

Surround sound decoders — 6

Multi-system ambisonic decoder

1 — Basic design philosophy

by Michael A. Gerzon, M.A., *Mathematical Institute, Oxford*

This series of articles describes a decoder capable of decoding all major existing and proposed two-channel surround-sound systems, including the Ambisonic System 45J, SQ, Regular Matrix, BMX and BBC Matrix H. For systems other than SQ, the decoder gives full psychoacoustically optimized results using NRDC Ambisonic decoding technology. In addition, Ambisonic playback of mono, stereo and of three-channel studio-format signals is provided. The decoder is suitable for three-amplifier/four-speaker, four amplifier/four-speaker, and four-amplifier/six speaker reproduction.

The NRDC Ambisonic project has developed a comprehensive technology for creating, encoding and decoding sound¹⁻⁴. While this ambisonic technology can give of its best only with optimized programme material and encoding (i.e. the System 45J described in reference 1), the methods can be applied to getting improved results from nearly all existing surround-sound systems. Later articles in this series will give the detailed circuit and some constructional details for a decoder for all the above-listed two-channel systems. As this decoder does a great deal more than any previously-published decoder designs, it is necessary to describe its many facilities in some detail, as well as giving some idea of what the circuitry is intended to do.

The aim of any surround-sound decoder is to provide the listener with an illusion of sounds coming from all (horizontal) directions around him. Moreover, if the decoder is well designed, the directions should be those intended by the recording engineer, and should be heard by a listener through the usable listening area. Conventional "quadraphonic" decoder designs give very poor images for sounds in inter-loudspeaker directions, especially at the sides, resulting in a rather gimmicky 'ping-pung-pang-pong' effect at the four corners.

The full theory used to design ambisonic decoders is mathematical, and computing facilities are required to

carry out the extremely complex design calculations involved. It is clearly not possible to give full details here of why the various parts of the decoders have the exact values that they do, but some idea of what is going on can be given.

Two previous articles in *Wireless World*^{1,5} have described some, but not all, of the psychoacoustics lying behind ambisonic decoder design. Essentially, the ears use different methods of localizing sounds, not just one or two. The more of these hearing mechanisms that are satisfied the better the sound and accuracy of the result. In particular, if a decoder satisfies several different methods of hearing, the brain has to do far less work to unscramble the complex sound reaching the ears, and the result is particularly "relaxed" listening, with little listener fatigue. While the mathematics of the design is aimed at getting accurate sound localization for all directions, it is undoubtedly true that the biggest musical benefit comes from this consequent low listening fatigue, rather than from any ability to "shoot the pianist".

Some aspects of sound that ambisonic decoders are designed to optimize are now described. At low frequencies, below 500 or 700Hz, there are three important aspects of sound localization: the "Makita" direction of a sound (the direction one turns to to face the apparent sound direction), the "velocity magnitude" (the degree to which the sound stays in its correct localization as one turns to other directions), and the "phasiness" (the degree to which unwanted components of sound not in phase with the desired sound are heard). It turns out that for all systems other than SQ, it is possible to design a decoder matrix below 700Hz to get the Makita localization correct for all encoded sound directions. In addition, a careful adjustment of the gain of the various signal components in the decoder permits the velocity magnitude to be made correct also. Thus, at low frequencies, a listener will hear all directions correctly reproduced in direction, no matter which way one faces.

Phasiness is more of a problem with two-channel systems, as it is not

possible to design decoders that get rid of it altogether. The effect of phasiness is not only to blur the sound image, but also to create an unpleasant sensation often described as "pressure on the ears" that actually makes some people feel sick, although others don't seem to notice it much. Studies by the BBC^{6,7} and NHK⁸ have given a good indication of how much phasiness can be tolerated. In addition, it is found in practice that phasiness is more acceptable for sounds behind the listener than for sounds in front. Two-channel ambisonic decoders are therefore designed for very low phasiness in the front sector of sounds, while giving rather higher phasiness in the rear.

