



HAL
open science

Room acoustic auralization with Ambisonics

Jean-Dominique Polack, Fábio Leão Figueiredo

► **To cite this version:**

Jean-Dominique Polack, Fábio Leão Figueiredo. Room acoustic auralization with Ambisonics. *Acoustics 2012*, Apr 2012, Nantes, France. hal-00811357

HAL Id: hal-00811357

<https://hal.science/hal-00811357>

Submitted on 23 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ACOUSTICS 2012

Room acoustic auralization with Ambisonics

J.-D. Polack and F. Leão Figueiredo

LAM/IJLRA - Université Pierre et Marie, 11 rue de Lourmel, 75015 Paris, France
jean-dominique.polack@upmc.fr

During the year of 2009, the room acoustics group of the LAM (Équipe Lutheries, Acoustique, Musique de l'Institut Jean Le Rond d'Alembert - Université Pierre et Marie Curie, Paris) performed a series of acoustical measurements in music halls in Paris. The halls were chosen in regarding their importance to the historic, architectural or acoustic domains. The measured ensemble of fourteen rooms includes quite different architectural designs. The measurements were carried out with a Soundfield microphone, in order to afterward recreate the sampled sound field in the listening room at LAM. The presentation describes the tools used to realise the auralization, then moves on to the subjective tests realised with the system. Statistical analysis was carried out on the results of the subjective tests. The results draw insight into the qualities of auralization for reproducing sound field, but also on its limitations.

1 Introduction

During the year 2009, the room acoustics group at LAM (Équipe Lutheries-Acoustique-Musique, Institut Jean Le Rond d'Alembert, Université Pierre et Marie Curie, Paris) performed a series of acoustical measurements in concert halls and theatres in Paris. The halls and theatres were selected for their historical, architectural, or acoustic interest.

Statistical analysis of the measured acoustical indices is presented in another session of the present congress [9]. Therefore, this presentation focuses on auralization and the subjective tests set up to check the system.

2 Measuring equipment

2.1 Source

The measuring equipment consists of a dodecahedral sound source (*Outline GSR*), and a subwoofer (*Tannoy Power VS10*) connected to the source, both supplied with their amplifiers.

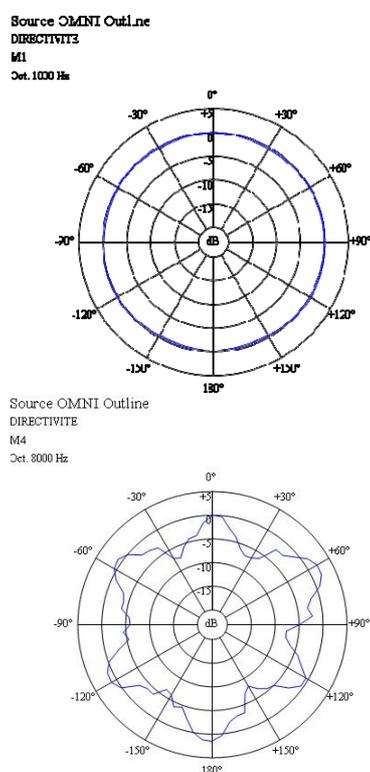


Figure 1: Polar responses of the source at 1kHz and 8kHz.

At its frequency of operation, the subwoofer radiation is omnidirectional, and so is the dodecahedron up to the 1kHz

octave, as depicted in Figure 1. However, at higher frequencies, the directivity departs from omnidirectionality, but variations remain within 5dB in the 8kHz octave band (Figure 1). No figure is given for the 16kHz octave.

2.2 Microphone

All measurements were carried out with an Ambisonics SoundField ST 250 microphone, connected to a multichannel soundcard driven by a laptop.

The Soundfield microphone contains four sub-cardioid capsules mounted in a tetrahedral arrangement. By combining the output of the four capsules, a pressure microphone and three gradient microphones, at right angles from each other, can be reconstructed. This four-channel signal is known as the Ambisonics B-Format.

Figure 2 presents the pressure responses of the pressure microphone (upper trace) and the gradient microphones (lower trace) reconstructed from the SoundField ST250. The omnidirectional response is constant within 1dB from 60Hz to 4kHz, and the figure-of-eight response with ± 1 dB within the same range, extending in fact up to 2kHz.

