Quadraphonic impulse responses for acoustic enhancement of audio tracks: measurement and analysis

Davide Bonsi, Diego Gonzalez, Domenico Stanzial
The Musical and Architectural Acoustics Laboratory FSSG-CNR, c/o Fondazione Cini,
Isola di San Giorgio Maggiore, 36100 Venezia, Italy, {**davide.bonsi, diego.gonzalez, domenico.stanzial}@cini.ve.cnr.it**

The purpose of the present paper is the description of a method for employing energetic properties of sound fields as useful physical parameters to be used in evaluating the performance of sound reproduction systems. Following some previous works, it will be shown that a complete description of these physical properties in any point of space requires the acquisition of the so called quadraphonic impulse response (sound pressure plus the three components of acoustic velocity). This has led to the realization of a measurement system based on the new generation Microflown® intensimetric probe. It will be shown how this approach has the twofold advantage of a rigorous acoustical interpretation and the possibility of use as an intensimetric-based audio technology, since the pressure-velocity coding is formally compatible with the standard Ambisonic/B-format representation of sound fields.

1 Theoretical background

In the linear acoustic theory, any sound event can be described by a scalar field, the kinetic potential ϕ, from which both the sound pressure and the particle velocity are derived:

$$p(\mathbf{x},t) = -\rho_0 \frac{\partial \phi(\mathbf{x},t)}{\partial t}\ ; \quad \mathbf{v}(\mathbf{x},t) = \nabla \phi(\mathbf{x},t) \qquad (1)$$

All these functions are not only solutions of the D'Alembert wave equation but are interconnected by acoustical Lorentz transformation in the 4-D (space-time) acoustic Minkowski space so that, from a mathematical physics viewpoint, they form an entity known as four-vector which is related to the generalized momentum of the acoustic field [1]:

$$P = \{ p/c, \rho_0 v_1, \rho_0 v_2, \rho_0 v_3 \}. \qquad (2)$$

It is then reasonable to state that the stimulus giving rise to the sound event, as it could be potentially perceived by a listener, is identified by the simultaneous determination of all the above four functions. This means that a correct procedure for achieving a complete acoustical reconstruction of a sound stimulus in a point of space must be implemented in a quadraphonic control system.

Given a point-like pressure source term in $\mathbf{r}_M$ then pressure and velocity in the position $\mathbf{r}_S$ may be written as follows [2]:

$$p(\mathbf{r}_M, \mathbf{r}_S; t) = \int_{-\infty}^{t} g_p(\mathbf{r}_M, \mathbf{r}_S, t-\tau) s_p(\tau) d\tau \qquad (3)$$

$$v_i(\mathbf{r}_M, \mathbf{r}_S; t) = \int_{-\infty}^{t} g_{v_i}(\mathbf{r}_M, \mathbf{r}_S, t-\tau) s_p(\tau) d\tau \qquad (4)$$

where $i = 1,2,3$. The above convolution integral are a direct consequence of the linearity of the wave equation: they mean that acoustic field arises by "merging" the information embedded in the time history of the source pressure signal with any physical transmission properties of sound waves due to a specific source-receiver configuration placed in the environment, where the real sound event takes place (the so-called boundary condition). These environment specific properties are specified for all of the field quantities through the functions of the set $g_{quad} := \{ g_p, g_{v_1}, g_{v_2}, g_{v_3} \}$, whose physical meaning is that of a 4-dimensional propagator of the acoustic field, which we can name "quadraphonic impulse response". In practice, g_p and each of the g_v are given by the acoustic pressure-velocity response to an excitation being impulsive in time and point-like in space (e.g. a "pistol shot"). As a consequence, the transmission of sound energy from the excitation to the reception point depends on the way how the function representing the source term is transferred into the four components of pressure and velocity by the proper propagators. In a practical context, once g_{quad} has been experimentally determined, equations (3) and (4) can be used for the simulation of a spatialization process involving all the quantities needed for a complete description of the sound field in $\mathbf{r}_S$ due to a source signal $s_p(t)$ placed in $\mathbf{r}_M$. Since g_{quad} is supposed to carry all the acoustic information of the field except for the source, the latter has to be recorded following a procedure where wave contributions due to sound reflections are absent or negligible (e.g. into an anechoic environment).

