# A Geometric Model for Two-Channel Four-Speaker Matrix Stereo Systems*

**MICHAEL A. GERZON**

*Mathematical Institute, University of Oxford, Oxford, England*

Some of the properties of a convenient geometric model are described which is variously known as the "energy sphere," "Stokes' sphere," "Poincaré sphere," or "representation sphere," for describing two-channel information systems, such as matrix four-speaker systems. Several of its uses are outlined, including the study of directional ambiguities, compatibility, the behavior of variable-matrix systems, the invention of systems, and stereo imaging. Once the relationship between the sphere and more direct representations of two-channel systems is established, it proves to be possible to survey the system's properties with virtually no need for mathematical formulas.

**INTRODUCTION:** In the following we describe a geometric model for two-channel matrix systems that was independently discovered by Scheiber [1] and the present author [2], [3], and which in fact was first discovered in 1852 by Stokes [4] and in 1892 by Poincaré [5] in connection with polarized light.

The aim of this paper is describe how this model, which we shall term the "energy sphere," may be used to visualize a wide range of properties not discussed by previous authors, so that many of the properties of systems that are normally obscured in mathematical detail, such as compatibility, the effect of variable-matrix circuitry, stereo imaging properties, and the effect of using pairwise pan-potting, may be examined easily. Experience indicates that there is no other representation of information about matrix systems that presents so much information in such an easily graspable form, although naturally there are some types of information it does not describe conveniently.

In order to keep the paper within reasonable bounds, we omit proofs. Some results are proved in previous papers on this model [1], [2], [4]–[6], and some of the others are implicit in the physics literature on the rotation and Lorentz groups, notably [29], [24], and somewhat more obscurely but with more results, in [7] specialized to the spin ½ case.

We would emphasize that uses of the following model are by no means confined to four-speaker matrix systems, and it may be applied equally to stereo recording (e.g., see Appendix II), to the study of interference in double-sideband AM reception, or to any other situation in which two related channels of analog information occur.

The rest of this introduction largely repeats Scheiber's account [1], but in a terminology convenient for subsequent use.

Consider a signal recorded on two channels with a gain $\alpha$ on the left channel (we shall hereafter shorten this to "recorded with $\alpha$ on $L$") and with $\beta$ on $R$ (i.e., with gain $\beta$ on the right channel). Then the direction of motion of a stylus tracking a phonograph recording of this two-channel signal will make an angle $\theta = \arctan(\alpha/\beta)$ with the direction of the motion of the right groove wall (see Fig. 1**a**). If $\theta$ is increased by 180° (i.e., if the gains $\alpha$ and $\beta$ are changed to $-\alpha$ and $-\beta$), then the line of motion of the stylus will not change, so that the sound will still be recorded in the same stereo position, with only an inversion of the sound waveform. Thus the effective range of stereo positions in the groove causing straight-line motions of the stylus is only 180° (e.g., $-90° \leq \theta \leq 90°$).

Many of the early matrix systems [8]–[10] used only these straight-line motions to feed the four speakers, as does the regular matrix system [11], [12] for noninterior positions. These stereo positions may conveniently be represented by points around a circle, where a point on the circumference at an angle $2\theta$ from the righthand point of the circle represents sounds recorded with an angle of stylus motion $\theta$ from the direction of motion of the right groove wall (see Fig. 1**b**). By this means, the available 180° of direction of stylus motion available in a groove is represented by a full 360° around a circle.

Other stereo positions are available in the groove by applying phase shifts to the sounds recorded on each wall; for sine-wave sounds, the stylus is then caused to trace an elliptical path—the motions are more complex for nonsinusoidal waveforms. Consider a signal recorded with a gain $\alpha$ and a phase shift $\psi+\phi_L$ on the $L$ channel, and with gain $\beta$ and a phase shift $\psi+\phi_R$ on the $R$ channel, where $\phi_L$ and $\phi_R$ are independent of frequency, and where $\psi$ is a frequency-dependent phase shift that is, unfortunately, inevitable with realizable phase difference networks [13]. In studying systems, it is convenient to ignore $\psi$, which acts

* An earlier version of this paper was read to the A.E.S. British Section on February 8, 1972, under the title "Matrix Systems for Four-Speaker Stereo."

equally on all signals, and thus to consider the sound as being recorded with $\alpha$ (phase shifted $\phi_L$) on $L$ and with $\beta$ (phase shifted $\phi_R$) on $R$. If $\alpha$ and $\beta$ have the same sign, then the relative phase of the $L$ channel with respect to the $R$ is $\phi = \phi_L - \phi_R$; if they have opposite signs, then clearly the relative phase is $\phi = 180° + \phi_L - \phi_R$.