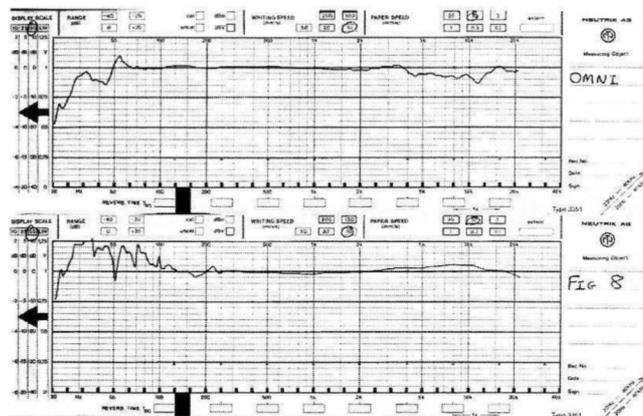


Figure 2: Responses of ST250 microphone.

2.3 Signal

An exponential sweep-sine signal was used as original signal, because it allows *a posteriori* elimination of harmonic distortions from the sound source, as well as efficient signal-to-noise ratio [4,7]. It was recorded and processed with the Aurora plug-ins, developed by Angelo Farina from Parma University.

The sweep sine signal is generated 20 Hz up to 20 kHz in 30 seconds. A relatively long duration was selected because the signal-to-noise ratio is proportional to the sweep time.

Figure 3 presents the spectrum of the sweep signal radiated in the large anechoic chamber at LNE (upper trace) together with the spectrum of the compensated sweep

signal (intermediate trace). Evident in Figure 3 is the fact that compensation allows rectifying the signal over a large band, from 60Hz to 5.5kHz, that is sufficient for the usual acoustical indices [6]. However, post processing makes it possible to further extend the bandwidth from 40Hz to 18kHz, at the cost of a light reduction in the level (lower trace in Figure 3). This extra bandwidth is necessary for the auralisations.



Figure 3: Original spectrum of the signal and the two steps of compensation.

For all measurements, the compensated signal was radiated in the halls. After recording and computation of the impulse responses, post processing was applied to the four channels of the B-format response and further processed for the auralization.

3 Measurement protocol

3.1 The 14halls

The halls were selected for their historical, as well as architectural and acoustic interests.

Table 1: The 14 halls.

	Volume (m3)	Seats	Abbr.
Théâtre des Abbesses	4500	396	ABE
Théâtre de l'Athénée	3366		ATH
Opéra Bastille	26000		BAS
Chapelle Royale de Versailles	14400		CHP
Théâtre du Châtelet		2300	CHT
Cité de la Musique	13400	1200	CIT
Salle Cortot	2580	400	COR
Opéra Garnier			GAR
Maison de la Culture du Japon	6300	400	JAP
Auditorium du Louvre	4500		LOU
Théâtre de la Porte St. Martin		1000	MAR
Auditorium du Musée d'Orsay	1700	347	ORS
Salle Pleyel	17800		PLE
Maison de Radio France	10000		RAD
Théâtre du Ranelagh	1920		RAN
Théâtre de la Ville	5120	1012	VIL

As our goal is not to evaluate acoustical excellence, but rather to develop a typology of halls based on acoustical

criteria, we were looking for a representative set of halls with broad ranges of such characteristics as: volume, form, wall materials, number of seats, and artistic usage.

Table 1 lists the 14 halls selected for the campaign, together with their volumes and numbers of seats. It also indicates the abbreviations used to refer to them.

3.2 Positions

In each hall, ten microphone positions were selected (except for the smaller rooms, as indicated in ISO standard 3382), trying to preserve a standard distribution of positions while respecting the physical possibilities of the rooms. Microphone positions were therefore selected according to the following scheme:

- Positions 'a', 'b' and 'c' on the central longitudinal axis ('a' nearest and 'c' furthest from stage).
- Positions 'd' and 'e' on lateral longitudinal axis ('d' nearer and 'c' further from stage).
- Positions 'f' and 'h' on central longitudinal axis, first and second balcony respectively.
- Positions 'g' and 'j' on lateral first and second balcony, respectively.