From the point of view of the energy transfer properties within sound fields two physical quantities like the intensity vector **j**, expressing the instantaneous energy flux density at any fixed field position, and the sound energy density w have to be considered:

$$\mathbf{j} = p\mathbf{v}\ ; \qquad w = w_P + w_K = \frac{p^2}{2\rho_0 c^2} + \rho_0^2 \frac{\mathbf{v}^2}{2} \qquad (5)$$

One way of analyzing confined fields in terms of their energetic properties is through the calculation of a set of intensimetric indices:

$$\eta = \frac{\langle \mathbf{j} \rangle}{c\left(\langle W_P \rangle + \langle W_K \rangle\right)}\ ; \qquad \sigma = 2\frac{\sqrt{\langle W_P \rangle \langle W_K \rangle}}{\langle W_P \rangle + \langle W_K \rangle} \qquad (6)$$

η is the *sound radiation index*, which is related to the fraction of travelling energy, and σ is the *energy balance index*. This kind of description of sound fields has proved to be a valuable tool for the acoustic characterization of spaces as pointed out for instance in [3].

Due to the linearity properties expressed by (3) and (4), the quadraphonic impulse response plays a fundamental role in the statistical analysis of sound fields during the transient state of random (wide band noise) excitations. The main functions of this analysis are the intensity and energy decay which are obtained through the Schroeder's backward integration method, commonly employed in room acoustics for the quadratic pressure and the corresponding reverberation time [4]:

$$\overline{\mathbf{j}}_{imp}(t) = \int_t^\infty \mathbf{g}_\mathbf{j}(\tau)d\tau = \int_t^\infty g_p(\tau)\mathbf{g}_\mathrm{v}(\tau)d\tau\ , \qquad (7)$$

$$\begin{aligned}\overline{W}(t) &= \overline{W_P} + \overline{W_K} = \int_t^\infty g_{W_P}(\tau)d\tau + \int_t^\infty g_{W_K}(\tau)d\tau \\ &= \frac{1}{2\rho_0 c^2}\int_t^\infty \left[g_p(\tau)\right]^2 d\tau + \frac{\rho_0^2}{2}\int_t^\infty \left[\mathbf{g}_\mathrm{v}(\tau)\right]^2 d\tau\end{aligned} \qquad ,(8)$$

where $\mathbf{g}_\mathbf{j}(t) := g_p(t)\mathbf{g}_\mathrm{v}(t)$, $g_{Wp}(t) := \left[g_p(t)\right]^2 / 2\rho_0 c^2$ and $g_{W_K}(t) := \rho_0^2\left[\mathbf{g}_\mathrm{v}(t)\right]^2/2$ are respectively the sound intensity, potential energy and kinetic energy corresponding to an impulsive excitation.

An important relation allows to express time averages of quadratic quantities in p and v through integrals of the corresponding impulse responses, as shown by the following example:

$$\begin{aligned}\left\langle \left[p(t)\right]^2 \right\rangle &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T}\left[\int_{-\infty}^{t} g_p(t-\tau)s_p(\tau)d\tau\right]^2 dt \\ &= \int_0^\infty \left[g_p(t)\right]^2 dt\end{aligned} \qquad (9)$$

where the last expression holds thanks to the autocorrelation property of white noise $N(t)$:

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} N(t)N(t+t')dt = \delta(t') \qquad (10)$$

It then follows:

$$\left\langle \mathbf{j}_{stat}(t) \right\rangle = \overline{\mathbf{j}}(0)\ ; \quad \left\langle W_{stat}(t) \right\rangle = \overline{W}(0) \qquad (11)$$

Thus, stationary energetic indices of Equation (6) can be calculated from statistical averages of the impulsive quantities.