Such a stereo position may be represented by an extension of the circle picture of Fig. 1**b**. Let the horizontal great circle of a sphere represent stereo positions with 0° or 180° relative phase shifts as in Fig. 1**b**, and let the point $A$ represent a sound with $\alpha$ on $L$ and $\beta$ on $R$. When a relative phase shift $\phi$ of $L$ with respect to $R$ is introduced, then we rotate this horizontal circle about its left/right axis by an angle $\phi$ upward to obtain from $A$ a new point $B$ on the sphere (see Fig. 2). This new point represents the stereo position of a sound recorded with $\alpha$ (phase shifted $\psi + \phi$) on $L$ and with $\beta$ (phase shifted $\psi$) on $R$ for any phase shift $\psi$ that affects both channels equally.

Any recorded stereo position can thus be represented by a point on the surface of the sphere. Fig. 3 shows what various points on the sphere correspond in terms of recorded gain and phase shift, and in terms of the stylus motion in the groove when tracking sine tones. For reasons that will become apparent, this sphere representing the various modes of placing a monophonic signal into two channels will be called the "energy sphere." Note that changing the gain or the phase equally on both channels does not affect the energy sphere point of the recorded sound, since the stereo position remains unchanged. This means that the way a mono sound is encoded into two channels is fully specified only if its sphere point is supplemented by information about its energy gain.

This paper departs from Scheiber's account [1] in several respects. It shows how the relative phase relationships between the (say) four channels prior to matrixing down to two channels may be characterized by means of the pan locus. The pan locus is that curve on the surface of the energy sphere that the stereo position of an encoded sound traces as the sound is panned 360° around a listener. There are two pan loci of importance associated with each matrix system, the "optimal" or "ideal" locus and the "pairwise" pan locus which is that obtained when a four-channel tape

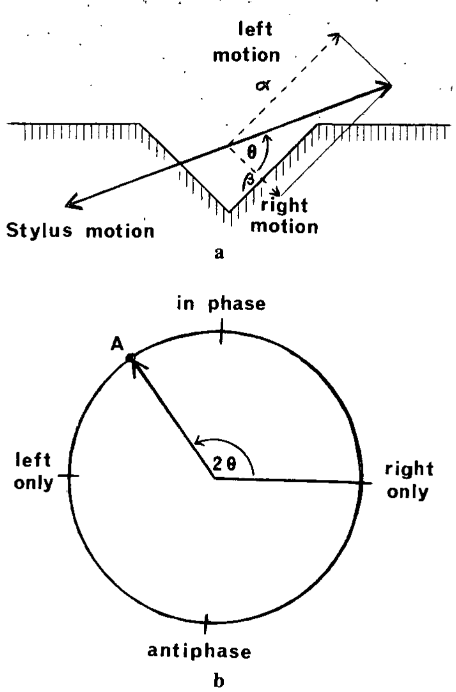



Fig. 1. Point $A$ at an angle $2\theta$ from right point of **b** represents a direction of stylus motion at an angle $\theta$ from the right wall's direction of motion as in **a**.

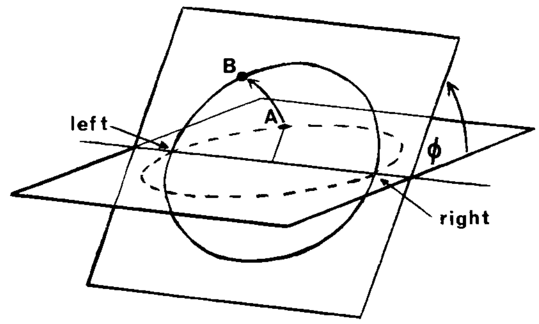



Fig. 2. Representation of a stereo position on a sphere. Point $A$ in horizontal plane represents a sound recorded with $\alpha$ on $L$ and with $\beta$ on $R$ as in Fig. 1. Point $B$ represents a sound recorded with $\alpha$ (phase shifted $\phi$) on $L$ and with $\beta$ on $R$.

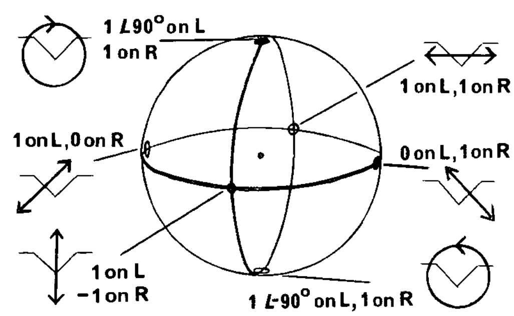



Fig. 3. Stereo positions and sine-wave stylus motions corresponding to various points on energy sphere.

mixed using pairwise pan-potting is encoded. While pairwise pan-potting is nonideal even for direct four-speaker reproduction [14]–[16], the pairwise pan locus also serves as a pictorial characterization of the relative phase properties of the (say) four encoded channels for those systems which encode from only four channels, even if pairwise mixing is not used on a four-channel input to the encoder. In the following, the pan loci discussed will always be the pairwise ones.

