PREPRINT NO. 1571 (A6)

PRACTICAL PERIPHONY: THE REPRODUCTION OF FULL-SPHERE SOUND

Michael A. Gerzon, Mathematical
Institute, University of Oxford,
England

**Presented at
the 65th Convention
1980 February 25 through 28
London**

AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

PRACTICAL PERIPHONY: THE REPRODUCTION OF FULL-SPHERE SOUND

Michael A. Gerzon, Mathematical Institute, University of Oxford,

Oxford OX1 3LB, England.

Draft of lecture presented at Audio Engineering Engineering Society 65th Convention, London, 25 February 1980

Periphony is a term meaning the recording and reproduction of sound via loudspeakers from a full sphere of directions, including both all horizontal directions and all elevated and depressed directions also.

The title of this lecture refers to "practical periphony", so let's begin by seeing what impractical periphony is like. Figure 1 shows everyone's idea of impractical full-sphere sound reproduction, i.e. the use of 12 speakers at the face-centres or of 20 speakers at the vertices of a regular dodecahedron. The reasons for the impracticality of this is self-evident to all whose living rooms are not anechoic chambers with wire-mesh floors.

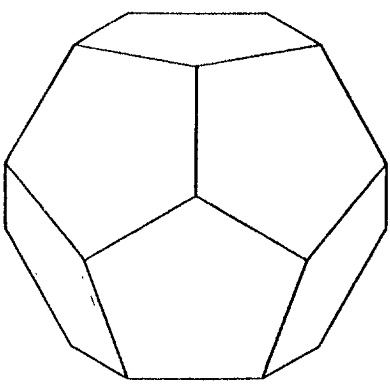

Figure 1.

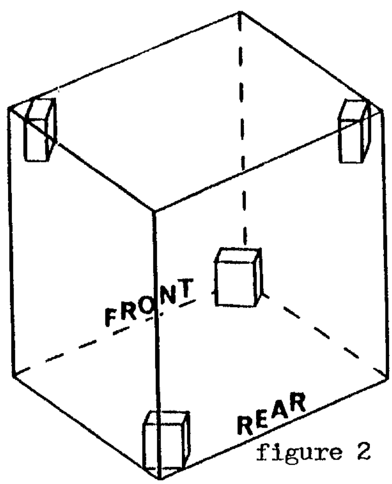

figure 2

Another impractical means of full-sphere reproduction is the tetrahedral system shown in fig. 2, or any other system using a tetrahedron of speakers. Such systems were first proposed by Pierre Schaeffer around 1952, and rediscovered by Granville Cooper, the present author and Jerry Bruck around 1970. In this case, the reason for impracticality is that the reproduced results have a number of defects (which we shall describe a little later). These defects are inherent in the tetrahedral speaker layout. Since the early 1970's research has continued in finding methods of periphonic reproduction avoiding both domestic unfeasability (fig. 1) and poor subjective results (fig. 2).

This progress has become possible due to the development of a fairly comprehensive theory of the psychoacoustics of directional reproduction, which is cast in a mathematical form that allows theorems to be proved that greatly simplify design of equipment for optimum subjective results. The technology based on this theory is termed Ambisonic. Hitherto, only ambisonic equipment for horizontal surround-sound has been available, but the theory is equally applicable to the vertical dimension also. Apart from showing the design equations for periphonic decoders at the end of the lecture, we shall avoid mathematics and explain the principles involved, along with design trade-offs, as simply as possible.

First we describe some of the psychoacoustic theory used, in a form adapted for design work. Oversimplifying, it may be said that the ears use two different types of method of locating sounds, one at low frequencies below 700 Hz, and one at high frequencies above that. Low frequency localisation is determined by the phase difference between the two ears and high frequency localisation by intensity differences. Two theories of sound localisation based on these mechanisms are the Makita theory and the Energy vector theory. The Makita localisation of a reproduced sound is that direction in which the head has to face in order that the interaural phase difference is zero. (In 3 dimensions one has to be a little careful what one means by this, since one can face a given direction and yet still tilt one's head from side to side. We mean that the interaural phase difference stays zero despite one rotating the head around the direction in which it faces. The Makita localisation may well not be horizontal). In a similar way, the energy vector localisation is the direction the head has to face in order that there be no interaural amplitude difference at high frequencies.

