Proceedings of The Institute of Acoustics

AMBISONIC SURROUND SOUND

P.B. Fellgett

Dept. of Cybernetics, University of Reading, Reading, England.

SOME HISTORY

Directional effects in reproduced sound seem to have been first noticed, apparently accidentally, at the Paris Electrical Exposition of 1881 [1]. Ear-pieces were provided each fed from a separate carbon microphone, indeed in the absence of amplifiers at that time there was virtually no alternative. The microphones were placed in the neighbouring Opera and the directional effects were noticed when one ear piece was applied to each ear.

The observation remained a curiosity for many years. Even when the idea of 'high fidelity' reproduction began to replace the 'mellow-toned radiogram' from 1940 onwards, there was little or no public awareness of the need to reproduce more than monophonically. Meanwhile however, the famous Philadelphia-Washington relay of 1933 [2] had publicly demonstrated directional effects. These were based on multi-channel spaced-microphone and loudspeaker methods. In England, Alan Blumlein had by 1931 [3] developed a systematic technology of stereo reproduction based on what are now called coincident-pair directional microphone methods. It is only comparatively recently that Blumlein's genius has come to be more fully recognised, and on re-reading his work in the light of modern knowledge one is astonished at the depth of his insight into audio. Nor was his contribution confined to audio; it was he who was largely responsible for the development of high-definition cathode-ray television in Great Britain, replacing the earlier Baird mechanical system.

Monophonic recording and broadcasting remained the norm throughout the change-over from the 78 rpm shellac pressing to the modern microgroove disc, and stereo was not effectively introduced into the market until the 1960s. Even then, 'stereo' remained for a considerable time a cult word with sybaritic connotations, record companies kept dual mono and stereo inventories, and record shops had separate stereo sections. Today of course it is taken for granted that equipment and records are stereo unless otherwise stated. Technically, this means that two audio channels are available; a point to which we shall return.

The inherent limitation of stereo is that it provides directional information only over a frontal soundstage limited in practice to about 60°. In seeking to expand the audio market, several companies had the idea of extending directional information to all around the listener. This is defined as *surround sound*. The introduction of stereo had been accompanied by some inter-company confusion, and with hindsight it appears that in order to avoid being left behind firms introduced surround sound systems prematurely before the necessary research and development had been sufficiently completed.

These early attempts at surround reproduction were known by the Latin-Greek hybrid name 'quadraphonic'. Substantial news of them began to appear in magazines etc. in 1970, but it was immediately apparent (at least to a number of people in the world) that this attempt suffered from defects of assumption, aim and method which would militate against satisfactory performance. Indeed over the next decade or so the various 'quadraphonic' systems failed to gain a satisfactory market share, and the industry suffered substantial losses in consequence.

The present author was stung by this situation to formulate ideas leading towards a more soundly-based technology for surround reproduction. The academic tradition is, and rightly so, for the free dissemination of knowledge. It was however concluded, with some hesitation, that in this case practical benefits would be realised only through commercial exploitation. The ideas were therefore submitted, as was customary at the time, to the National Research Development Corporation (NRDC) resulting in two British patents being assigned in 1971 and 1972 to NRDC under a revenue-sharing agreement [4]. The author was later joined by industrial partners, Mr. David Brown, Mr. John Hayes and Mr. John Wright, and later Mr. Michael Gerzon of the University of Oxford, who had independently formulated similar ideas, joined the partnership and became its theoretician. The practical work remained in the Department of Cybernetics at Reading, supported mainly by NRDC.

As a result of this research over almost exactly a decade, knowledge became available for the new technology for surround reproduction which will be described briefly in what follows. The name 'Ambisonic' for this technology was coined by the present author and published by him in 1973 [5]. Note that 'ambisonic' is the adjective, and 'Ambisonics' the noun, strictly analogously to 'electronic' and 'Electronics'.

In the course of the research it was found that related and partly complementary ideas had been developed by Professor Duane Cooper, University of Illinois, USA, and by Dr. T. Shiga and others of the Nippon Columbia Company Japan. NRDC eventually arranged a patent-sharing agreement with these parties, thus obviating the possibility of technologically counter-productive conflicts.

