ENERGETIC ANALYSIS OF THE QUADRAPHONIC SYNTHESIS OF SOUND FIELDS FOR SOUND RECORDING AND AURALIZATION ENHANCING

PACS: Sound recording and reproducing systems, general concepts (43.38.Md)

Davide Bonsi, Domenico Stanzial
The Musical and Architectural Acoustics Laboratory, FSSG-CNR
Fondazione G. Cini, Isola di san Giorgio Maggiore,
I-30124-Venezia, Italy
Tel: ++39 041 5207757
Fax: ++39 041 5208135
E-mail: davide.bonsi@cini.ve.cnr.it, domenico.stanzial@cini.ve.cnr.it

ABSTRACT

A methodology for fully recovering the acoustic information of the sound field is here presented and its potentiality for enhancing both the sound recording and the auralization processes is remarked. Through the rigorous acoustical interpretation of a commercial quadraphonic coding of audio signals, an advanced intensimetric analysis of coded signals is performed on the basis of the four-vector (sound pressure-particle velocity) description of the sound field. Some case studies where sound pressure and particle velocity signals have been synthesized by convolution with quadraphonic impulse responses are then presented and two energetic indicators (sound radiation and resonance), calculated from the obtained signals, are proposed as statistical parameters for controlling the degree of acoustical fidelity of sound recording and reproduction.

INTRODUCTION

In the linear acoustic theory the stimulus giving rise to the sound event, as it could be potentially perceived by a listener, is identified by the simultaneous determination of the 4 solutions of the wave equation: sound pressure and particle velocity. A recent study of linear acoustics [1] showed that these solutions form a four-vector which is related to the generalized momentum of the acoustic field. In fact, the conservation principle applied to this four-vector gives the D'Alembert wave equation.

In this context, it is reasonable to state that the correct procedure for a true acoustical reconstruction of sound stimulus must be quadraphonic. On the other hand, some indications of this need are evident when considering the directionality of microphones used for high-fidelity recordings in professional audio. These are classified, according to their directional properties, as a combination of the following patterns: omnidirectional (circular pattern) and bidirectional (figure-of-eight pattern). The behaviour of the corresponding signals are respectively the ones predictable for sound pressure and particle velocity in a one-dimensional sound field.

In a recent work the authors of the present paper have dealt with the four-vector acoustical interpretation of audio recordings based on Ambisonics® technology [2]. It has been shown that the acquisition of sound pressure and the three components of particle velocity can be implemented by the so-called "B-format" coding, as the one based on the post-processing performed over the 4-capsules' tetrahedral array of the Soundfield microphone. Some measurements from a preliminary experiment substantially confirmed this property, even if some inevitable limitations involved in the practical employment of such a system for strict measurement purposes were pointed out.

Given the consistency of the B-format coding and the acoustic four vector representation of sound fields, the next step of our research program will be the development of a rigorous acoustical methodology for the quadraphonic recording and synthesis of sound and new hardware developments in the direct acquisition of particle velocity signals (micro-flown® technology) [9].

FUNDAMENTALS OF ENERGETIC ANALYSIS

The acquisition of sound pressure and the three components of particle velocity of a sound field allows us of perform an acoustical analysis based on energetic quantities: the sound intensity vector $\mathbf{j}$, expressing the instantaneous energy flux density at any fixed field position, and the total energy density e:

$$\mathbf{j} = p\mathbf{v}, \qquad e = e_P + e_K = p^2/2\rho_0 c^2 + \rho_0 \mathbf{v}^2/2 \tag{1}$$

These quantity are linked by the conservation equation $\nabla \cdot \mathbf{j} + \partial e/\partial t = \wp$, where $\wp$ represents the instantaneous power density, which depends in general both on the source and the boundary conditions. Starting from these quantities an investigation of energetic properties in architectural spaces for music was undertaken in the past years [3]. In particular, this has led to the introduction of a pair of local field indicators [4]:

$$\mu = \frac{\sqrt{\langle \mathbf{r}^2 \rangle}}{c\langle e \rangle}, \qquad \eta = \frac{\langle p\mathbf{v} \rangle}{c\langle e \rangle}. \tag{2}$$

μ is the *resonance indicator*, which accounts for the fraction of the energy density trapped into the normal modes excited by the source ($\mathbf{r} = \mathbf{j} - p^2 \langle \mathbf{j} \rangle / \langle p^2 \rangle$ is the *oscillating intensity* term), η is the *sound radiation indicator*, which is related to the fraction of travelling energy due to local absorption. This kind of description of sound fields has proved to be a valuable tool for the acoustic characterization of spaces as pointed out in [5].