There is a way of picturing these directions, and also aspects of localisation not covered by these theories. Imagine a loudspeaker layout, such as fig. 3 (for ease of illustration we only show a simple horizontal layout.) Draw a vector as shown from the centre of the loudspeaker layout to each loudspeaker, giving each vector a length proportional to the "amount" of sound emerging from the speaker it points towards. At low frequencies, this "amount" is the amplitude gain of the sound in each speaker, and at high frequencies the energy gain of the sound in each speaker. (At low frequencies, an antiphase sound will have negative gain, in which case its vector will point away from the speaker!). Now add up the total magnitude (length) of all vectors and we get a "total amount" of sound at the centre. Also add up the vectors. The direction of the resultant vector is the Makita localisation (at low frequencies) or the energy-vector localisation (at high frequencies).

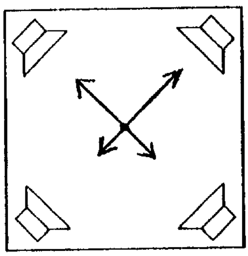

figure 3

When the head is pointing in directions other than the Makita or energy-vector direction, the perceived localisation will in general differ, so that as one rotates ones head, the sound image will move. The image will be stable under rotation only if the magnitude of the resultant vector is precisely the same as the total amount of sound from the loudspeakers, since this condition obviously holds if there is a live sound, ie just one "loudspeaker". The ratio of the length of the resultant vector length to the total amount of sound is called the vector magnitude of the sound, and should ideally equal one. Good decoder design consists of getting both the Makita localisation and the energy vector localisation correct for all sound directions at all frequencies, getting the low frequency vector magnitude r_V equal to one at low frequencies and the energy vector magnitude r_E as close to one as possible at high frequencies. In practice, it turns out that r_E is always less than one, so that one aims to get it as big as possible.

The first results that help design optimal playback apparatus are the following very useful assertions.

DIAMETRIC DECODER THEOREM

MAKITA AND ENERGY-VECTOR LOCALISATION COINCIDE IF :-

(1) ALL SPEAKERS ARE SAME DISTANCE FROM CENTRE OF LAYOUT

(2) SPEAKERS ARE PLACED IN DIAMETRICALLY OPPOSITE PAIRS

& (3) THE SUM OF THE 2 SIGNALS FED TO EACH DIAMETRIC PAIR IS THE SAME FOR ALL DIAMETRIC PAIRS

BONUS :-

DIAMETRIC SPEAKER LAYOUTS CAN BE FED <u>OPTIMAL</u> SPEAKER FEEDS USING ONLY n+1 CHANNELS OF AMPLIFICATION VIA A "SPEAKER MATRIX", WHERE n IS THE NUMBER OF PAIRS OF SPEAKERS.

i.e.	4	SPEAKERS	NEED	3	AMPLIFIERS
	6	"	"	4	"
	8	"	"	5	"

The tetrahedral speaker layout shown earlier in fig. 2 does not satisfy the diametric decoder theorem, and the Makita and energy-vector localisations do not coincide. In fact computations of the energy vector localisation show that sounds at high frequencies are very much drawn towards the four loudspeakers of the tetrahedral layout, as shown for the left-front-up octant in figure 4. The centre of this picture is one of the loudpeakers, and it will be seen, for example, that the energy vector localisation of a "left-front" horizontal sound is actually pulled to the other side of the speaker! This problem of sounds being pulled towards the speakers was in fact noticed in early experiments in tetrahedral recording, and is the reason why other speaker layouts must be used.