The 'quadraphonic' experience turned the industry and public for some years against surround reproduction, and the industry has sought to expand more through the development of new media, such as digital audio and the compact digital disc (CD). These media provide of themselves no new facilities, although they are claimed to do the old things better than hitherto. Either they or the older stereo media are however compatible with Ambisonics. Indeed there are approaching two hundred ambisonic records issued in the regular commercial market [6]. Many of these, and other ambisonic material, have been broadcast by the BBC. Ambisonic broadcasts have also been made by the IBA, Netherlands Radio, and by National Public Radio in the USA. Ambisonic technology is also reported to be used by the Library of Congress for archive purposes. With increasing recognition that the new technology does not repeat older 'quadraphonic' mistakes, the time appears ripe for more widespread adoption of Ambisonics, and the author and his colleagues are hopeful that the matter is now in the effective hands of private enterprise.

ENCODING

The inception of the 'quadraphonic' attempt at surround reproduction seems to have been the observation that strong experiences could result from playing the four tracks of a multi-track tape into a corresponding set of four loudspeakers disposed about the listener. This was before the multi-track explosion, at a time when four-track tape was an industry standard. Three tracks had been the norm at a slightly earlier stage, and with hindsight one wonders whether surround technology would not have developed earlier and better had three-track ideas been in vogue at the relevant time. As it was, the 'four channel' ideas were reinforced by the undeniable fact that most domestic rooms have four corners, although not necessarily or even usually four corners into which it is domestically acceptable to place loudspeakers.

The problem thus appeared to revolve around sending four channels of audio signal to the domestic listener, either by multiplexing (as in carrier-disc recordings) or by pretending to have audio channels that did not physically exist; these latter were known as 'matrixed channels'. It was not noticed sufficiently explicitly that even had these stratagems been entirely successful, they were still insufficient to provide acceptable surround reproduction in general. The problem of how to use the available information to give the listener correct directional impressions had not been solved, or indeed adequately addressed.

So far as it was ever clearly stated, the idea seems to have been that each pair of adjacent loudspeakers would give phantom images of the stereo type, i.e. what may be deemed 'four-wall stereo'. Elementary observation shows, as has been frequently pointed out, that this cannot work satisfactorily. The stereo illusion is unsatisfactory, with a hole-in-the-middle impression and a tendency for the sounds to run to either speaker with the slightest movement of the listener, when the loudspeakers subtend as much as 90° of front sector. The illusion is still worse at the rear, and hardly operates at all at the sides. Moreover it was not clear how the loudspeakers should be fed from the original source. Four stereo-pairs of loudspeakers would require four stereo-pairs of microphones, giving a total of eight signals which have somehow to be distributed to the four recipient loudspeakers, and there seems to be no satisfactory way of doing this. Once it is understood that the need is not to encode four or any other number of 'channels' (which is semantically and technically impossible anyway) but to encode *direction*, matters become both simpler and more favourable. Two channels already provide two parameters which can be used for encoding direction, namely relative amplitude and relative phase. These quantities can conveniently be displayed on the Poincare-Stokes representative sphere [7] or its stereographic projection. Although in principle both altitude and azimuth angles could be encoded, subsidiary requirements confine two-channel surround systems in practice to horizontal surround only. The complete coding information is then represented by the position of the locus, on the representative sphere, corresponding to the azimuth circle, together with information on the correspondence between each azimuth and the corresponding point of the locus, and the relative loudness with which each azimuth is encoded. Evidently the encoding locus should have left-right symmetry, and the disposable parameters may be characterised as the amount of tilt of the mean plane of the locus, the extent to which it deviates from a great circle, how much it is bent out of plane, and (for smoothness) one parameter each for the distribution of azimuths and of loudness. In the ambisonic two-channel UHJ encoding, these design parameters have been carefully adjusted to optimize the overall performance in mono, stereo and surround playback. The stereo compatibility which results is evidenced by ambisonic recordings having received major awards *judged as stereo* without taking account of their additional surround capability.