The acquisition of pressure and velocity signals for measurements purposes is usually done by using intensity probes based on the p-p technology; this consists in the calculation of the three components of particle velocity from the temporal integration of sound pressure gradient obtained from the approximation of Euler equation.

QUADRAPHONIC FIELD ENCODING BY THE AMBISONICS METHOD

Ambisonics is an audio technology implementing a quadraphonic signal encoding, called B-format, which allows the directional properties of sound fields to be reproduced in an environment by means of an array of loudspeakers [6]. Thanks to its innovating characteristics, Ambisonics technology has always been considered a reference point among the various recording methods. in particular, it proved effective as a sound field encoding method for auralization applications.

We first want to understand the meaning of the B-format encoding from the viewpoint of the four-vector representation of sound fields; for this we start illustrating the physical principles underlying a generic electrostatic combination microphone. As known, the design of the device is done in such a way to make it sensitive to a superposition of both pressure and pressure-gradient [7]. Let us show this for the simple case of a plane monochromatic wave detected by a microphone capsule oriented along a generic direction $\mathbf{n}$; pressure and pressure gradient take the form

$$p(\mathbf{x}, t) = \hat{p}\, e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)}, \qquad \nabla p(\mathbf{x}, t) = i\mathbf{k}\hat{p} e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)}. \tag{3}$$

p is a scalar quantity, so its value does not depend on the device orientation; on the contrary, the pressure gradient is a vector, therefore its detected amplitude is maximum (minimum) for a parallel (orthogonal) orientation of $\mathbf{n}$ with respect to $\mathbf{k}$ and behaves like $\cos\alpha$ ($\alpha = \arccos(\mathbf{k} \cdot \mathbf{n})$) for intermediate angles, thus giving the so-called figure-of-eight pattern.

A combination microphone capsule can be schemed as a small rectangular air cavity having the sound diaphragm on one side and an opening with an acoustic resistance on the back side. Denoting with Δl the cavity width, the relation between p_1 (pressure at the diaphragm) and p_2 (pressure at the resistance position) is given by $p_2 \simeq p_1 + (\nabla p_1 \cdot \mathbf{n})\Delta l$. The quantity giving rise to the recorded signal is the pressure difference across the diaphragm ($p_D = p_1 - p_2$): this may be determined analyzing the system in terms of the dynamical analogy as done in [7]. The expression of p_D may be written as:

$$p_D = A\left[p_1 + iB(\nabla p_1 \cdot \mathbf{n})/k\right] \tag{4}$$

where $A = ZR / \{ZR - i[(R+Z)/wC]\}$ (being Z the diaphragm acoustic impedance, R the rear acoustic resistance and C the diaphragm acoustic compliance) and $B = \Delta l / cRC$ is a real positive number determining the polar pattern of the capsule (e.g. $B = 0 \Rightarrow$ omnidirectional, $0 < B < 1 \Rightarrow$ subcardioid, $B = 1 \Rightarrow$ cardiodid, $1 < B < \infty \Rightarrow$ hypercardioid and $B = \infty \Rightarrow$ figure-of-eigth). Eq. (4) states that the output contains a term proportional to sound pressure and one proportional to the component of pressure gradient along the direction **n**. For a steady state excitation, the pressure gradient is linked to particle velocity by the linearized Euler equation: $\nabla p(\mathbf{x},t) = iz_0 k\mathbf{v}(\mathbf{x},t)$. Then, for any given frequency, Eq. (4) may be rewritten as a sum of a sound pressure term and a particle velocity term:

$$p_D = A\left[p_1 - z_0 B\mathbf{v}_1 \cdot \mathbf{n}\right]. \tag{5}$$

The Soundfield microphone consists of four subcardioid-type electrostatic capsules assembled as a regular tetrahedron. The signals of the four capsules (A-format) are linearly combined to give the B-format according to the following matrix transformation:

$$\begin{pmatrix} W \\ X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} L_F \\ L_B \\ R_F \\ R_B \end{pmatrix} \tag{6}$$