There is another reason why phase shifts cause poor directional reproduction that comes into action below 300Hz. In a real-world listening room, the loudspeakers are at a finite distance from the listener (often about 2.5 metres for British listening rooms), which means that the sound wave from each loudspeaker arrives as a curved wave-front at the listener, rather than as a plane wave. This curvature can be shown to cause the "phasiness" components of the reproduced directional sound to be converted into rotations of sound images around the listener at low frequencies. However, it is possible to remove these low-frequency errors by means of two high-pass filters in the "velocity signal" paths in the decoder. These speaker-distance compensation filters are RC types with -3dB points at about 20Hz for 2.7m speaker-to-listener distance. This may seem too low to worry about, but listening tests here confirmed the design theory and show that image displacements of as much as 15 to 30° can occur for instruments such as double basses when situated behind the listener unless distance compensation is used. Distance compensation does not turn a bad decoder into a good one, but it does give a "tighter" and more well-defined sound to an already good decoder design.

The ears use different methods of locating sounds at higher frequencies say from 700Hz to 5kHz. However, a rather magical result emerges from the

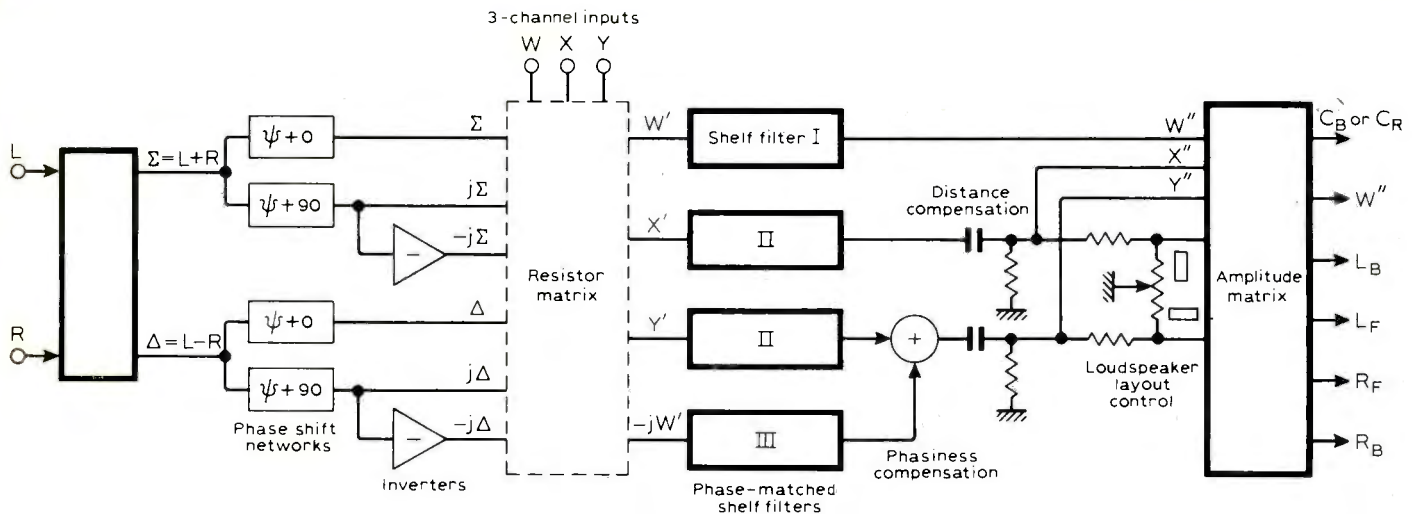


Fig. 1. Block diagram of multi-system ambisonic decoder, switching arrangements omitted. Shelf filters, inoperative for SQ decoding, depend on system being decoded, as does the resistor matrix. Also not shown is switching for C_B or C_R output.

design theory that states that, in effect, the basic sound localization of a decoder will automatically be the same at low and higher frequencies provided that the loudspeaker outputs of the decoder are derived via a particular type of amplitude matrix, the matrix involved depending only on the loudspeaker layout being used by the listener.

