path between the source and the receiver is modelled as a FIR filter. Each path contains a medium attenuation *effect* and reflection and/or transmission *effects* for indirect sounds.

The final *effect* is obtained by element-wise multiplication of all eight-element vectors and then multiplied by the distance attenuation factor. A final FIR filter is built using coefficients from the obtained final effect. The delay (D) connected with the length of the sound path expressed in samples can be calculated using equation (14)

$$D = \frac{r}{c} F_s, \tag{14}$$

where r is the length of the sound path (in meters), c is the sound speed in a medium (m/s) and F_s is the sampling frequency (in Hz). In the case of a FIR filter of order 512, the delay for the output signal is about 10 milliseconds.

4. Efficiency

Filtration of the audio signal by a FIR filter by definition can be done in the time domain by linear convolution, but complexity $O(n^2)$ of such an approach is unacceptable here. In this case dozens or hundreds audio streams with the sampling rate of 48 kHz are filtered simultaneously in real-time. In order to omit an unacceptable computational complexity of linear convolution, the authors utilize the properties of a circular convolution known as *fast convolution*. An efficient implementation of the Fast Fourier Transform algorithm allows to obtain a significantly lower complexity. Thanks to this method it is possible to process about two hundred of audio streams in real-time, utilizing only a part of the CPU's power.

Audio processing operations in the described library are divided into two stages. The first is related to calculations of the digital filters in the frequency domain for each found sound path during geometry scanning. Then each sound path represented as a stream of audio samples has to be filtered separately by the obtained FIR filters. The results presented below prove that the performance is sufficient to provide a real-time simulation of all the considered acoustic phenomena. The mean rendering time for one 30 ms sound frame is presented in Table 3. Most optimizations were made in the frame rendering procedures, thus the performance in the AVX mode is a bit more efficient than in the SSE one.

Table 3. Mean duration of preparation of a 30-ms long audio frame for a single audio stream on different architectures. DSP operations are performed using one thread.

	i5 with SSE	i5 with AVX	i7 with SSE
Filter calculations [µs]	59	60	89
Sound frame rendering [µs]	39.7	37.7	61
Rendering of 200 pipelines [ms]	19.75	19.54	30

The whole library has been built in the x64 mode using the Microsoft Visual C++ 11 compiler. The measurements were performed on two different machines running the 64-bit Windows 7 Professional operating system. The models of the mentioned processors are: Intel Core i5 2550OK and Intel Core i7 860. For the first processor (i5 from Intel's Sandy Bridge family introduced at the beginning of 2011) the full processing of the output audio frame takes by average less than 20 ms. Even the older Intel's Core i7 processor from 2009 (without Advanced Vector Extensions) processes two hundred audio streams in real time. The results prove that the presented approach works successfully on older machines.

4.1. Code statistics

The described library is written in C++ programming language. The current version consists of roughly 150 classes and above 65 thousand lines of the source code. Additionally, there are several smaller applications which provide the authors with diagnostic and visualisation tools. The module connected with acoustic phenomena modelling, finite impulse filters design, and digital audio processing takes about 25% of the code. All crucial parts contain complex optimizations related to the memory layout and SIMD using the SSE and AVX instructions.

5. Evaluation of results

In the previous sections the authors mention that the applied solutions give acceptable hearing experience which indeed is subjective. Defining an objective measure of the obtained results is possible, however, this will not give any useful information in this case. It is determined by the specific features of the described application. Computer games may – and in authors' opinion also should be – a kind of art and it is really difficult to find objective criteria for assessment. Writing about sufficiently good experiences, the authors mean subjective opinions of more than ten professional sound designers working in the computer game industry.

The simplest and the most robust method of evaluation of the obtained results with reality would be comparison of the obtained impulse response of a given modelled room with the result obtained in real measurements. Although it would give the authors valuable data allowing to verify how accurate the applied model is, they only will describe a final result. This will not help to improve models of reflection or transmission effect. The authors do not even expect that the presented results accurately represent the reality. There were a lot of simplifications applied. Without them, there would not be any possibility to achieve such good performance, which was much important than accuracy.

6. Conclusions

Acoustic phenomena modelling is a complex issue particularly when all physical properties of the sound wave have to be taken into account. Accurate modelling requires not only significant mathematical derivations and amount of computational resources. Equally important are accurate measurements of the results, which requires well-equipped laboratories.

Geometrical methods, though widely used in acoustic simulations, have to be treated carefully, particularly for low frequencies. In the presented case this does not have to be a disqualifying obstacle. As proved in this paper, less accurate, simplified acoustic modelling can provide acceptable results for specific cases. They suit well for real-time applications, i.e., computer games. Although the proposed methods do not fulfil a lot, always relevant physics laws, they provide sufficient good experiences for gamers, who pay attention not only to the sound, but also to other relevant aspects of the virtual reality.

Derivations presented in this article prove, that it is possible to simulate the acoustic properties of virtual scenes using many simplifications. Even though they sometimes neglect relevant physical laws, due to limited performance of CPU and memory, professional audio designers maintain that such a simplified solution can be successfully applied in the video game industry.

Acknowledgments

This work was supported by the National Centre for Research and Development under INNOTECH-K1/IN1/50/159658/NCBR/12 grant. The authors would also like to thank the whole team working on RAYAV project, especially Tomek Pędzimąż, Szymon Pałka, Irek Gawlik, and Paulina Bugiel.

References

1. Bellmann J., Michel F., Deines E., Hering-Bertram M., Mohring J., Hagen H. (2008), *Sound tracing: Rendering listener specific acoustic room properties*, EG Computer Graphics Forum, **27**, 943–950.
2. Funkhouser T., Tsingos N., Carlbom I., Elko G., Sondhi M., West J. (2002), *Modeling Sound Reflection and Diffraction in Architectural Environments with Beam Tracing*, Forum Acusticum, page 8, Seville, Espagne.
3. Hassan O. (2009), *Building Acoustics and Vibration – Theory and Practice*, World Scientific Publishing Co. Pte. Ltd.
4. ISO, PN-EN ISO 354:2005P *Acoustics – Measurement of sound absorption in a reverberation room* [in Polish: *Akustyka – Pomiar pochłaniania dźwięku w komorze pogłosowej*].
5. Kamisiński T., Burkot M., Rubacha J., Brawata K. (2009), *Study of the effect of the orchestra pit on the acoustics of the Kraków opera hall*, Archives of Acoustics, **34**, 481–490.
6. Kulowski A. (2007), *Acoustics of halls* [in Polish: *Akustyka sal*], Wydawnictwo Politechniki Gdańskiej, Gdańsk.
7. Kuttruff H. (2009), *Room acoustics*, 5th ed., Spon Press.
8. Long M. (2006), *Architectural Acoustics*, Elsever Inc.
9. Makarewicz R. (2009), *Sounds and waves* [in Polish: *Dźwięki i fale*], Wydawnictwo Naukowe UAM, Poznań.
10. Ozgura E., Ozisb F., Alpkocaka A. (2004), *Daad: a new software for architectural acoustic design*, The 33rd International Congress and Exposition on Noise Control Engineering Internoise, Prague.
11. Pałka S., Głut B., Ziółko B. (2014), *Visibility determination in beam tracing with application to real-time sound simulation*, Computer Science, **15**, 2.