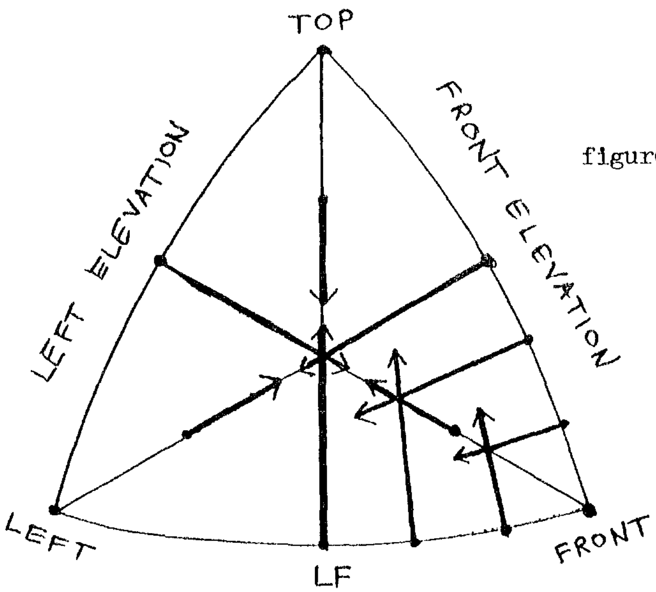

figure 4

Three speaker layouts satisfying the diametric decoder theorem requirements are shown in figures 5-7. These are the "cuboid", the octahedron and the "birectangle" layouts. In all these layouts, it is not nessary that the different rectangular sides have the same lengths, as long as the signals fed to the speakers are suitably compensated for the layout shape. The birectangle layout has the distinct advantage that it provides a conventional stereo speaker pair for reproducing older stereo recordings.