Besides getting the basic high frequency localization correct, it is necessary to minimize phasiness in this frequency range also, and to ensure that the sound image does not move around as the listener faces other directions. To get this last requirement right, it turns out that the best decoder design at higher frequencies involves different relative signal gains from those apt at low frequencies, so that the decoder has to be made frequency-dependent. The

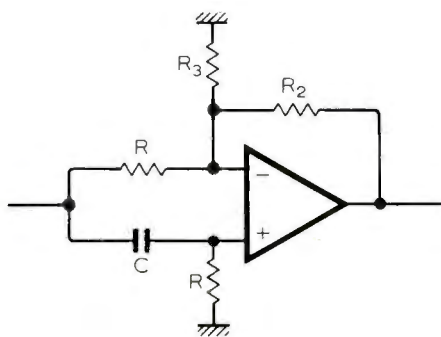


Fig. 2. Phase-compensated shelf-filter circuits allow frequency-dependent decoding. Conventional RC shelf filters would cause unwanted phase differences between signal paths.

effect of minimizing the image movement as the listener rotates his head is to avoid an unpleasant "in-the-head" sensation often, but incorrectly, described as "closeness" of sound by other authors⁹.

There are numerous other detailed aspects of decoder design, particularly those involving the way reverberation is reproduced (where a careful choice of encoding system such as 45J¹ can help), the effect heard by listeners seated away from the centre of the listening area, and the tone quality of the sound. It may seem strange that absolutely flat frequency response reproduction can sound coloured when reproduced through several speakers, the coloration depending on the precise speaker feeds used in the decoder. Many simple "matrix quadraphonic" decoders suffer from a tubby bass or harsh treble due to these effects. However, it is possible to account for most of these effects by the psychoacoustic design theory, and to minimize them in the decoder design. In practice, sounds encoded at the back are allowed to sound a little more colored than frontal sounds in ambisonic designs, because the ears appear to be more tolerant of marginal faults at the back, although one should assume that they are infinitely tolerant!

All decoder designs for two-channel encoding systems are a compromise between conflicting factors, and no design can achieve perfect performance in all ways. In this respect, the design of decoders is akin to loudspeaker design in being in the final analysis an art based on experience and listening. However, the science (i.e. the comprehensive psychoacoustic theory) is a very necessary part of reducing the almost infinitely complex design problem to a point where the designer can be sure of achieving his particular compromise as well as possible. The compromises inherent in these designs are based on the requirements:

- for front-stage material, the surround-sound should be subjectively superior to stereo for musical listening

(few existing designs meet this minimal requirement!),

- good results for listeners facing non-frontal directions and in non-central listening positions, especially behind-centre, and
- "musicality" of effect on both "ambient" and "surround" programme material, leading to low listening fatigue.

To some extent these requirements conflict with those based solely on the localization of direct sounds, such as in the experimental results quoted in ref. 10, where image sharpness for a forward-facing central listener has been achieved in a simple matrix decoder at the expense of "in-the-head" sound and severe image mislocation for non-forward-facing and non-central listeners. Good single-sound localization for most directions and listening positions can be achieved using a signal-actuated variable matrix decoder, but such decoders give a high level of listening fatigue on music due to the constant variation of signal parameters. Such decoders may be useful for surround-drama, where accuracy of localization becomes more important than "musicality" or low listening fatigue, and a fully-fledged ambisonic 'variable matrix' design is under development for such specialist applications. However, there is no doubt that a non-variable decoder is going to remain the preferred method of serious listening to music despite its superficially less "impressive" performance.

The basic diagram of the ambisonic decoder to be described in detail in later parts is shown in Fig. 1. Left and right signal inputs are fed to a sum-and-difference matrix to derive the sum $\Sigma = L + R$ and difference $\Delta = L - R$, because this leads to simplification of the later parts of the decoder, as well as to a slightly greater tolerance to small component errors. These two signals are each fed to 0° and 90° phase shift networks, and the 90° -shifted signal is also phase inverted to yield a -90° phase shifted signal in each case. The