Other positions were used occasionally, depending on architectural specificities of the rooms. As for the source, it was positioned on the centre of the stage, or on its left and right. The last two source positions allow for auralization with stereophonic recordings.

At all these positions, impulses responses were measured in order to derive the traditional set of indices [1,6]. These indices, together with their statistical analysis, are presented in another session of the present congress [9]. Impulse responses were also measured in order to carryout auralizations in a listening room, as described in the remaining of this paper. Consequently, a calibrated Pink Noise signal was also played through the source and recorded at each measuring position in order to adjust the reproduced sound level for auralizations.

4 Auralization

Auralisation is carried out in two steps: convolution; and Ambisonics decoding. Convolution of the impulse responses with an anechoic musical excerpt aims at recreating the impression of listening to the excerpt as if it was played in the room where the impulse responses was recorded. Since the B-format impulse responses are audio files with 4 channels, the convolution tool must support this format. We selected *Voxengo Pristine Space*, that enables multichannel convolutions in real-time.

The Ambisonics decoder receives the 4 signals and distributes them to the 12 loudspeakers of the listening room according to the Ambisonics protocol. We retained the *Decopro* decoder, which enables to insert the coordinates of all the loudspeakers so that the decoder itself corrects the differences in arrival times from each loudspeaker at the listener's position. The listener is placed at the centre of the room, since Ambisonics has a narrow spot [2]. Sound level corrections due to the geometrical irregularities of the system, however, are made by the user: level at the listener's position must be the same for all the

loudspeakers. *Decopro* and *Voxengo Pristine Space* are VST plug-ins, hosted in our case in *Audio Mulch*.

Once the configuration and the impulse responses are loaded in the decoder in the convolution tool, and before playing the excerpts for the auralization, sound levels must be adjusted to their values in the original rooms. Pink Noise was used for that purpose (see Section 3.2), and we adjusted the configuration of the system so that the noise level at the listening position in the listening room is equal to the level at the measuring position in the original room. Once set, the configuration was never changed. We repeated the same procedure for all the positions of measurement, and obtained a bank of 235 configurations corresponding to each position of measurement.

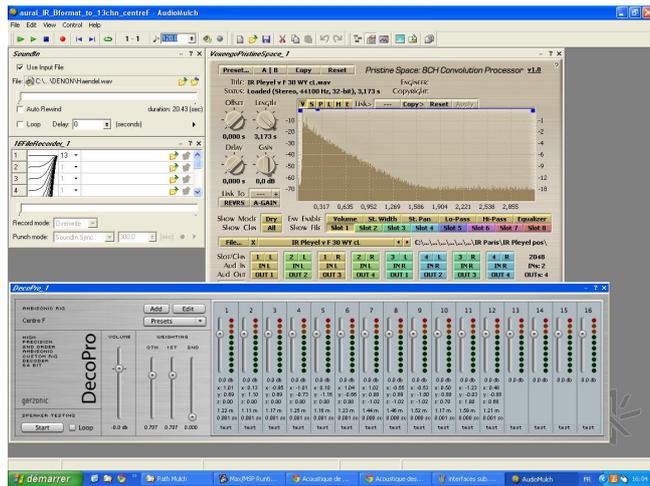


Figure 4: Interface for auralization

In the end, the decoder was fed with the anechoic excerpts and delivered for each position of measurement 13 convolved calibrated channels, which were recorded in a 13-channel audio file. The files were played at demand through a purpose-designed MAX/MSP interface in the test room at LAM. It is a very damped room of size 2.77x3.24x3.62 m built on a floating floor. It contains 12 “Studer A1” loudspeakers positioned in dodecahedral form and a “JBL 4645C” subwoofer. These loudspeakers are hidden behind visually opaque, but acoustically transparent, fabric panels and, so that listeners cannot see the loudspeakers (Figure 5).

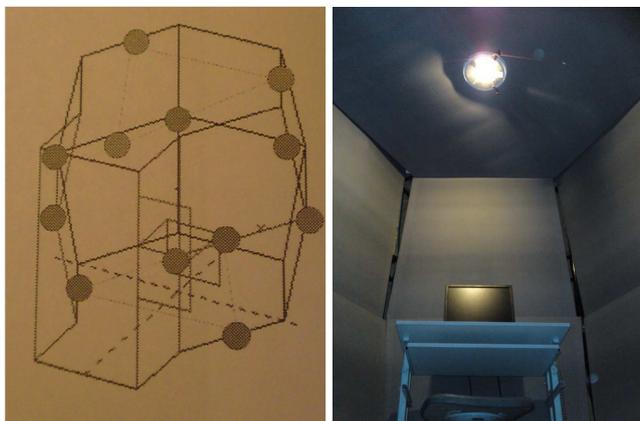


Figure 5: The listening room

The auralization hardware was composed of a PC of last generation, a DIGI 96 soundcard which played the digital

signal files to two RME ADI-8 Pro converters. These digital-to-analog converters sent in turn the analogical signals to the loudspeakers and the subwoofer.

5 Subjective tests

In order to evaluate the perceptual relevance of our database, we opted for a free categorisation test [5]. However, free categorisation of 235 configurations is unpractical, and we had to select a subset of the database. For practical reasons, a subset of 10 configurations was selected more or less at random, arbitrarily covering the different types of halls. So the first task was to check that the subset is representative of the whole database, and we used Principal Component Analysis (PCA) for this check.

5.1 Selecting the configurations

The ten configurations selected for auralization are listed in Table 2.

Table 2: The 10 configurations.

Séquence	Position
A	ABE b LR
B	ATH c LR
C	CHP b LR
D	CIT b LR
E	COR b LR
F	JAP ref c LR
G	LOU b LR
H	MAR c LR
I	ORS ref c LR
J	PLE b LR

Principal Components Analyses (PCA) was then carried out (Figure 6 and 7).

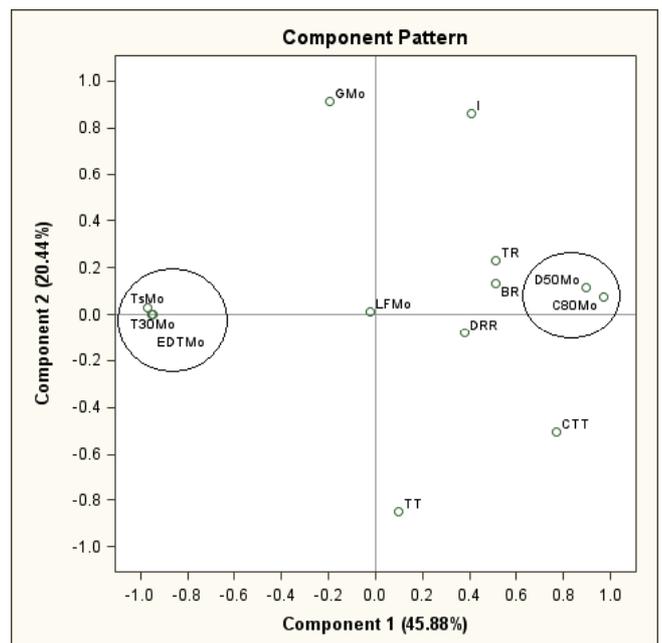


Figure 6: PCA for mean octave values of the indices

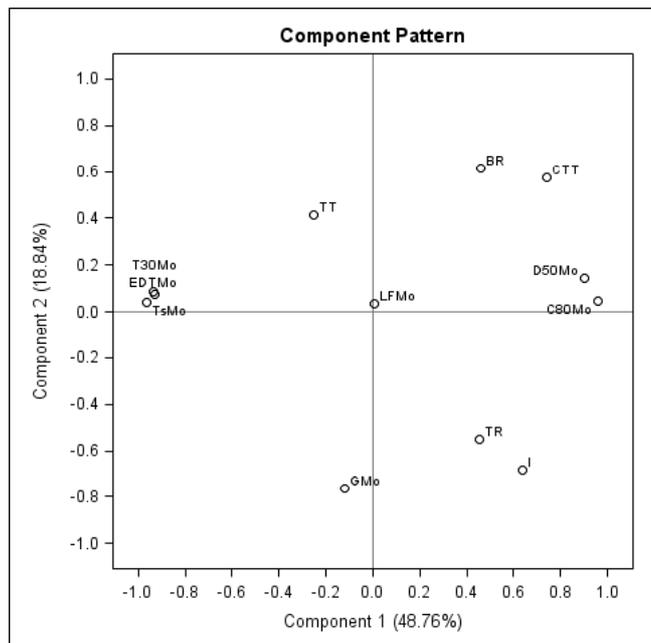


Figure 7: PCA for selected subset

They were carried out with the mean octave values of the indices previously used in the statistical analysis [9]. Figure 6 presents the plane built by the two principal components, when PCA is carried out for the full database, and Figure 7 the same plane for the selected subset of the 10 configurations of Table 2. Each of the 13 indices is represented in this plot by a point.