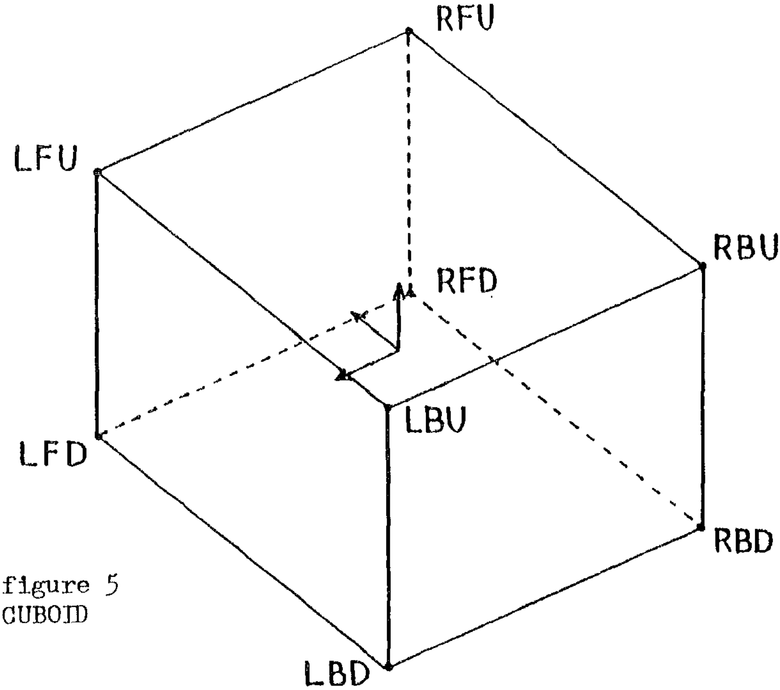

figure 5
CUBOID

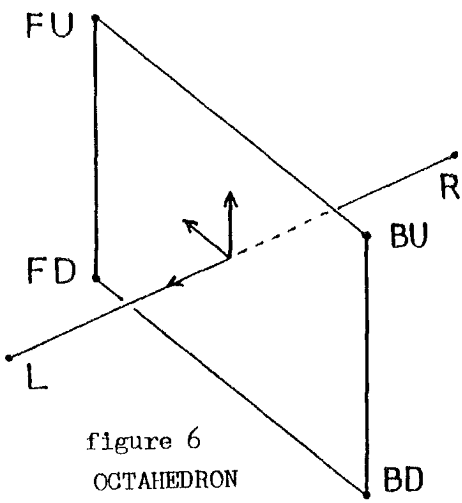

figure 6
OCTAHEDRON

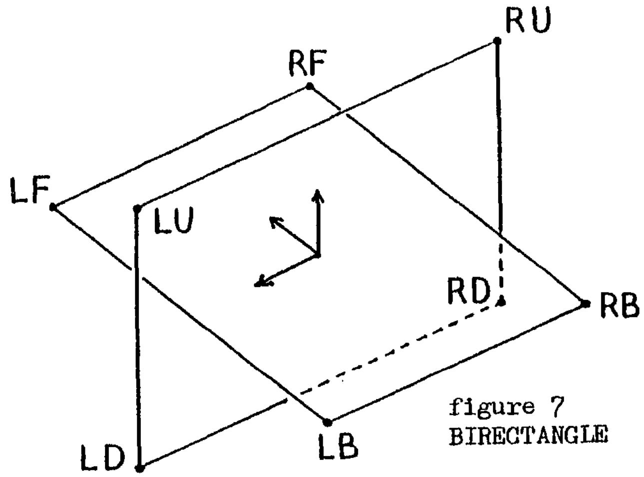

figure 7
BIRECTANGLE

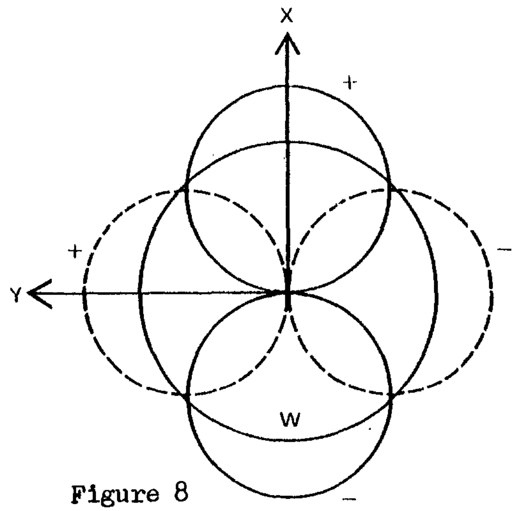

Figure 8

B - Format, Z Points up.

Now we discuss the process of deriving suitable speaker feed signals for layouts such as the one shown above. The signals required to feed a periphonic decoder must effectively treat all directions equally, and suitable signals are the 4 channels of the ambisonic B-format. This consists of 4 signals W,X,Y,Z (see figs 8 & 9) such that W is an omnidirectional signal, i.e. one containing sounds from all directions with equal gains, and X, Y, Z are three figure-of-eight signals (i.e. incorporating sounds with a cosine-law directional gain characteristic) pointing respectively forward, leftward and upward. In order to make B-format signals carry more-or-less equal average energy, X,Y,Z have a gain of $\sqrt{2}$ in their directions of peak sensitivity.

Such B-format signals can also be derived via a suitable phase -amplitude matrix from 4 periphonically encoded signals in encoding systems such as UHJ (Universal HJ) which incorporate a 4-channel periphonic specification. The overall form of a decoder for periphony is shown in fig.10. This shows an input matrix to derive B-format, followed by shelf filters to modify the low frequency vector magnitude r_V as the frequency increases so that the energy vector magnitude becomes optimal at higher frequencies. Suitable shelf filtering characteristics