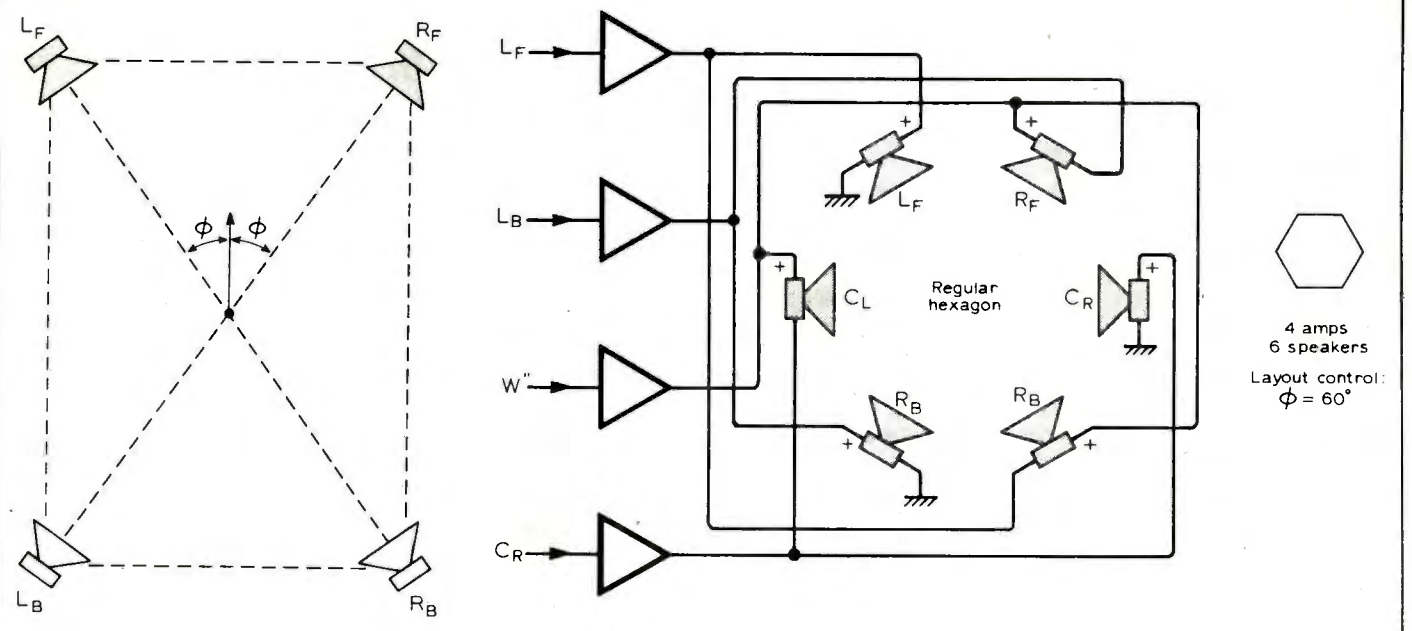
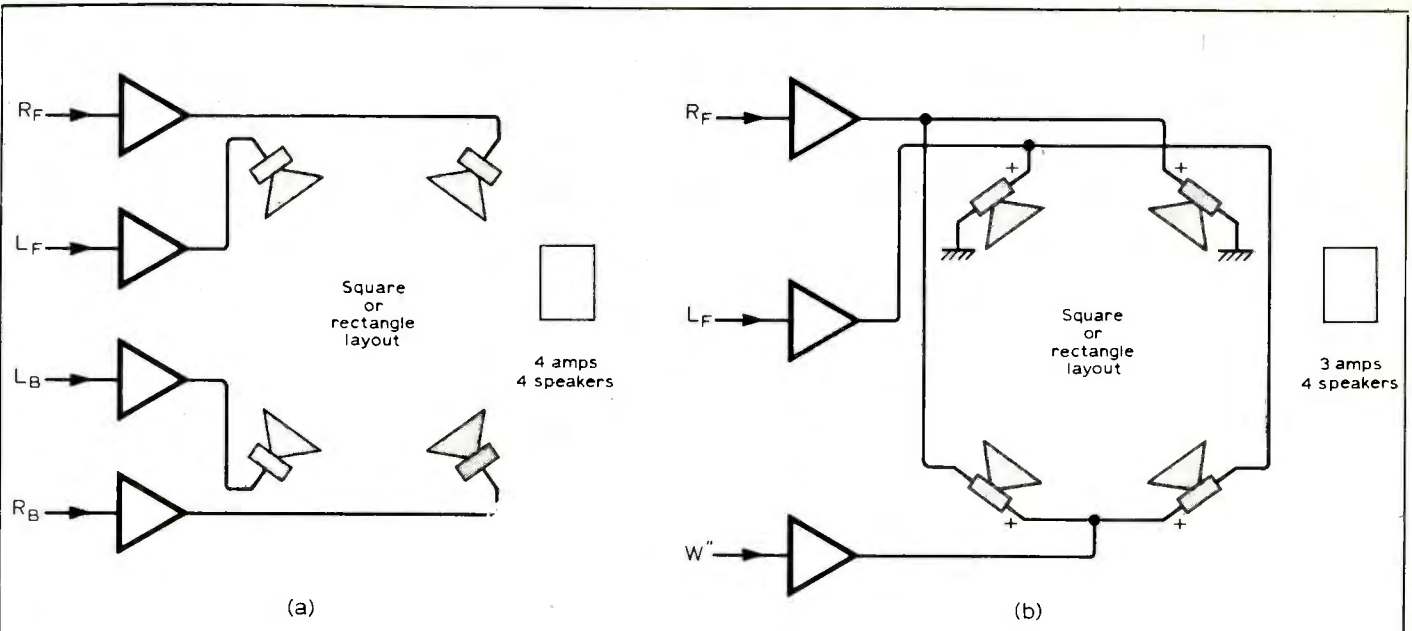