1.1 Acoustic quadraphony

"Acoustic quadraphony" is the audio process developed by FSSG-CNR acoustics laboratory as an application of the previously illustrated acoustic intensimetric treatment of sound fields to the audio domain [6]. It thus stands for the specific coding – together with its related technological process – by means of which it is possible to identify in a physically rigorous manner, the mechanical state of stationary sounds at an arbitrary point of space where even sound transient events take place. This identification can be obtained combining the knowledge of the particular set of four solutions of the sound wave equation given in Eq. (2) with the time-average energetic properties of general sound fields as given for instance by the indicators reported in Eq. (6). Acoustic quadraphony can be considered an audio "meta-format", since in addition to the audio signal itself it contains also all acoustical information needed eventually to synthesize the spatial attribute of a pseudo-anechoic audio signal by convolution with a 4-IR experimentally measured in a real environment.

2 The measurement method

The first practical task to be accomplished for acoustic ambience identification and reconstruction based on intensimetric techniques is the development of a measurement system capable of extending the standard measurement of pressure impulse response to the four vector-case. In this context, the most critical issue is the choice of a transducer capable of giving, both the sound pressure and the particle velocity signals. Usually, this is done just in the context of sound intensity measurements (e.g. for sound power calculation of noise sources) by using intensity probes based on the *p-p* technology and the implementation of the acoustic Euler equation: these probes (see for instance the B&K 4135) consist of pairs of close pressure microphones by means of which the three components of particle velocity are calculated from the time integration of the finite difference approximation of the pressure gradient. Nevertheless, this technique is not suitable for audio applications due to some practical reasons: the limited frequency bandwidth (varying with the microphone spacer), the high number of channels involved (up to 6 for the measurements along the three axes) and the systematic errors due to channels phase mismatch and the finite difference approximation.

On the other hand, in the field of professional audio, one interesting solution is given by the Ambisonics®/B-format technology as implemented for instance by the *Soundfield®* microphone [8,9]: this transducer provides 4 signals referred to 4 nominally coincident microphones, one of them being omnidirectional and the remaining three bi-directional and oriented along three Cartesian axes. Such a coding is achieved by linearly combining the signals given by four mixed pressure-gradient condenser capsules assembled on a tetrahedral array. Through a proper matrix conversion the four basic signals are recovered giving the so-called "B-format" encoding (W,X,Y,Z).

It was proved experimentally that when used into an anechoic environment this microphone provides 4 signals that are consistent with those of pressure and velocity. This type of quadraphonic coding has been recognized as an useful starting point for implementing the entire process. Some surveys in architectural spaces were in fact made employing this technique and some preliminary tests on the acoustical four-vector synthesis of sound were performed [5].

In practice, Ambisonics technique is satisfactory for standard audio applications but not as much for physical measurements involving the calculation of the energetic parameters requested by the acoustic quadraphony approach. This is due to a number of assumptions whose reliability cannot be taken for granted in any field condition. The most remarkable are:

- sensitivity of capsules to the same perfect mix of sound pressure and pressure gradient at all frequencies;
- validity of linear version of Euler equation for the indirect measurement of acoustic velocity;
- perfect spatial coincidence of the capsules;
- absence of any significant diffraction effect due to microphone size.

Nonetheless, the current subject of investigation regards the possibility of using new generation transducers capable of directly measuring particle velocity signals, in such a way to overcome the limitations inherent in all of the traditional indirect methods based on pressure microphones signals.

The recent developments in the transducers' technology represented by the *Microflown®* probe (by 'Microflown Technologies' B.V., Zevenaar, The Netherlands [10]) have been indicated as promising valuable tools for future implementations of the process. Our efforts are currently devoted to test this technology with the aim of building an innovative system for the quadraphonic coding of sound fields implementing both the acquisition of 4-D impulse responses to be used in the convolution-based synthesis process and the quadraphonic recording of generic sound sources. In the fall of 2003 the manufacturer introduced a new sensor (TITAN model), which is characterized by a 15 dB(A) improved selfnoise compared to the previous element, so that the effective measurement range reaches an upper frequency of about 15 kHz.

