ABSTRACT
A programme that joins measurements and simulation tools for room acoustics is currently being developed at the University of São Paulo. The work presented here describes the implementation of two traditional simulation methods, based on geometrical acoustics: the Image Sources and Ray Tracing methods. Results from computer simulations performed with both are compared to measurements performed in an auditorium, a small reverberation room and the reference room used in the third Round Robin on room acoustical computer simulation. The limitations and advantages of each method are analysed through the comparison of Impulse Responses, decay curves and room acoustical parameters. Beyond that, the implementation of an auralisation tool is described, including a comparison of results obtained through the use of different convolution algorithms. Its use is exemplified by a preliminary study on sound source localisation.

INTRODUCTION
Since the institutional introduction of acoustics as a teaching subject in Brazilian universities, such as the Federal University of Santa Catarina and Federal University of Rio de Janeiro, in the 1970's, this field of knowledge has gained specialised professionals in the country and experienced a growth of interest. Room acoustics, in particular, has attracted attention from diverse directions, including academic groups related to contemporary music composition. This interest may be particularly inspired by research and development performed in institutes like IRCAM (“Institut de Recherche et Coordination Acoustique/Musique”). Not only issues related to the physics of sound propagation in real rooms, but also related to possibilities created by the application of recent technologies in virtual environments and situations are part of the universe explored by composers and musicians.

The spatial information of a sound, for instance, is an element which may be consciously introduced in a music piece [1,2]. In this case, auralisation tools may be used in the composition process and in its performance. Within this context, a group of scientists from University of São Paulo, supported by the State of São Paulo Research Foundation (FAPESP) started the AcMus project in 2001. The goal was to develop an open source platform for measurement, simulation and analysis of room acoustics [3]. In 2003 this author joined the project with the goal of developing the simulation programmes.

Two methods were implemented so far, namely the Ray Tracing Method and the Image Sources Method. Both implementations are briefly described in this paper, so as an auralisation module. They are intended to be used not only as a tool for the acoustical design of rooms, but also as an experimental platform, where models related to sound source, receiver, diffraction and reflection can be tested. Beyond that, the programmes should be used to generate samples for psychoacoustics experiments.

SIMULATION METHODS
Ray Tracing
This well known method was implemented basically as described in [4]. Differences are related to the models for sound source, receiver and reflections. Instead of using a random algorithm for generation of rays’ directions, a routine for geodesic division of a regular icosahedron was implemented, as described in [5] and [6].

Two variations of a spherical receiver were implemented (in the same way as explained in [6]). One calculates the ray energy in proportion to its intersection length with the sphere. The other simply divides the energy by the maximal cross sectional area of the sphere.

The model implemented for reflection takes absorption and scattering properties into account. Once a ray strikes a surface, a random number between 0 and 1 is generated. If this number is greater than the scattering coefficient attributed to the surface (input data), the reflection occurs symmetrically in relation to the surface’s normal direction. Otherwise the ray is reflected in a random direction.

Image Sources
The traditional Image Sources Method was implemented, as explained in [7]. A single difference refers to the computation of the energy from each sound source at the listener’s position, which in our case is given by:

\[
E_i = \frac{R_0^2}{R_i^2} e^{m(R_0-R_i)} \prod_{j \in S} (1 - \alpha_j),
\]  

(Eq. 1)

where \( R_0 \) is the distance from the listener to the original sound source, \( R_i \) is the distance from the listener to the \( i \)th image source, \( m \) is the attenuation constant for sound propagating in the air, \( S \) is the set of surfaces used in the generation of the image source and \( \alpha_j \) is the absorption coefficient attributed to the \( j \)th surface.

AURALISATION MODULE
Both simulation programmes provide an output file which may be used to construct binaural impulse responses. The first step for that is to organise an impulse response matrix \( (IR) \), where directional information is stored in its lines, i.e., every line is a simulated impulse response for sounds that strike the listener from a single direction. Than, a Head-Related Impulse Response \( (HRIR) \) matrix is constructed, and every line also contains information related to a single direction. In order to achieve the binaural impulse response, each line from \( IR \) must be convolved with the correspondent line of \( HRIR \). The resulting signals from each line must be added at the end, to obtain the impulse response for the left or right ear.