Fig. 3. Methods of feeding four loudspeakers using four amplifiers (a), and three amplifiers (b). Speaker terminals marked + are the positive-phase terminals in each case.

Fig. 4. Compensation is provided for non-square layouts. Angle ϕ is set on a layout control.

Fig. 5. Better results can be obtained from four-amplifier, six-speaker regular hexagon decoding. Connections are shown for two hexagon layouts with the angle ϕ used in the equations of the output matrix.

phase-shifters used are high precision types, as the ears are capable of hearing very small errors in localization (as little as 2° in real life). Previous "quadrophonic" decoders have not required such high precision mainly because they gave in any case a poor decoded effect due to sub-optimal design.

The six signals are then fed to a resistor matrix, which derives the required combination of these signals to produce the correct pressure and velocity signals W' , X' , Y' , for the particular encoding system in use. (For a discussion of this aspect see ref. 1, in particular in connection with its Fig. 2). The resistor matrix used is different for different encoding systems, so that switching is provided for different matrix resistance values. The resistor matrix, which involves no active circuitry, also includes a switched three-channel input option suitable for use with three-channel ambisonic mastertapes. In a later article we hope to describe live ambisonic recording for the keen tape enthusiast. These three channel inputs only cost a few resistors and input sockets in the present decoder, and so come virtually for free; in addition, they provide useful test signal inputs for setting-up purposes. We shall give the resistor matrix formulas for the various encoding systems for the signals W' , X' , Y' in part 2 of this series. An output $-jW'$ is provided for phasiness-control in some systems, as described in references 1 and 3.

The frequency-dependent aspects of the decoder are provided by the shelf filters which give one decoding matrix at low frequencies and a second at high frequencies, with the transition centred at 400Hz. Were conventional RC shelf filters to be used, there would be phase shifts between the various signal paths, which would cause quite bad localization errors. For this reason, the shelf filters are designed to give phase shifts identical to one another by making them "all-pass" types. The basic circuit of the phase-compensated shelf filters is shown in Fig. 2. The particular arrangement shown has an input impedance of R at all frequencies, which means that it is seen by the resistor matrix as a resistive load, suitable for terminating a matrix circuit. The value of R_2/R controls the ratio of high-to-low-frequency gain of the shelf, and R_3 provides extra h.f. gain to make up the losses of the preceding resistor matrix.

Thus the shelf filters are made to do five different jobs: terminate the matrix circuit, provide gain, give a different matrix circuit at low and high frequencies, give matched phases over the transition frequency band, and give an overall flat frequency response to the decoder at all frequencies.

An additional complication arises because different methods of encoding require different shelf filters in the decoder. In practice, the shelf filters required for BMX, RM, 45J and BBC H

differ only slightly, so that a compromise choice has been made to do all these systems. Decoding mono, stereo and three-channel studio format requires, for best results, a different set of shelf filters, and SQ requires that no shelf filters be used. (SQ decoders cannot be designed to give full ambisonic results; there is a mathematical theorem to this effect. The decoder for SQ provided is, however, less phasey in quality than the SQ designs on the market, and was designed specifically for incorporation into this design. It is not in accordance with CBS Laboratories' SQ specification, but in the author's opinion, it is better than decoders that are.)

The switching of the shelf filters involves equipping the op-amps of Fig. 2 with several filter circuits, which are switched in and out as required.

The outputs of shelf filters acting on Y' and $-jW'$ (see Fig. 1) are added to reduce front-stage phasiness, and the velocity signals are then subjected to the RC high-pass distance compensation. This gives us three signals W'' , X'' and Y'' representing respectively the signal pressure, forward component of acoustic velocity, and leftward component of acoustic velocity heard by the listener. These are fed to an output amplitude matrix, which includes a layout control adjustment to adjust the outputs of the decoder to match different shapes of rectangular loudspeaker layout in the room.

The decoder provides six different outputs L_B (left back), L_F (left front), R_F (right front), R_B (right back), W'' (pressure) and either C_B (due back) or C_R (due right), switched. The way these six outputs can be used is itself an interesting story, for they can be used to provide decoding via four loudspeakers in a wide range of rectangle shapes using either four amplifiers, Fig. 3(a), or, remarkably, using just three power amplifiers as shown in Fig. 3(b). The three-amplifier set-up in no way means that there is any compromise in the psychoacoustics of the decoded signal, as precisely the same speaker signals are produced as in Fig. 3(a)!

To see this, we first remark that the outputs of the decoder are given by the formulae

$$\begin{aligned} L_B &= \frac{1}{2}(W'' - \sqrt{2} \sin \phi X'' + \sqrt{2} \cos \phi Y'') \\ L_F &= \frac{1}{2}(W'' + \sqrt{2} \sin \phi X'' + \sqrt{2} \cos \phi Y'') \\ R_F &= \frac{1}{2}(W'' + \sqrt{2} \sin \phi X'' - \sqrt{2} \cos \phi Y'') \\ R_B &= \frac{1}{2}(W'' - \sqrt{2} \sin \phi X'' - \sqrt{2} \cos \phi Y'') \end{aligned}$$

where ϕ depends on the setting of the layout control, being 45° for a square layout, and being equal to the angle ϕ shown in Fig. 4 for a rectangle layout. From these formulae

$$L_B + R_F = L_F + R_B = W'',$$

so that

$$\begin{aligned} L_B &= W'' - R_F \\ R_B &= W'' - L_F, \end{aligned}$$

and it will be seen that the rear speakers of Fig. 3(b) indeed are connected so that the potentials of their "positive phase" terminals relative to their negative phase terminals are $W'' - R_F$ and $W'' - L_F$

respectively.

Even more remarkable however, are the four-amplifier six-loudspeaker arrangements possible with this decoder. It has been known for several years that decoders using six loudspeakers are capable of better results than is possible using four, no matter how well-designed the decoder may be. If properly used, the extra speakers give more solid image location over a larger area, with less tendency for the image to hug the loudspeakers than when using four, particularly on difficult waveforms such as audience applause. It has not been possible to market six-speaker equipment; few homes could properly accommodate it, and the market for such special equipment was thus too small to justify manufacture. However, the ambisonic decoding method permits the same decoder and the same four amplifiers to be used for six speakers for the few who can manage it, making this improved form of decoding domestically available for the first time. We emphasise that in no way does the use of four amplifiers imply substandard results: exactly the same speaker signals are given as one would design a psychoacoustically optimized six-amplifier decoder to give.

The six-speaker connections for two shapes of regular hexagon layout are shown in Fig. 5. The three speakers that are fed in a "simple" manner in each case form an equilateral triangle of speakers; this helps minimize the subjective effects of slight mismatches of amplifier gain. The signals C_B and C_R are

$$\begin{aligned} C_B &= \frac{1}{2}(W'' - \sqrt{2} X'') \\ C_R &= \frac{1}{2}(W'' - \sqrt{2} Y''). \end{aligned}$$

Although detailed instructions for calibrating and using the decoder will be given at the end of this series of articles, it is worth emphasizing now that all amplifiers and loudspeakers

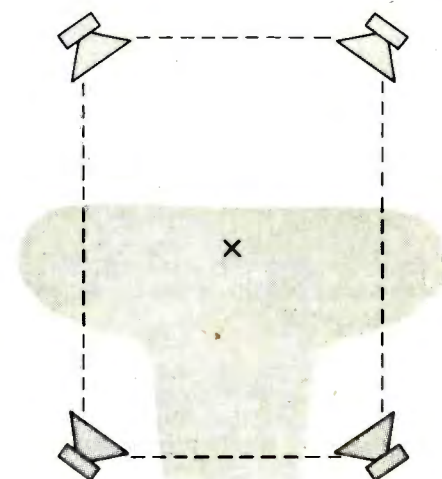


Fig. 6. Approximate listening area for ambisonic decoding (shaded) with a rectangle speaker layout obtained for BMX, 45J, Matrix H and RM systems as well as stereo. Optimal listening is at the centre (X).

must be accurately matched for correct ambisonic results. Unlike "quadraphonic" decoders, both front and rear loudspeakers co-operate to produce sounds in any direction. Thus, for example, the rear speakers provide outputs that help to reinforce the localization of sounds that are reproduced in front of the listener. Thus one cannot try turning the rear speakers up or down in the mistaken idea that the front and rear are independent of one another. When the outputs are not matched, the sound field tends to "fall apart"; in fact turning down the rear speakers often makes them much more audible (as distracting noises at the back) than in a correct balance.

While it is not absolutely necessary to have all power amplifiers of the same make, they should be adjusted for identical gains and phases, and one should check (e.g. by using an X-Y oscilloscope display) that they have substantially identical phase responses over the audio band. If in doubt, identical amplifiers should be used. While identical speakers should be used, a small number of speaker manufacturers (e.g. IMF Electronics) have taken trouble to match the different models in their range carefully, and in such cases different models can be used at front and rear. Again, if in doubt, use identical speakers for best results.

The decoder not only reproduces surround sound from a variety of existing systems, but also handles mono and stereo, using ambisonic techniques to get the most natural possible reproduction (using four or six speakers) from existing records and broadcasts. Except for exceptional stereo material, the decoder does not create "pseudo surround sound", but reproduces stereo over a conventional frontal stage with a subtle enhancement over two-speaker stereo, and without any gimmickry. The mono decode mode reproduces a mono source from straight in front, but the

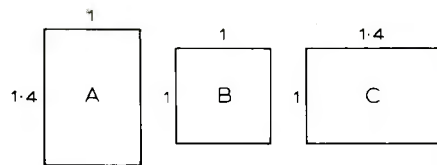


Fig. 7. Most stable front images are provided by arrangement (a), most stable side images by (c), while (b) is a compromise between these extremes.

rear speakers help to lock the image solidly in space behind the front loudspeakers. Neither mono or stereo decode modes enhance badly recorded material, but neither do they degrade it any further. In practice, many ambient SQ classical recordings, such as those released by EMI, will be found to reproduce better in the stereo decode mode than via SQ decode mode, owing to the inherent limitations of the SQ system.

Fig. 6 shows the approximate usable listening area for most decoding modes (excluding SQ) in a typical domestic room using a rectangle layout. The listening area will in practice depend on the loudspeakers used, the room acoustics, the layout shape used, the programme material, and also on the system being decoded. The type of listening area shown has been obtained both with ambisonic recordings made in concert halls, and with commercial "easy listening" (*sic*) music in the BMX, 45J, Matrix H and RM systems, as well as with stereo material played in stereo decode mode.

It is found that a longer-than-wide layout of four speakers as in Fig. 7(a) gives the most stable front images for non-central listener and the least stable side images. A wider-than-long layout as in Fig. 7(c) gives excellent stable side images for most listeners, although the front stage tends to be drawn over to the nearest speaker. A square layout,

Fig. 7(b), is a compromise in terms of image stability between these extremes. Extremely thin rectangles (whether long or wide) cannot be expected to give good results, although the layout control adjustment will help to minimize the inevitable defects.