As regards the hardware components the system consists of a laptop PC connected to a MOTU® 896HD firewire audio interface. The sound source is a dodecahedron loudspeaker LOOKLINE D301 equipped with an equalizing circuitry capable of giving an almost flat frequency excitation (approx. ± 2.5 dB) in the range 60÷16000 Hz, so as to allow the recovery of wide-band impulse responses to be used for faithful convolution applications in the audio domain

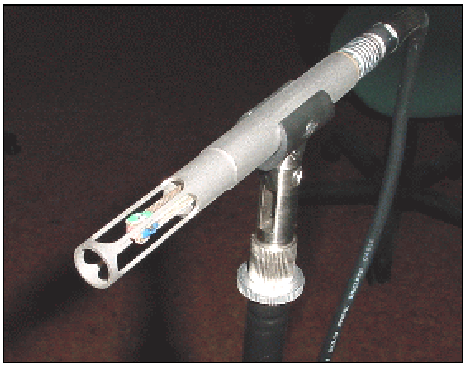

Figure 1: The 3D Microflown probe on the tripod.

The measurement method is based on the standard cross-correlation between excitation and response. A monophonic pressure signal is sent to the environment (Maximum Length Sequences or Swept sines) and its quadraphonic response is synchronously acquired. The measurement session consists in the following steps:

1. Choice of the measurement technique (Maximum Length Sequence or Swept Sines) and parameters settings;
2. Definition and storing in memory of excitation signal;
3. Sound field generation and quadraphonic recording followed by memorization of acquired data;
4. Calculation of impulse responses by Fast Hadamard Transform or Fast Fourier Transform (depending on the chosen technique), normalization of data according to sensors sensitivities and file saving.

The entire procedure is implemented in a MATLAB routine based on the Data Acquisition Toolbox for signal input/output. A graphical user interface has been developed for setting the measurement parameters.

3 Sample measurements

The above illustrated measurement system has been tested by collecting a library of quadraphonic impulse responses in rooms characterized by different acoustical properties. These were some historical buildings of the “Cini Foundation” on the Island of San Giorgio Maggiore in Venice, venue of the FSSG-CNR Acoustics Laboratory. The survey was one of the FSSG tasks in the RACINE-S European project, IST-2001-37117, ended in December 2004 [7]. About 50 data sets have been taken in different positions of 7 rooms.

3.1 Impulse responses analysis

Figure 3 reports an example of quadraphonic impulse response taken in a big shoe-shaped conference room ($L \times W \times H \cong 30\ m \times 12\ m \times 7\ m$) with the probe and the source located as shown in Figure 2. Some results of the intensimetric analysis performed over the impulse response are also reported: Figure 4 shows the first 80 *ms* of the impulsive intensity $\mathbf{g_j}(t)$ along X and Y, from which it is possible to ascertain a prevailing “radiative” behaviour along X (i.e. the probe-source joining line) and an “oscillating” behaviour along Y. Figure 5 and 6 respectively show the corresponding Schroeder plots (see Equation (7)) and the octave-band values of η.

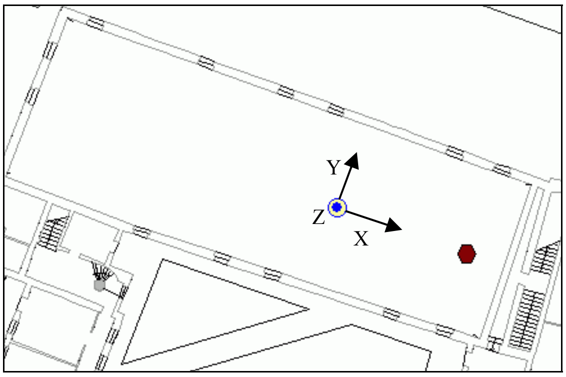

Figure 2: Room plan with symbols indicating the probe position and the source location (hexagon).

3.2 Sound fields rendering

As an example of quadraphonic convolution (see Equations (3) and (4)) Figure 7 reports the four waveforms obtained by convolving the previous quadraphonic impulse response with an anechoically recorded signal (sound of a flue organ pipe). These plots highlights some properties characterizing the four components of the sound field like the differing shape of the attack and the differing amplitude and phase values during the stationary condition.

4 B-Format compatibility

Ambisonics technique has been developed as a low-order approximations to the holophonic reproduction of sound fields. The four signals, as those provided by the Soundfield microphone, are not used as a pressure-velocity coding, but as the first elements of the spherical-harmonics decompositions of pressure field [11]. B-format coding is then linearly transformed in such a way to recover the signals corresponding to a virtual group of coincident microphones with polar patterns varying from omnidirectional to figure-of-eight and pointing in arbitrary directions of space. In this way it is possible to feed a loudspeakers system assembled in the same geometrical pattern as this first-order microphone array.