Two convolution algorithms in the time domain and two in the frequency domain were implemented, with the goal to compare their efficiency. In the time domain, the convolution integral was implemented “as it is”. However, because the lines from \( IR \) may be sparse, a second method was implemented, which uses the following property related to the convolution integral:

\[
\delta(t-t_0) \ast s(t) = s(t-t_0),
\]  

(Eq. 2)

where \( \delta(t) \) represents the Dirac Delta function and “\( \ast \)” represents the convolution operation.

Through the use of this property, the convolution operation is performed simply by including the \( m \)th line in \( HRIR \) in the position of the non-zero component in the \( m \)th line in \( IR \), multiplied by its magnitude. If there are two adjacent components, the resulting values are summed. In the frequency domain, FFT-based algorithms were implemented. One of them represents a “traditional” FFT algorithm, while another one used a C library named FFTW (after [8]), which is optimised. The convolution operation is used again to obtain auralised signals from anechoic recordings. In both steps (to obtain binaural impulse response and to auralise anechoic audio signal) the user may choose which convolution method will be applied. Two programmes were created, for different HRIR data. One of them uses data from the Institute of Technical Acoustics from Aachen University, while the other one uses KEMAR/MIT data, also with the option to select ear’s size [9].
EXEMPLARY RESULTS FROM ACOUSTICAL SIMULATIONS

Reverberation Room
A 72 m$^3$ reverberant room was simulated, and the results compared to measured curves and values. The walls are mainly made of painted concrete blocks, the ceiling is painted concrete and the floor is ceramic. 7.8 m$^2$ of the floor were covered by absorptive foam. Absorption coefficients assumed for each material at 1 kHz are: 0.07 for painted concrete blocks, 0.5 for the floor (ceramic tiles + absorptive foam), 0.02 for painted concrete.

The Schroeder frequency estimated for the room under the described condition is 130 Hz. Figure 1 shows the decay curves obtained from the simulated and measured impulse response at 1 kHz (for the same sound source and receiver positions). Two variations of spherical receiver were tested, as explained (receiver type 1 and receiver type 2 in the figure). Scattering was not taken into account.

![Figure 1.- Measured and simulated decay curves for reverberant room (1000 Hz).](image1)

![Figure 2.- Detail of measured and simulated impulse response.](image2)

The potential of the Ray Tracing Method to predict the overall behaviour of the sound decay is clearly seen in the figure. As known, the Images Sources Method has limitations...
related to computation time, which grows exponentially with the reflection order. In this case, a reflection order of 12 was not enough to reasonably predict the sound decay. From estimations based on [10], a reflection order above 20 would be necessary. An advantage of the Image Sources Method would be its capacity to predict more precisely the distribution of sound reflections in the impulse response. In impulse responses obtained through Ray Tracing the time resolution is limited by the diameter of the sound receiver. A detail of the impulse response obtained with Image Sources is presented in Figure 2, compared to a measured one. The result is not identical, but the potential ability of the method to represent the distribution of first reflections is illustrated.

Auditorium
The auditorium from the School of Civil Engineering at the University of Campinas was also used for comparisons between measured and simulated results in the context of this work. Its volume is approximately 690 m$^3$ and the geometry can be described by a relatively simple concave polyhedron. There are no specific elements to provide diffuse reflections and the audience area is where sound reflections may assume more complex patterns. In the model, this area was approximated by a plane with an equivalent absorption coefficient. Simulations were performed considering (0.6 at 1 kHz) and not considering a scattering coefficient for this surface. An additional simulation was performed, where a scattering coefficient of 0.05 was attributed for all surfaces, but the audience area (for which the 0.6 value was maintained). Figure 3 shows the decay curves obtained from measurement (filtered in 1 kHz octave band) and simulations. Results obtained with only one type of receiver are shown, since no significant differences were observed. The simulation with Image Sources considered 12 reflection order.