The complete set of directly-compatible UHJ encodings include also two-and-a-half, three and four channel encodings with compatibility upwards and downwards by the addition or omission of channels, without modification to those that remain. The so-called two-and-a-half-channel SHJ employs a third channel of restricted bandwidth, the three-channel THJ employs the *optimum* number of channels for domestic horizontal-surround, and the four-channel PHJ can provide periphonic playback, i.e. full spherical surround including height information.

Ambisonic encoding is of course easily provided synthetically by ambisonic pan-pots etc., but the *Soundfield Microphone* (SFM) also provides direct transduction of live soundfields. Based on the theory of sampling on the surface of a sphere, the principle of the SFM has proved so good in overcoming earlier limitations that further development of capsule technology has been needed to exploit it to the full. A soundfield microphone is equivalent to any number of truly coincident first-order microphones (i.e. of the cardioid family) each independently steerable horizontally and in height. It thus provides unprecedented operational flexibility (either during a recording or broadcast, or afterwards if it is recorded) as well as high quality because of the avoidance of comb-filter effects of spacing, and because pressure and velocity responses can be equalised separately. This flexibility has been illustrated by broadcast use of a single SFM suspended in the Royal Albert Hall with unknown mechanical orientation (since it is fully steerable electrically), and its discrimination is illustrated by individual voices being discernible in a recording of a large chorus and orchestra using a single SFM in this venue.

DECODING

The task of the decoder is to feed to the loudspeakers of the eventual listener such feeds as will radiate into the room sounds which cooperatively give the listener correct directional clues. Since these feeds must be adapted to the size and shape of the loudspeaker layout including the number of loudspeakers used, they cannot have existed at any earlier point in the chain. Ambisonic decoders are equipped with simple controls, partly analogous to the 'balance' control of stereo, enabling the necessary adaptation to be made.

The human ear-brain system uses a number of mechanisms for locating sound, the nature of which needs to be understood for optimum design. The present author turned at first to the results of perception psychologists, but in retrospect this is seen as a mistake and the essential key lies in the physics of diffraction of soundwaves at the human head, and the conditions this imposes on the types of information available to the ear.

The key facts had already been established by Lord Rayleigh before 1907. In that year he wrote in the Philosophical Magazine [8] 'It is some thirty years ago since I executed a rather extensive series of experiments in order to ascertain more precisely what are the capabilities of the ears in estimating the direction of sounds'. Rayleigh had earlier [9] solved the then long-standing problem of why the sky is blue by developing the theory of scattering eponymous to him, according to which the energy of a wave scattered by a body small compared with the wavelength varies as the fourth power of that wavelength. The scattering cross-section of a body large compared with the wavelength approaches in the limit the geometrical cross-section of the body.

AMBISONIC SURROUND SOUND

The changeover between these two regimes is approximately when the circumference of the body is equal to the wavelength. For soundwaves impinging on the human head this occurs at about 500 Hz.

Using the principle that the total field is the sum of the incident field and the scattering, it follows that at low frequencies, where the scattering is small, the head is effectively in an undisturbed soundfield. Only small differences are then possible in the intensity experienced by the two ears, and the phase difference corresponds to the free-space propagation time between them. At higher frequencies, substantial intensity differences are possible, but the phase information becomes ambiguous. Because of the fourth-power law, the changeover between the two regions is quite sharp, occupying approximately an octave. Thus Rayleigh, quoting from his 'Theory of Sound' [10] concludes that the maximum possible difference of intensity at 250 Hz amounts to only (in modern terms) 1 dB, but can be over 3 dB at 500 Hz. He estimates that the ambiguous phase-difference of ± 180° occurs at about 600 Hz.