Past studies showed that, from the viewpoint of its formal working principle, B-format representation is equivalent to the acoustic quadraphonic pressure-velocity coding [5]. Some tests are currently in progress for evaluating the compatibility of the Microflown pressure-velocity signals as B-format input to the Ambisonics based decoding matrices (mainly for UHJ stereo and 5.1 reproduction [12,13]).

5 Conclusions

A short account of the theoretical framework and sample implementation of a new audio technology called “acoustic quadraphony” has been presented. This methodology has been first conceived as a tool for enhancing the acoustic ambience audio effect of regenerated cinema audio tracks within the Racine-S EU Research Program. Anyway, this program was not concerned about the problem of the actual quadraphonic acoustical rendering of audio tracks so, after the regeneration process, the actual playback was accomplished in a stereo format. The problem of facing the real acoustical quadraphonic rendering of soundtracks is now a running research task at FSSG-CNR Lab within the IP-Racine program whose accomplishment is programmed within 2007. At the present state of the research a strict comparison between the existing multichannel sound rendering systems and the control system envisaged for the real acoustic quadra-sound rendering is under study. The here reported results represent an encouraging step on this research route.

6 Acknowledgements

The present research has been supported by EU projects RACINE-S and IP-RACINE respectively under contracts IST-2001-37117 and IST-2-511316-IP [14].

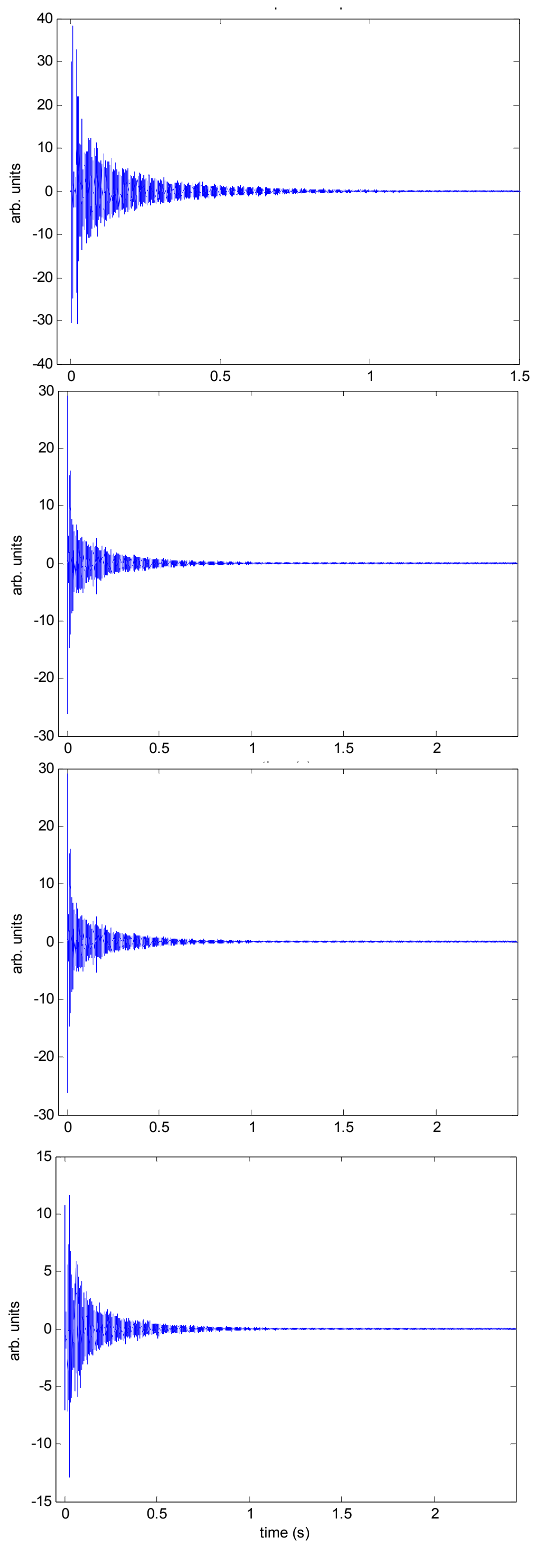

Figure 3: From top to bottom, pressure and X,Y,Z components of acoustic velocity impulse responses of the conference room.