![Figure 3. Measured and simulated decay curves for auditorium (1000 Hz).](image)

When considering scattering coefficients, the tendency is that the curve becomes closer to the measured one, especially in the range from 0 to -20 dB. Although the difference is small, one clearly observes how EDT comes closer to the measured value (see Table I). In the Image Sources simulation, reflection order around 60 should be considered for a reliable simulation, according to estimations based on the expected decay.

| Table I.- Measured and simulated parameters (Auditorium, 1 kHz). |
|---|---|---|---|---|
| Measured & Ray Tracing & Ray Tracing & Ray Tracing & Image |
| Specular | Scattering in audience | Scattering all surfaces | Sources |
| EDT | 1.16 | 1.03 | 1.19 | 1.22 | 0.81 |
| $T_{20}$ | 1.21 | 1.33 | 1.45 | 1.28 | 0.82 |
| $D_{50}$ | 0.40 | 0.54 | 0.52 | 0.50 | 0.56 |
Reference room from the 3rd Round Robin on room acoustic computer simulation

Data from the first phase of Round Robin III [11] were used to check the quality of the implemented methods. In this phase, scattering was not taken into account. The Image Sources Method was not tested in this case, because of the limitation in the maximal reflection order. Comparative results for the 1 kHz frequency band, and position S2R2 are shown in Table II.

Table II.- Parameters simulated in Round Robin III (S2R2, 1 kHz). Mean values and standard deviations from 21 different programmes are shown.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Ray Tracing</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDT</td>
<td>1.8</td>
<td>0.14</td>
<td>1.7</td>
</tr>
<tr>
<td>T30</td>
<td>1.8</td>
<td>0.16</td>
<td>1.7</td>
</tr>
<tr>
<td>D50</td>
<td>34.7</td>
<td>2.66</td>
<td>34.3</td>
</tr>
</tbody>
</table>

PERFORMANCE AND USE OF AURALISATION MODULE

The convolution algorithms used in the auralisation module were compared in relation to computation time and quality of results. Regarding the latter, differences in results were not relevant. However, the computation time required to obtain a binaural impulse response may vary substantially, according to the convolution method used.

Figure 4 presents the time needed to obtain a binaural impulse response as a function of the number of non-zero components in a simulated impulse response. Only results obtained through the method which uses the property described by Eq.4 and with the FFTW algorithm are shown. The time required to compute the binaural responses through "traditional" convolution methods, in time and frequency domain, was at least 120 s.

![Figure 4.- Time necessary to compute the binaural impulse response with different convolution algorithms.](image)

A preliminary study on sound source localisation was performed with sound samples created with the image sources programme and the auralisation module. The virtual “rectangular” room created had a short reverberation time, so that the room simulation was reliable and not limited by the required computation time. The programme which uses HRIR data from the Institute of Technical Acoustics from Aachen University was used in this case.

The virtual listener position was set still in the study, and speech was presented through headphones. The virtual speaker could be in six different positions, divided in two sets: 1) three in front and to the right of the listener's head and 2) three behind and to the left. For each set of speaker position, the angular differences between them varied in 5, 10 or 15 degrees.
Two levels of precision in localisation ability were investigated. At first, the subjects were asked to localise the virtual speaker, in general terms, by saying: “he is behind me” or “he is in front of me, on my right”. Later, the subjects were asked to compare two sound samples and tell their relative position (“Listener A stands on the left of Listener B”). Complete results are not presented here.

CONCLUSIONS AND FINAL COMMENTS
As observed in the results from room simulations, the Ray Tracing algorithm is reliable for determining decay curves and room acoustical parameters. The use of a model for reflection which considers diffusion is important for its reliability (even a simple one), although a scattering coefficient data base from common surfaces is still not available. When the goal is to create virtual environments, there is a need to find exactly what features are important. The time resolution existing in Ray Tracing results may be not precise enough for creating a realistic impression and this is where the use of the Image Sources becomes more important. However, only a hybrid approach would provide the possibility of simulating more complex rooms with this method. The simulation methods presented here will now be implemented in the AcMus platform, in a hybrid version. Psychoacoustic experiments using the programmes described here and the auralisation module will be used for different purposes.

ACKNOWLEDGEMENTS
The author is grateful to the State of São Paulo Research Foundation (FAPESP) for supporting this work, the undergraduate students João Gorenstein Dedeca and Rhenan Giorgiani do Nascimento, who programmed and developed solutions in the auralisation module and Ray Tracing, Prof. Stelamaris Rolla Bertoll for receiving me at the School of Civil Engineering, Prof. Michael Vorlaender and Tobias Lentz for providing the HRIR data base and all people from the AcMus project at the University of São Paulo.

References: