FORUM

RECORDING CONCERT HALL ACOUSTICS FOR POSTERITY

MICHAEL A. GERZON

Mathematical Institute, University of Oxford, Oxford, England

Heyser [1] has asked if there is any way of incorporating into modern recordings information enabling a future technology to recover the original sound field from the recording. In the following we observe that it is possible to preserve useful information concerning concert hall acoustics, and that this may be used to "reverberate" recordings artificially with a given acoustic. Proposals are made for recording the acoustic itself (in a way independent of the music) for the purpose of studying architectural and musical acoustics.

Recent developments [2]–[4] in ambisonic recording and reproduction technology (i.e., the recording, storage, and recreation of the sound field pressure and three-dimensional velocity at a point; see also [5], [6]) have shown that it is possible both to transduce the lower order directional components of a sound field at a point with good accuracy up to frequencies around 7.5 kHz, and to create via four or more loudspeakers horizontally (or six or more for periphonic reproduction [7]) a reasonably faithful illusion of the original directional field, including the effect of sound source distance, whether closer or further than the loudspeaker distance used in reproduction.

The primary use of such a sound field microphone array [3] is to take down as accurately as possible the relative balances and directionalities of individual sound sources, early reflections, and the reverberant field. In this sense, such a recording using four channels[1] is in itself a "Rosetta stone" [1] for the sound of the original performance. The only "subjective" element of control by the recordist is the positioning of such a microphone, although he may also subsequently use the already recorded sound field as a basis for further processing to obtain a desired less natural effect.

In cases where a nonnatural recording is made (using close microphones and multitrack tape), it is clearly not possible to use this information plus further information about the hall acoustics (whatever the form in which the latter information may be stored) to recover an accurate recreation of the sound of the original performance. There are a number of reasons for this impossibility. First, the number of channels of information required to store all the sound field information over the audio frequency range even over a listening region of only 2-m diameter is around 400 000 by sampling theorem arguments, with the number of channels increasing proportionately to the listening area covered. Second, each of the recorded sound sources is a directional emitter of sound, and the excitation of the hall by this unknown frequency-dependent polar diagram is unpredictable, especially as the close microphone placement can only sample one part of the polar diagram. Third, even small movements or rotations of a performer will change the hall's response to his/her sound. Fourth, normal changes in the disposition of people and objects in a hall (and changes in temperature and humidity) will change the hall response in an unpredictable fashion during the performance.

Also, the layout of performers during a multimicrophone recording session is usually quite different to the optimum performance layout in order to give improved microphone separation. The increased spacing of performers gives acoustic time delays (often in excess of 0.1 s) that affect the timing of cues heard by the musicians, and hence the musical interpretation. It is thus not evident that a recreation of the "studio" performance in the original hall acoustics would be satisfactory in musical terms, even if technically possible. It must also be observed that the empty halls often used for recording sessions rarely have an optimum acoustic. Adding to a recording an acoustic different from that heard by the performers is likely to make those aspects of the original performance depending on the use of the original hall acoustic seem to be musically ill-judged.

Nevertheless, it is perfectly feasible to record the acoustic of concert halls for posterity subject to the abovestated limitations. The procedure is essentially a well-known one. An accurate impulse (e.g., a spark discharge) is launched into the hall from a position A_i and picked up by

[1] This 'tetraphonic' use of four channels to represent components of the sound field is, of course, quite distinct from the use of four audio channels to convey speaker-feed signals for a square loudspeaker layout.

sound field microphone arrays placed at positions B_1, $B_2, \cdots, B_n$ chosen to be representative of good listening positions. The four-channel impulse information picked up by the array placed at B_j is the hall's four-channel convolution kernel at B_j for an omnidirectional source at A_i. When reprocessing a multimicrophone recording, one first has to decide which position A_i each track is to be placed at. Clearly, different performers should occupy differing positions A_i of the stage, both laterally and in distance. One must then decide which listening position B_j one wishes to simulate. If the kth track of our four-channel recording has a recorded impulse $h_{ij}^{(k)}(t)$ corresponding to the B_j listening position and A_i source position, then the resultant four-channel signal will be, for $k = 1, 2, 3, 4$,

$$g_j^{(k)}(t) = \sum_{i=1}^{m} \int_0^\infty h_{ij}^{(k)}(\tau) f_i(t - \tau)\, d\tau \qquad (1)$$

where $f_i(t)$ is the audio signal track assigned to the position A_i. The computation of the B_j position sound (1) can even now be carried out (at great expense and time) via digital processing techniques, although one expects that the technology of the near future should enable the signals $g_j^{(k)}(t)$ to be produced without any very great difficulty.

There are a number of possible improvements to the basic proposal outlined above. The signal-to-noise ratio of the recorded four-channel impulse response may be improved if the impulse is replaced by a wideband pseudorandom sequence or a frequency sweep, and if the recorded "reverberated" sequences picked up at B_j are digitally or otherwise deconvoluted back to reverberated impulse responses. The speed accuracy of analog recorders is not high enough for the deconvolution to be performed accurately off tape, and digital recording or direct digital processing will be required. The pseudorandom or sweep signals may be reproduced in the hall through high-quality full-range electrostatic loudspeakers.