Part 2 will give details of the decoding matrices used. Patent rights in circuits described in this and subsequent parts of this article are owned by the National Research Development Corporation. A kit of parts for the decoder will be available from Integrex Ltd - see advertisement.

References

1. Gerzon, M.A. NRDC surround-sound system. *Wireless World*, vol. 83, April 1977, pp.36-9.
2. Gerzon, M.A. Ambisonics Part II: Studio techniques, *Studio Sound*, vol. 17, Aug. 1975, pp.24-6. 28, 30, Correction: *ibid* vol. 17, Oct. 1975, p.60.
3. Gerzon, M.A. Optimum Choice of Surround-Sound Encoding Specification, preprint A5, 56th AES Convention, Paris, 1st March, 1977.
4. Gerzon, M.A., Design of precisely coincident microphone arrays for stereo and surround sound, in collected papers, 50th AES Convention, London, March 1975.
5. Gerzon, M.A. Surround sound psychoacoustics, *Wireless World*, vol. 80, 1974, pp.483-6.
6. Bower, J. S. Subjective effects of inter-channel phase-shifts on the stereophonic image localization of wideband audio signals, BBC Research Department Report, BBC RD1975/27, Sept. 1975.
7. Bower, J. S. Subjective effects of inter-channel phase-shifts on the stereophonic image localization of narrowband audio signals, BBC Research Department Report, BBC RD1975/28, Sept. 1975.
8. Nakabayashi, K. Method of analyzing the quadraphonic sound field, *J. Audio Eng. Soc.*, vol. 23, 1975, pp.187-93.
9. Meares, D. J. and Ratliff, P.A. Development of a compatible 4-2-4 quadraphonic matrix system. BBC Matrix H, *EBU Review* (Technical) no. 159, Oct. 1976.
10. Ratliff, P. A. and Meares, D. J., BBC Matrix H. *Wireless World*, vol. 83, May 1977, pp.41-5.

Broadcasting satellite receivers

Mullard's research laboratories have released details of the microwave receivers which they, in co-operation with Philips at Eindhoven, have built for picking up broadcast television signals from "Hermes", the Canadian-American Communications Technology Satellite (News, April, p.40). The receivers convert the 12GHz f.m. broadcast signals picked up by small dish aerials to a form suitable for feeding into the aerial socket of standard NTSC colour television sets. Five receivers and associated 1.6m (or in some cases 1.2m) diameter metal-coated glass-fibre-reinforced polyester parabolic aerials were supplied by Philips, and one receiver of somewhat different design but using the same aerial was

supplied by Mullard. To accommodate satellite drift provision was made for remote control motor tracking of the aerials on a single axis.

An outdoor part of the receiver with an integral waveguide horn is placed at the focus of the aerial and provides fixed-tuned conversion from 12GHz to 410MHz, with a noise figure of about 7dB. It consists of a microstrip Schottky-barrier-diode mixer, followed by a 40dB amplifier. The local oscillator is a Gunn device delivering about 10mW at approximately 11.7GHz and, in the Mullard version, this is located in an aluminium cavity integral with the converter. Dielectric temperature compensation maintains frequency drift within 5MHz (over the range -40°C to $+40^{\circ}\text{C}$) which is well within the range of the automatic frequency control applied to the second mixer. This is located in an indoor unit fed with the 410MHz signal by a coaxial cable.

The indoor part of the receiver provides further conversion to 120MHz prior to limiting and frequency discrimination to yield the NTSC composite video colour signal and 5.14MHz f.m. sound signal. The last-mentioned is converted to the normal 4.5MHz inter-carrier frequency and with the video is used to amplitude-modulate a 500MHz carrier to provide a 10mV signal suitable for the aerial input of a standard NTSC colour television set. Extensive use is made of integrated circuit techniques.

The Mullard receiver is installed at the Canadian government communications research centre in Ottawa. The aerials were aligned first by a simple level of compass then final adjustment was made by means of the satellite signal itself and a signal strength meter. Mullard say excellent picture and sound quality was achieved and the received signal strength was close to the expected level of -105dBW .