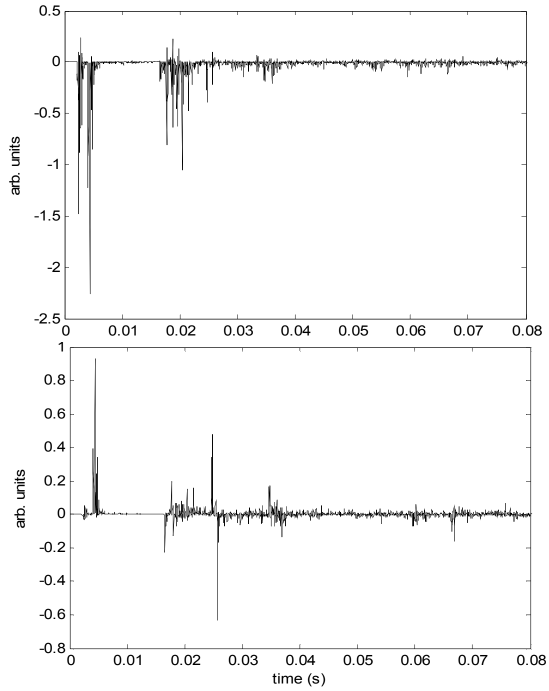

Figure 4: Impulsive intensity components along directions X (top plot) and Y (bottom plot).

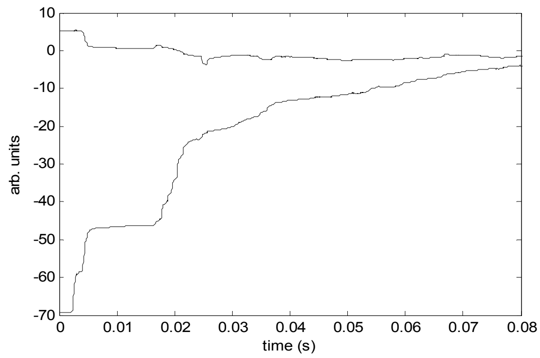

Figure 5: Intensity Schroeder plot along directions X (bottom line) and Y (top line).

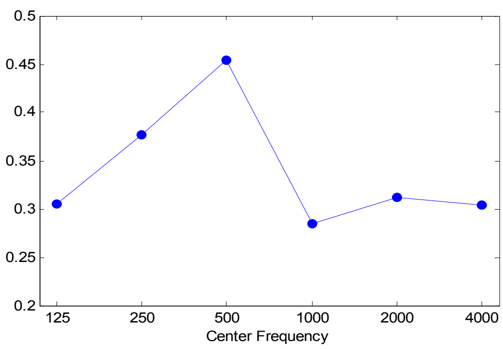

Figure 6: η indicator calculated in octave bands from the sample quadraphonic impulse response.