where in the right-hand side the symbols denotes the orientations of the four capsules, which in spherical coordinates (j, J) are: L_F (left-front): $45°, 54.7°$, L_B (left-back): $135°, 125.3°$, R_F (right-front): $225°, 54.7°$, R_B (right-back): $315°, 125.3°$. As it can be seen by calculating the matrix transform with the help of Eq. (5) (that is combining pressure and velocity terms of the capsules according to Eq. (6)), this consists of a term proportional to sound pressure (W) and three terms proportional to the particle velocity components along the Cartesian axes (X,Y,Z).
Note that in practice a compensation is applied by the microphone control-unit in order to have signals referred to the same acoustic center [8].

SOUND FIELD SYNTHESIS BY CONVOLUTIONS

According to the linear acoustics theory, the sound event is described by a scalar field: the kinetic potential f, from which both the sound pressure and the particle velocity may be derived

$$p(\mathbf{x},t) = -r_0 \frac{\partial f(\mathbf{x},t)}{\partial t}, \quad \mathbf{v}(\mathbf{x},t) = \nabla f(\mathbf{x},t). \tag{7}$$

When considering the wave equation in terms of sound pressure, the solution in a given point $\mathbf{x}$ for a point-like source may be written as a convolution between a function g, the pressure impulse response, and the function describing the time-dependence of the pressure source:

$$p(\mathbf{x},t) = \int_{-\infty}^{t} g(\mathbf{x}, t-t')s(t')dt' =: g(\mathbf{x},t) * s(t). \tag{8}$$

The kinetic potential may then be written as

$$f(\mathbf{x},t) = -h(t) * p(\mathbf{x},t)/r_0 = -g(\mathbf{x},t) * h(t) * s(t)/r_0 \tag{9}$$

where $h(t)$ represents the Heaviside function ($h = 0$ if $t<0$, $h = 1$ if $t \geq 0$). It then follows:

$$\mathbf{v}(\mathbf{x},t) = -\left[h(t) * \nabla g(\mathbf{x},t)/r_0\right] * s(t) = \int_{-\infty}^{t} \left\{ \int_{-\infty}^{t-t''} -\left[\nabla g(\mathbf{x},t')/r_0\right]dt' \right\} s(t'')dt''. \tag{10}$$

Eq. (10) states that the particle velocity signal may be expressed as a convolution between the particle velocity corresponding to the pressure impulse response and the sound pressure source function. The vector $\mathbf{g}_v = -(1/r_0)h * \nabla g$ may then be named “velocity impulse response”.

EXPERIMENTAL TESTS

<u>Validation of the quadraphonic convolution</u> The theoretical statement explained above was tested experimentally: some B-format recordings were made in a room excited with simple stationary signals sent to a loudspeaker. Then, keeping the configuration loudpspeaker-microphone unchanged, the corresponding 4 impulse responses $(g_p, \mathbf{g}_v)$ were taken: these

were subsequently convolved with the monophonic input signals used for the recordings (the function $s(t)$ of Eqs. (8) and (10)) and the obtained waveforms for W and X, Y, Z were compared with the recorded ones.
As an example, we report the results for a stationary signal consisting of 5 harmonics (fund. = 200 Hz). The spectra of the X signal obtained from the two methods (Fig. 1) show a very good correspondence, the only difference being the noise level, which is higher for the recording. Fig. 2 reports the time histories of W and X: the most remarkable fact to note here is the preservation of the relative phase difference between the two signals.

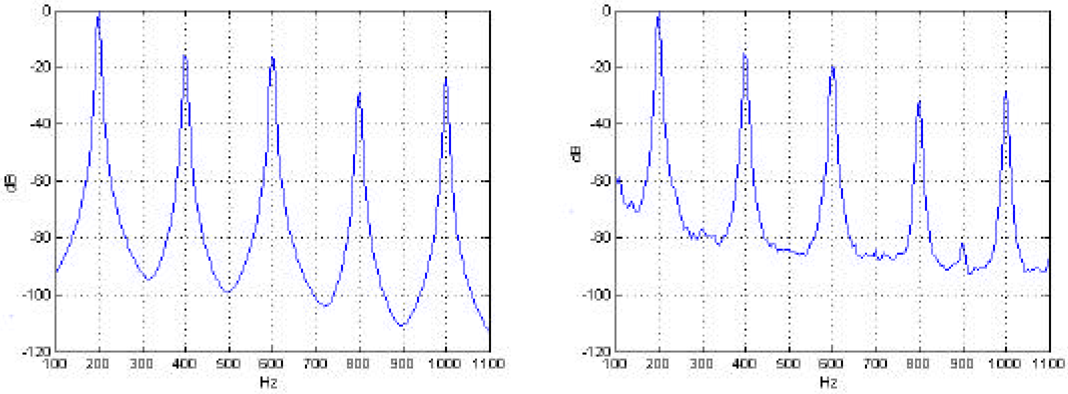

Fig. 1. - Spectra of signal X obtained by convolution (left plot) and recording (right plot).

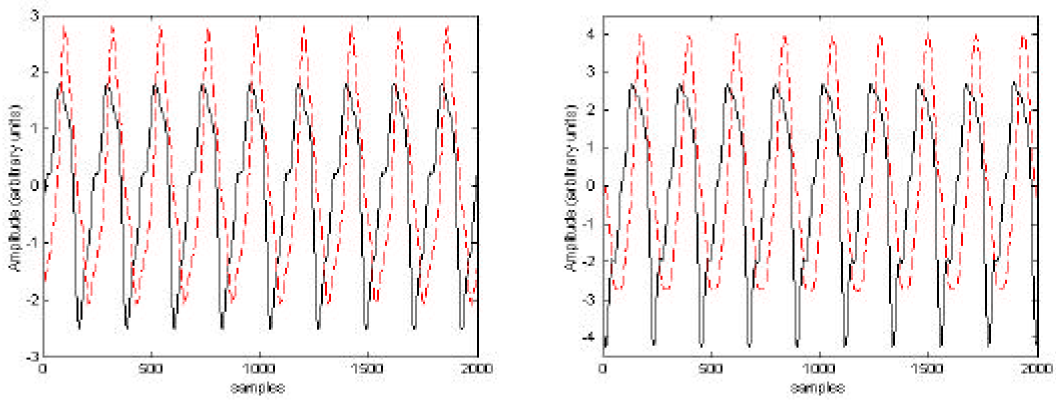

Fig. 2. – Samples of the time histories of W (solid line) and X (dashed lines) obtained from the convolution (left plot) and the recording (right).

<u>Energetic indicators of synthesized wide-band signals</u> The comparison with the recordings showed that the quadraphonic convolution based on a single input signal is the correct procedure for the synthesis of the four vector of the sound field since the amplitude and phase information between the components are preserved. Assuming the B-format encoding equivalent with the pressure-velocity detection, it is then possible to characterize the synthesized sound field by calculating the energetic quantities from the "B-format/pressure-velocity" convoluted signals. This may be of interest for achieving an acoustical characterization of the syntesized sound fields by studying the "virtual" energy transfer, which is the result of the interaction between the anechoic signal (input) and the four-vector impulse response of the room.
The calculation of m and h requires the physical normalization of the signal W with X,Y,Z in such a way their relative ratio is the same as that of p with $r_0 c v_x$, $r_0 c v_y$ $r_0 c v_z$ (see Eq. (1)). To this purpose, the Soundfield microphone was placed in a large reverberant room excited with a stationary white noise. The ratio $b = \langle W^2 \rangle / \left(\langle X^2 \rangle + \langle Y^2 \rangle + \langle Z^2 \rangle \right)$ was then calculated ($b \approx 0.32$) and from the assumption $\langle e_P \rangle = \langle e_K \rangle$, which holds when a large superposition of

modes is present, this was taken as our normalizing parameter for the subsequent tests ($X_v = \sqrt{b}X,\quad Y_v = \sqrt{b}Y,\quad Z_v = \sqrt{b}Z$).
We report the calculation of m and h for a white noise sound field synthesized from a set of impulse responses taken in a theatre (“Teatro Accademico” of Castelfranco Veneto, Italy), as shown in Fig. 3.