It can be seen in both Figure 6 and Figure 7 that the group formed by T30, EDT Ts and C80 strongly contributes to the first component. In both Figures, this first component accounts for roughly the same proportion of the variance of the data, 46% and 49% respectively. In a similar fashion, index G contributes to the second component, which respectively 20.5% and 19% of the variance. The only difference is that the second component is reversed in Figure 7. Similar results are obtained for the next 2 components, though with some rotations of the components, as can be seen when considering indices BR and TR in Figure 6 and 7. All in all, comparison of the two PCA proves that the selected subset is representative of the full database.

5.2 Categorisation

Subjects listen to a set of 10 sound sequences, the objects, corresponding to 10 different auralizations of the same anechoic excerpt, a 30s excerpt from Bruckner's Symphony no. 4 [3]. Subjects had to freely group together sequences that sound similar. They could listen to the sequences as many times as they wanted, and build as many groups as they wanted. Thus, each subject produced a partition of the set. 31 subjects participated to the test.

The data set to be analyzed is, therefore, a collection of partitions of the objects. From this collection, one builds a matrix of dissimilarities between the objects. The method of additive trees of similarity, proposed by Sattath and Tversky [10,8], allows to represent the structure of the objects in the shape of a tree (a set of nodes connected by edges). The objects correspond then to the leaves of the tree, and dissimilarity between two objects is represented

by the length of the path that joins them. The *AddTree* software, in the version of Barthelemy and Guénoche (1988), was used. It features a topological organization of the various groupings emerging from the individual data. From the tree, classes are obtained by cutting some branches, which correspond to categories under specific conditions. Applying the algorithm to the 31 partitions obtained in our test produced the tree of Figure 6.

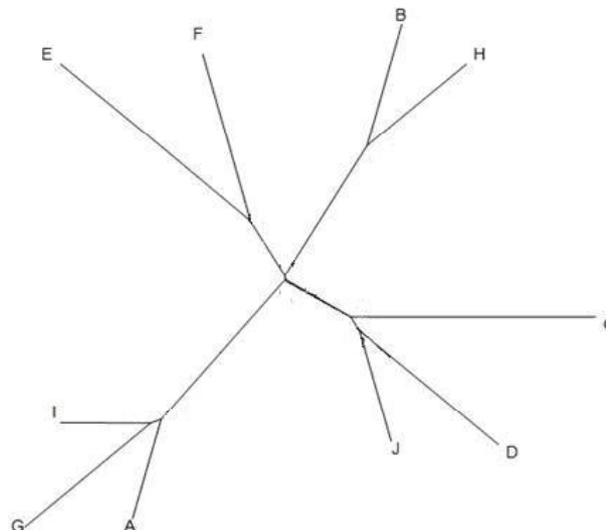


Figure 8: Categorisation tree.

Beyond representing the distances between the objects by the lengths of the edges, the tree of Figure 8 gives important information: it enables to identify four groups, or classes, of minimal dissimilarities, formed by objects BH, CDJ, ACTED and EFF. The corresponding identity of the objects is given in Table 3.

Table 3: Groups resulting from categorisation.

Groupe	Salle
BH	ATH, MAR
CDJ	CHP, CIT, PLE
AGI	ABE, LOU, ORS
EF	COR, JAP

These groups are similar to those obtained by cluster analysis [9]. The pair Athénée and St Martin were indeed clustered together, and so were Abbesses, Louvre and Orsay. However, Cortot and Japan also belonged to this second group, but are separated in the subjective test. Cité de la Musique and Pleyel, which belonged to a specific group "concert halls" in the cluster analysis, are here grouped with the Chapelle, certainly because of the higher reverberation time.