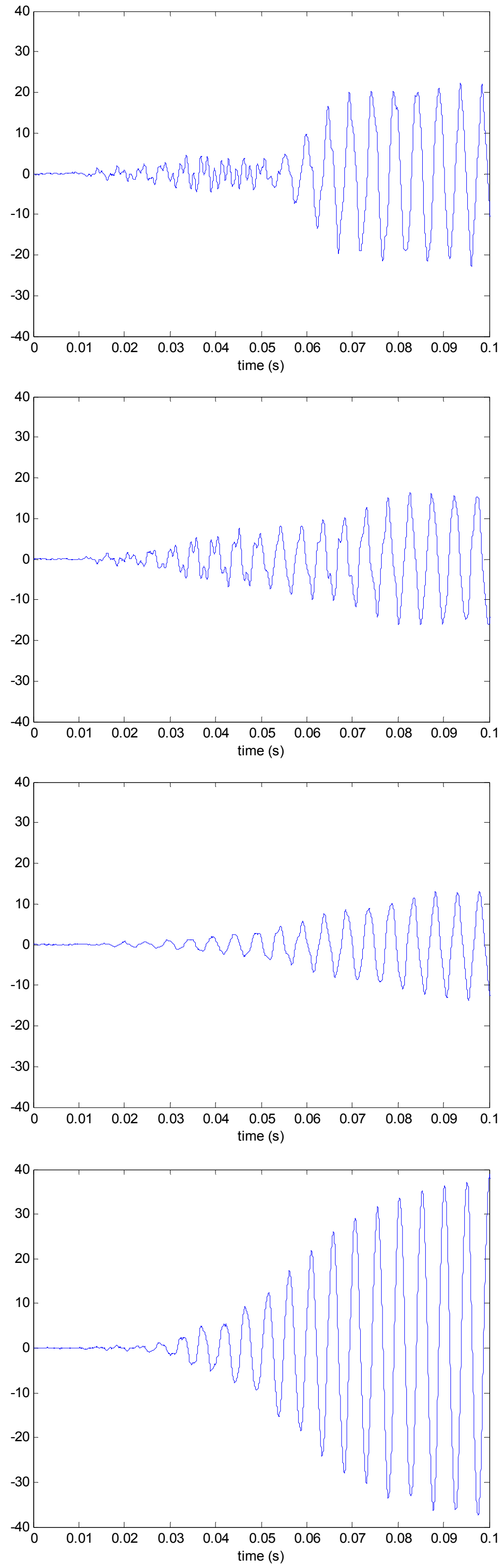

Figure 7: From top to bottom, pressure and X,Y,Z components of acoustic velocity of an organ pipe tone synthesized by convolution with quadraphonic impulse response of Fig. 3.

References

[1] D. Stanzial, D. Bonsi, G. Schiffrer, *Four-dimensional Treatment of Linear Theory of Acoustic Fields and Radiation Pressure*, Acustica-Acta Acustica, **89**, pp. 213-244 (2003).

[2] D. Bonsi, D. Stanzial, *The particle velocity impulse response: theoretical basis and measurement method*, Proc. 6th International Congress on Sound and Vibration, Copenhagen (1999).

[3] D. Stanzial, D. Bonsi, N. Prodi, *Measurement of new energetic parameters for the objective characterisation of an opera house*, J. of Sound and Vib. **232** (1), pp. 193-211(2000).

[4] M. Schroeder, *New methods for measuring reverberation time*, J. Acoust. Soc. Am **37**, pp. 409-412 (1965).

[5] D. Bonsi, D. Stanzial, *Energetic analysis of the quadraphonic synthesis of sound fields for sound recording and auralisation enhancing*, Proc. EAA Forum Acusticum, Sevilla (2002).

[6] D. Stanzial, D. Bonsi, G. Schiffrer, *Equipment for collection and measurement of quadraphonic sound data and metadata and related procedure*, European Patent n. 04425909.1, registered: Dec 2, 2003.

[7] D. Bonsi, F. Fontana, D. Gonzalez, D. Stanzial, *Report on the construction of a library of impulse responses for acoustic enhancement*, RACINE-S project deliverable D4.1.3 (2004).

[8] M. Gerzon, *Periphony: With-Height Sound Reproduction*, J. Audio Eng. Soc., **21,** pp. 2-10 (1973).

[9] K. Farrar, *Soundfield microphone*, Wireless World, 85, pp. 48-103 (1979).

[10] H-E. de Bree, *An overview of Microflown Technologies*, Acustica-Acta Acustica, **89**, pp. 163-172 (2003).

[11] M. Poletti, *A Unified Theory of Horizontal Holographic Sound Sistems*, J. Audio Eng. Soc., **48**, No. 12, pp. 1155-1182 (2000).

[12] M. Gerzon, *Ambisonics in Multichannel Broadcasting and Video*, J. Audio Eng. Soc., **33**, No. 11, pp. 859-871 (1985).

[13] M. Gerzon, G.J. Barton, *Ambisonics Decoders for HDTV*, Preprint 3345 of the 92nd Audio Eng. Soc. Convention, Vienna (1992).

[14] http://www.ipracine.org.