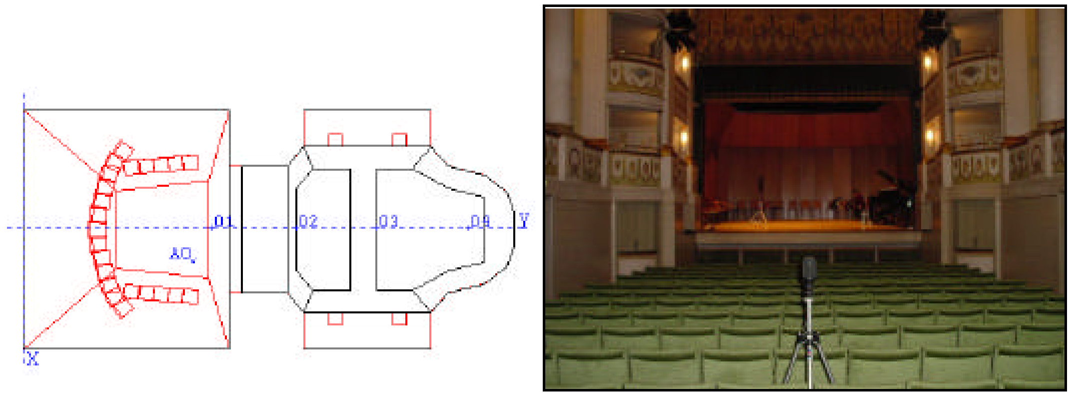

Fig. 3. – Left: the theatre plan with the source (A0) and the receiver positions (1-4); Right: a picture of the measurement setup.

The obtained values in octave bands allowed us to verify the sensitivity of the indicators to a variation in the boundary conditions, consisting in the insertion of an orchestra shell on the stage. Fig. 4 shows the comparisons of the results for the two cases (i.e. with and without the shell) referred to two positions of the Soundfield microphone: on the stage and in the centre of the stalls (point 01 and 03 in Fig. 3). The source was an omnidirectional loudspeaker placed on the right side of the stage (point A0 in Fig. 3).
It may be observed that on the stage the indicators are affected by the variation in a noticeable manner: h and m are respectively lower and higher when the shell is present. In fact, this behaviour is consistent with the increased field confinement caused by reflecting and near surfaces, whose effect is that of a decrease of local absobtion of energy (less radiant intensity compared to the energy density) and an increase of the energy of the normal modes (more modal intensity compared to the energy density).

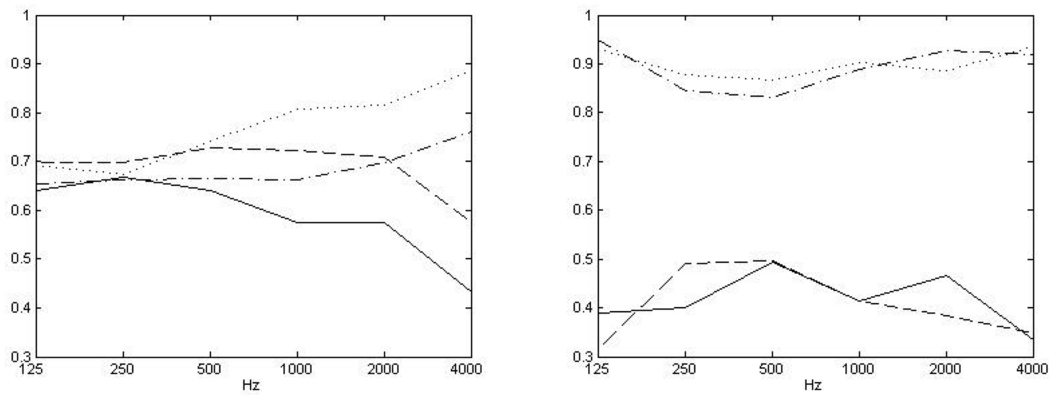

Fig. 4. – Energetic indicators in the theatre. Left plot: stage; Right plot: stalls. Solid line: h-no shell; Dotted line: m-no shell; Dashed line: h-with shell; Dashed-dotted line: m-with shell.

<u>Synthesis of a musical sound</u>. As a final application of the methodology, we present the quadraphonic synthesis of a flue organ pipe (note D_3 with a fundamental frequency of 166,6 Hz), whose anechoic signal was recorded inside an anechoic chamber. Fig. 5 shows two samples of the B-format “b-normalized” signals during the stationary phase as obtained by the convolution with the impulse responses in the stage of the theatre, with and without the orchestra shell. An evident change in the waveform shape and in the relative time relation between the components may be observed. As regards the indicators, the values found are:

$h = 0.81$, $m = 0.28$ for the stage with no shell and $h = 0.74$, $m = 0.35$ for the stage with the shell, confirming the decrease of h and the increase of m as noted above. Anyway, it must be emphasized that the values are quite different from those obtained in the case of the wide-band analysis (left plot of Fig. 4); of course, this has to be interpreted as an effect of the spectral content on the energy transfer inside the field.