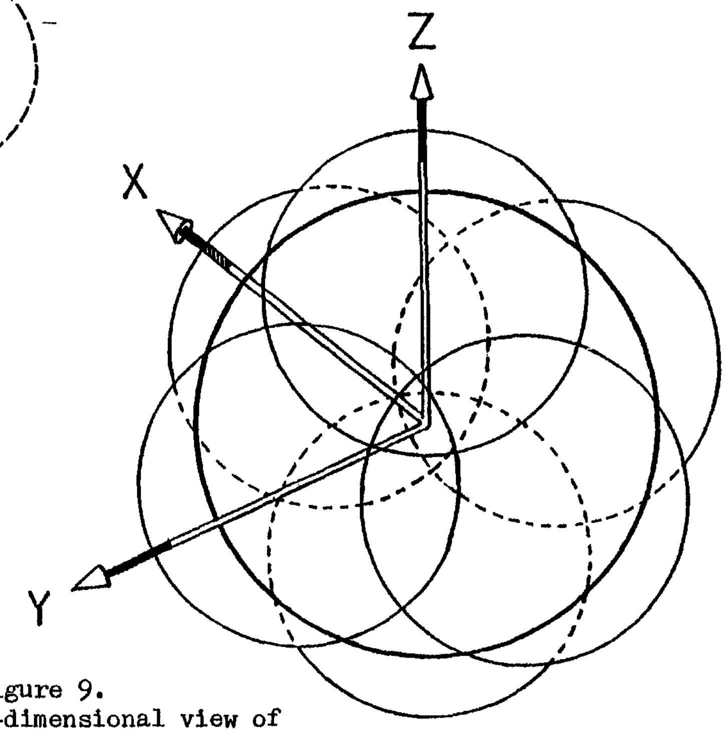

Figure 9.
3-dimensional view of
B-format.

are shown in figure 11, both for conventional horizontal ambisonic decoders and for full-sphere loudspeaker layouts. Note that the shelf filters required in the two cases are different. The decoder of fig. 10 then incorporates high pass filters (acting at about 20 Hz) to compensate for the finite distance of loudspeakers. This is a standard feature of ambisonic decoders which has been described elsewhere. The output amplitude matrix has to be adjusted to the shape of loudspeaker layout in use. In practice, a convenient implementation shown in fig. 10 is to have the output matrix

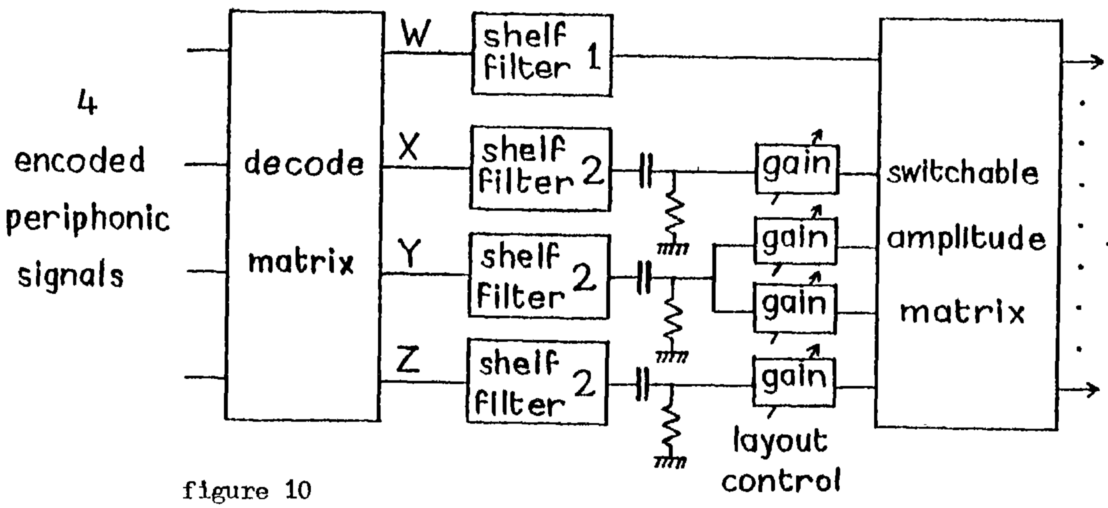

figure 10

switchable for the basic type of layout (e.g. figs. 5, 6 or 7), but to implement the continuously variable gains required to trim the decoder to the precise layout shape as a "layout control", i.e. as potentiometers in the X,Y and Z signal paths.