The remaining problem is the actual procedure used to record our great concert halls. Most halls sound at their best with the acoustic absorption of a (possibly noncapacity) audience and orchestra. Ideally, therefore, the recording of the hall convolution kernels should be taken in the presence of a *very* quiet audience and orchestra. This may be done by setting up the microphones B_j in good listening locations among the audience, and by recording the impulse from A_i simultaneously from all the microphones on several tape recorders. It would be useful if the original impulse fed to the hall were also to be recorded at the same time, so as to preserve time-delay information between A_i and B_j. Recording the kernels for each source position should not take more than 5 or 10 seconds. It may be possible to get an audience to cooperate for a couple of minutes while 15 or 20 positions A_i are recorded. Similar cooperation has not been difficult to elicit in cases (such as the Royal Albert Hall, London) where acoustic modifications have required testing with an audience. The alternative is to cover the hall with a "replica" audience of acoustic absorbers; this seems expensive.

It may well be worthwhile to allocate funds for a project to record for posterity the world's concert halls and churches. Clearly, the equipment involved would be expensive for a "one-off" measurement, but quite reasonable as part of a continuing program of measurements. The data obtained would be useful to acousticians, musicians, historians of music, and even to recording organizations wishing to add a given "concert hall" to existing multimicrophone recordings when the technology becomes sufficiently economical. This project would seem to be ideally suited for organization and coordination by the learned societies in acoustics and audio.

The sound field transducing technology [3] was originally developed in connection with the ambisonic surround-sound project supported by the U.K. National Research Development Corporation and with the help and cooperation of Calrec Audio. Improved transducer technology to reduce the remaining imperfections above 7.5 kHz is under development.

REFERENCES

[1] R. C. Heyser, "A Rosetta Stone for Audio," *J. Audio Eng. Soc.*, vol. 22, p. 251 (May 1974).

[2] P. B. Fellgett, Ambisonic Reproduction of Directionality in Surround-Sound Systems," *Nature*, vol. 252, pp. 534-538 (Dec. 1974).

[3] M. A. Gerzon, "The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound," presented at the 50th Convention of the Audio Engineering Society, London, March 4, 1975.

[4] M. A. Gerzon, "Surround Sound Psychoacoustics," *Wireless World*, vol. 80, pp. 483-486 (Dec. 1974).

[5] J. J. Gibson, R. M. Christensen, and A. L. R. Limberg, "Compatible FM Broadcasting of Panoramic Sound," *J. Audio Eng. Soc.*, vol. 20, pp. 816-822 (Dec. 1972) (especially the Appendix).

[6] M. A. Gerzon, "Experimental Tetrahedral Recording" (in 3 parts), *Studio Sound*, vol. 13, pp. 396-398, 472-475, 510-515 (Aug., Sept., and Oct. 1971).

[7] M. A. Gerzon, "Periphony: With-Height Sound Reproduction," *J. Audio Eng. Soc.*, vol. 21, pp. 2-10 (Jan./Feb. 1973).

MORE ON "RECORDING CONCERT HALL ACOUSTICS FOR POSTERITY"

RICHARD C. HEYSER

California Institute of Technology,
Jet Propulsion Laboratory,
Pasadena, Calif.

Mr. Gerzon has struck the nail squarely on the head. There is genuine need for preserving not only the total acoustics associated with artistic performances, but also the acoustics of our major halls. The need for the proper use of the four parameters of pressure and the true three-dimensional velocity is completely correct. Readers may recall one of the results of another type of analysis presented a short while ago [8].

" . . .the prediction is that we must record both pressure (a scalar) and velocity (a vector) before we can truly begin approaching accurate reproduction from the standpoint of recreating a replica sound field at a later time and a different place."

The recreation referred to, of course, is the appropriate acoustic field including angle and space location of all significant sources relative to any arbitrary listening position. If that field is the associated point spread function of a concert hall, then this information may be used to isolate the hall as an acoustic entity. That would be the direct result of recording the proper acoustic signals at the time of a performance and was one of the intents behind my request for such a thing. Later processing could either extract that field from the recording made there or modify it to suit the reproduction methods of future technology. This conclusion was also stated by Davis [9].

The idea of setting up a "performerless" recording session and using a "Rosetta stone" signal for recording the hall's point spread function has great merit and I heartily endorse the concept.

The limited experience I have had causes me to shun loudspeakers as a reference acoustic source, particularly if the same sound is to be produced some generations hence. That is why I suggested an omnidirectional acoustic spark source. However, the advantage of loudspeakers in generating signals with a greater time-bandwidth product than such a spark may well cause them to be preferred from a practical standpoint in deconvolution. But this means we must more accurately measure the true loudspeaker response than is now our practice.

I only have one small misgiving about the preservation of a hall's acoustics. That regret is that it will probably be a short time before some entrepreneur uncorks the computer genii and electronically places performers in the glorious ambiance of good halls without the necessity for their physical presence there. The "computer hall" could well become the replacement for our present echo chambers. Such is progress.

REFERENCES

[8] R. C. Heyser, "Geometrical Considerations of Subjective Audio;" *J. Audio Eng. Soc.*, vol. 22, pp. 674-682 (Nov. 1974).

[9] D. Davis, "A Rosetta Tone?," *J. Audio Eng. Soc. (Forum)*, vol. 22, p. 430 (July/Aug. 1974).