On the objective side, one can ask which acoustic indices underlie these subjective regroupings. Since grouping together corresponds to choosing more homogeneous subsets, this also means less dispersion within the groups. Thus, by comparing the dispersions of the acoustic indices after grouping with the dispersions before grouping, one can find out for which indices the reduction of dispersion, or increase in homogeneity, is largest. It is then tempting to consider the same indices as principal references for the regrouping made by the jury.

We have calculated the coefficients of variation, that is, the reduced standard deviation, for each index in the original set, then in each of the four groups of Table 3, taking the average of the four groups. These coefficients are presented in Figure 9.

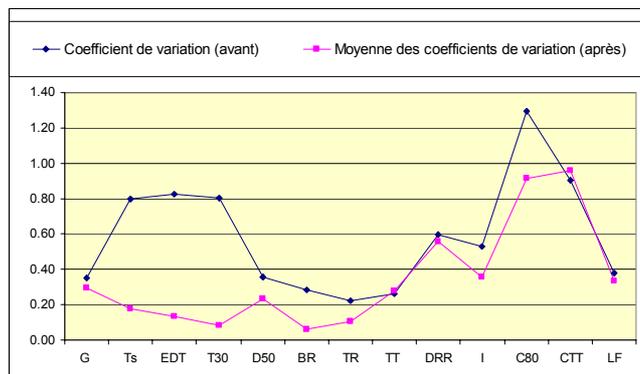


Figure 9: Coefficients of variation for all indices before and after categorisation

The indices mainly affected by categorisation are T30, EDT and Ts. We can therefore consider these indices as the most important subjectively for our jury, and the basis of their grouping.

6 Conclusion

The subjective tests presented in this paper, and its comparison with statistical analysis of the objective indices, confirm that the traditional room-acoustical indices accurately describe the subjective analysis of concert halls. And since the companion paper [9] has shown that the present database is representative of the variety of concert halls and theatres known in the literature, we can conclude that our selection of halls basically contains all ingredients for developing a typology of halls based on acoustical criteria correlated to perception. Even though subjective analysis must be carried out on other subsets of the database, the results obtained so far validate the measurement protocol and the experimental design selected for this study.

Acknowledgments

The authors thank the management of all the halls for granting permission to carry out acoustical measurements. They also thank Liu Shu for carrying out the statistical analysis. This work is part of a doctorate thesis supported by the “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)” of the Ministry of Education of Brazil.

References

- [1] L.L. Beranek, *Concert and Opera Halls. How they sound*, Acoustical Society of America, New York (1996)
- [2] S. Bertet, J. Daniel, S. Moreau, “3D Sound Field Recording with Higher Order Ambisonics – Objective Measurements and Validation of a 4th Order Spherical Microphone”, *120th AES Convention*, Paris (2006).
- [3] Denon, *Anechoic Orchestral Recordings Music and Test Signals for Evaluation of Room Acoustics*, Denon CD PG-6006 (1992).
- [4] A. Farina, P. Fausti, R. Pompoli, “Measurements in opera houses: comparison between different techniques and equipment.” *Proc. of ICA98 - International Conference on Acoustics*, Seattle (1998)
- [5] F. Guyot, *Etude de la perception sonore en termes de reconnaissance et d'application qualitative : une approche par la catégorisation*, PhD Thesis, Université Pierre et Marie Curie, Paris (1996)
- [6] ISO 3382 Acoustics, *Measurement of the reverberation time of rooms with reference to other acoustical parameters* (1997)
- [7] S. Müller and P. Massarani, “Transfer Function Measurements with Sweeps”, *Journal of the Audio Engineering Society* **49**(6), 443 (2001)
- [8] J. Poitevineau, “Méthode des arbres de similarité additifs de Sattath et Tversky : Illustration dans une tâche de catégorisation de situations d'incertitude”, *Cahiers du LCPE B44* (2002)
- [9] J.D. Polack, F. Leão Figueiredo and S. Liu, “Statistical analysis of a set of Parisian Concert Halls and Theatres”, *CFA 2012*, Nantes (2012)
- [10] A. Tversky “Features of Similarity”, *Psychological Review*, **84**, n.4, 327-352. (1977)