A decoder of the type shown in fig. 10 can be designed so that the requirements of the diametric decoder theorem are satisfied and such that the Makita (and hence also energy-vector) localisation coincides with the original localisation in the B-format signal. Some involved matrix analysis shows that the output amplitude matrix, including the setting of the layout controls, can be computed using the formulas shown in figure 12. We shall not attempt to prove this here.

Finally, we mention some of the design trade-offs in periphonic reproduction. Ideally, as we mentioned earlier, the energy vector magnitude r_E should equal one for ideal image stability. For horizontal ambisonic decoders working off B-format, it is not actually possible to have an r_E averaged over all horizontal directions that exceeds 0.707, although a suitable choice of horizontal speaker layout allows r_E to

figure 11

SHELF FILTER GAINS in dB		L.F	H.F.
HORIZONTAL LAYOUT	W	0.00	+1.76
	X & Y	0.00	-1.25
FULL-SPHERE LAYOUT	W	0.00	+3.01
	X, Y & Z	+ 1.76	0.00

TRANSITION AT APPROX 400 HZ

be increased to (say) 0.8 in some directions at the expense of being decreased (to say 0.6) in others. The 0.7 average r_E for horizontal decoders is quite satisfactory in practice. For 2-channel surround-sound material, a fixed horizontal decoder cannot have an average r_E of more than 0.5, which gives rather poor image stablity, although ingenious distribution of this fault around the circle of directions can mitigate it to some degree.

Full sphere reproduction from 4 B-format signals unfortunately can be proved to have an average r_E over all directions not exceeding 0.577, which is perilously close to being unsatisfactory. Thus it is important (far more so than in the horizontal case) to use shelf filtering carefully to optimise r_E. It is also important to choose the shape of the speaker layout to distribute the actual values of r_E in different directions in a manner optimising overall subjective results. In particular, a trade-off of r_E in different directions can be chosen (e.g. as in fig. 13) so as to give a rather smaller value of r_E in the vertical direction than horizontally. Such a trade-off involves a careful choice of speaker layouts.

THE DESIGN MATHEMATICS :-

LET n DIAMETRIC SPEAKER PAIRS LIE IN THE DIRECTIONS

$$\pm(x_i, y_i, z_i)$$

FOR $i = 1, 2, \ldots, n$. THEN THE RESPECTIVE SPEAKER FEED SIGNALS ARE

$$S_i^{\pm} = W \pm (\alpha_i X + \beta_i Y + \gamma_i Z)$$

WHERE

$$\begin{pmatrix} \alpha_i \\ \beta_i \\ \gamma_i \end{pmatrix} = \sqrt{\frac{1}{2}}\, nk \left[\sum_{j=1}^{n} \begin{pmatrix} x_j^2 & x_j y_j & x_j z_j \\ x_j y_j & y_j^2 & y_j z_j \\ x_j z_j & y_j z_j & z_j^2 \end{pmatrix} \right]^{-1} \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}$$

WHERE $k = 1$ AT LOW FREQUENCIES

Figure 12

ENERGY VECTOR TRADE-OFF

HORIZONTAL

$r_E = 1/\sqrt{2} = 0.71$ ALL ROUND OR

$r_E \simeq 0.8$ FRONT/BACK

$\simeq 0.6$ LEFT/RIGHT

FULL-SPHERE

$r_E = 1/\sqrt{3} = 0.58$ ALL ROUND OR

$r_E \simeq 0.69$ FRONT/BACK

$\simeq 0.58$ LEFT/RIGHT

$\simeq 0.39$ UP/DOWN

Figure 13