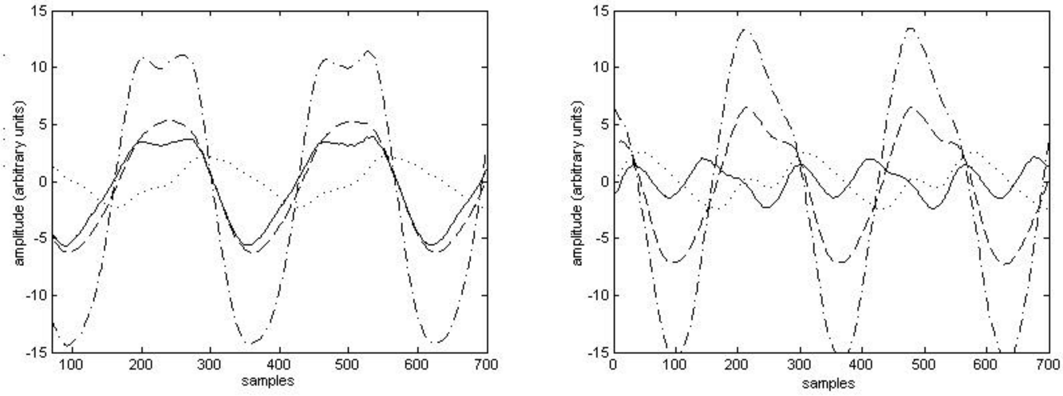

Fig. 5. – Sample time histories of quadraphonic synthesis of a organ pipe sound (normalized waveforms). Left plot: with no shell; Right plot: with shell. Dashed-dotted line: *W*; Dashed line: *X*; Solid line: *Y*; Dotted line: *Z*.

CONCLUSIONS

A rigorous acoustical methodology for the quadraphonic synthesis of sound fields (Elettra Sound®) has been presented and its potentiality for sound recording and auralization processes has been pointed out to some extent. Two sound intensity based quantities: the sound radiation and resonance indicators has been proposed as statistical parameters for quantifying the degree of acoustical fidelity of the recorded sounds. Finally, the commercial quadraphonic coding known as B-format has been recognized as an useful starting point for implementing the acoustical four-vector synthesis of sound and other recent hardware developments in the microphones' technology such as the 3-D micro-flown probe has been indicated as promising valuable tools for future implementations.

BIBLIOGRAPHICAL REFERENCES

[1] D. Stanzial, D. Bonsi, G. Schiffrer, Linear theory of acoustic field and radiation pressure, Submitted for publication on Acustica–Acta Acustica (Oct. 2001).
[2] D. Bonsi, D. Stanzial, Preliminary study of Ambisonics Technology as a method for sound intensity measurements, Proc. ISMA2001, Perugia-Italy, Sept. 10-13, 2001, pp. 449-454.
[3] D. Stanzial, N. Prodi, D. Bonsi, Measurement of New Energetic Parameters for the Objective Characterization of an Opera House, J. of Sound and Vib. 232(1), April 2000, pp. 193-211.
[4] D. Stanzial, D. Bonsi, An indicator based on modal intensity for testing the influence of the environment on the sound power of an acoustic source, Proc. ICA2001, Rome, Sept. 3-5, 2001.
[5] D. Stanzial, D. Gonzalez, D. Bonsi, Resonance monitoring by means of the modal sound intensity in the baptistry at Pisa, Proc. ISMA2001, Perugia-Italy, Sept. 10-13, 2001.
[6] M. A. Gerzon, "Periphony: With-Height Sound Reproduction", J. Audio Eng. Soc., 21, 1973, pp. 2-10.
[7] L. Beranek, Acoustics, Acoust. Soc. Am., 1954.
[8] K. Farrar, "Soundfield microphone", Wireless World, 85, 1979, pp. 48-103.
[9] See the "Research and Development - Scientific publications" section at the internet site: http://www.microflown.com