It follows that evolution has had no option but to equip us to use phase information below about 500 Hz, and intensity information above this frequency, if we are to have the power of localising sounds in both frequency ranges. It is only necessary to verify experimentally, as is easily done, that discrimination of direction is possible both above and below 500 Hz to prove that these must be the mechanisms employed in the respective frequency ranges; the physics permits no other possibility. Rayleigh indeed confirmed this conclusion by further experiments, and writes[8] 'the sensation of lateralness due to phase-difference disappears in the region of pitch where there would be danger of its becoming a misleading guide. ..... It is fortunate that when difference of phase fails, difference of intensity comes to our aid. Perhaps it is not to be expected that we should recognise intuitively the very different foundations upon which our judgement rests in the two cases'. Rayleigh also recognised a third region at still higher frequencies where discrimination, particularly as between front and back, 'depends upon an alteration of *quality* due to the external ears ..... In this matter it would not be surprising if individual differences manifested themselves .....a "paddle box" formation of the external ear, if not ornamental, may have practical advantages'. Such individual differences are of course now recognised as a difficulty with dummy-head recording techniques.

If azimuth angle θ is measured anti-clockwise from the front-centre position, the phase difference at low frequencies is evidently nearly proportional to sin θ, i.e. a L - R figure-of-eight. Slight rotation of the head about a vertical axis, often performed unconsciously and wrongly inhibited in some experiments in localisation, effectively differentiates this function to give cos θ, i.e. a front-minus-back figure-of-eight. Left-right tilt of the line joining the ears similarly gives information equivalent to the cosine of the altitude angle. It is for this reason that stereo phantom-images often appear above, or occasionally below, the line joining the loudspeakers, as was first described and explained by De Boer [11]; similar considerations apply *mutatis mutandis* in the frequency region above 500 Hz, with intensity replacing phase as the discriminant.

The psychoacoustic design of ambisonic decoders consists in putting together and optimising these facets, and in particular ensuring that the different mechanisms of localisation are presented with clues which agree with each other. Indeed Rayleigh himself reports [8] a series of ingenius experiments, both with tuning forks and with the then novel telephone ear-pieces, showing how the sense of direction disappears when this agreement is not achieved. To bring about this agreement, and to optimize the strength of the phase-clues at lower frequencies, and similarly the intensity-clues at higher frquencies, requires the characteristics of a well-designed decoder to change with frequency, and this is an especially significant feature of ambisonic decoders.

CONCLUSIONS

Ambisonics works for the same physical reasons as make the sky appear blue in colour. Perhaps this gives a new meaning to the phrase 'blue sky research'.

REFERENCES

[1] Clément Ader's 'Théâtrephone'. Ader was also a pioneer of aviation, claiming to have achieved take-off under steam power, but not sustained flight, in 1890. The claim is disputed, and an attempted demonstration with his 'Avion III' in 1897 failed to get off the ground.

[2] J.C. Steinberg and W.B. Snow, 'Symposium on Wire Transmission of Symphonic Music and its Reproduction in Auditory Perspective', Bell Syst. Tech. J., Vol. 13, 245-258, (1934).

[3] A.D. Blumlein, British Patent 394325 (1931).

[4] P.B. Fellgett, British Patents 3698/71 and 52008/71.

[5] P.B. Fellgett, 'Quadraphonics Queried and Ambiophonics Analysed',Hi-Fi News and Record Review, Vol. 18, No. 7, 1309-1311 (July 1973).

[6] A list of ambisonic recordings known to the author is available by sending an A4 SAE (100 grams) to the address at the head of this article.

[7] H. Poincaré, 'Theorie Mathematique de la Lumiere II', Georges Carré, Paris, 282-285, 302, (1892).

[8] Rayleigh, John William Strutt, 3rd Baron, 'On our Perception of Sound Direction', Phil. Mag., Vol. 13, 214-230 (1907).

[9] Rayleigh, 'On the Light from the Sky, its Polarisation and Colour', Phil. Mag., Vol. 41, 107, 274, (1871).

Rayleigh, 'On the Scattering of Light by Small Particles', Phil. Mag., Vol. 41, 447, (1871).

[10] Rayleigh, 'The Theory of Sound', London, Cambridge, Vol. I 1877, Vol. II 1878. 2nd Ed. (revised) Macmillan & Co., 1894-6.

[11] K. de Boer, 'A Remarkable Phenomenon with Stereophonic Sound Reproduction', Philips Tech. Rev., Vol. 9, No. 1, 8-13 (1947).