Other new departures here are the use of interior sphere points to characterize complex sounds, the use of "conformal" transformations of the sphere to represent $2 \times 2$ matrixing, and the use of the sphere to represent four-speaker decoders, with "pan loci" now used to represent phase relationships between decoder outputs.

Appendix I gives the formulas that allow one to compute the energy sphere point corresponding to recording with given gains and phase shifts on two channels, and vice versa.

Appendix II shows how the energy sphere can be used to predict the apparent positions of sounds with interchannel phase differences in two-speaker stereo reproduction according to several theories of sound localization.

## II. MATRIXING AND CONFORMAL TRANSFORMATIONS

In this section we show that the geometric properties of the energy sphere are related to matrixing in a natural manner. Suppose that two systems of encoding four channels or infinity channels into two are such that recordings made by one system can be converted into recordings indistinguishable from recordings made for the other simply by passing the two-channel signal of the first system through a linear $2 \times 2$ matrix involving only frequency independent coefficients and relative phase shifts. Two such systems are said to be "interconvertible," and any four-speaker playback effect obtained from one can equally be obtained (albeit via a modified decoding matrix) from the other. For example, for noninterior sound positions, the regular matrix [11], [12] and BMX [17] systems are interconvertible, although they are certainly not identical.

It transpires that the effect of a $2 \times 2$ matrix on a two-channel recording is to move the energy sphere points corresponding to all different stereo positions around. The precise distortion or transformation of the surface of the energy sphere thus caused by matrixing is of the type known as a "conformal" transformation of the sphere, i.e., a transformation of the points of the surface which preserves the angles between curves on the sphere, and that takes circles into other circles on the sphere.

The simplest conformal transformation of the surface of a sphere is obtained simply by rotating the sphere about its center. It can be shown [5], [18] that a $2 \times 2$ matrixing of two channels into two causes a rotation of the energy sphere points representing stereo positions if and only if, whenever a sound that is recorded with $\alpha$ (phase shifted $\phi_L$) on $L$ and with $\beta$ (phase shifted $\phi_R$) on $R$ is converted to a sound recorded with $\alpha'$ (phase shifted $\phi_{L'}$) on $L$ and with $\beta'$ (phase shifted $\phi_{R'}$) on $R$, $\alpha^2 + \beta^2$ is equal to some fixed multiple of $\alpha'^2 + \beta'^2$, i.e., if and only if the energies of all recorded sounds remain unaltered (apart possibly from an overall gain change) after the matrixing. For example, a phase shift applied to the left channel causes a rotation of the sphere about the left/right axis as in Fig. 2. Similarly, if the left and right signals $L$ and $R$ are matrixed to become $L \cos \theta + R \sin \theta$ in the left channel and $-L \sin \theta + R \cos \theta$ in the right (thereby rotating the direction of stylus motion by $\theta$ anticlockwise), then the energy sphere is rotated anticlockwise by an angle $2\theta$ about the vertical axis.

There is another important type of conformal transformation, known to physicists as a boost, which deforms the shapes of geometric figures (such as pan loci) on the sphere. Suppose that the two channels are subjected to a matrixing such that two opposite points $D$ and $D'$ (e.g., corresponding to in-phase and out-of-phase signals) on the energy sphere as in Fig. **4a** are not altered in position, but such that the relative gain of one (e.g., the in-phase) of the corresponding signals is increased with respect to that of the other. (Such a transformation occurs, for example, when a width, or "panoramic" control is used, or when a balance control is used.) Then the points of the sphere are moved towards the former point $D$ on the sphere, and away from the latter $D'$, as shown in Fig. **4b**. Such a transformation is called a boost. If $x$ (between $-1$ and 1) represents the height of a point above the center of the sphere as a proportion of the radius, measured in the direction of the axis joining the opposite points $D$ and $D'$ as illustrated in Fig. **4a**, then the effect of the boost transformation is to raise that point so that its new $x$ becomes equal to $(x + \lambda)/(1 + \lambda x)$ for some fixed number $\lambda$ between $-1$ and 1. Fig. **4b** shows the effect of the boost for which $\lambda = 1/3$.

Note that points which are opposite on the sphere before a boost (e.g., the pairs $A$, $A'$ or $B$, $B'$ in Fig. **4a**) are not in general opposite afterward. Thus, while opposite points represent "orthogonal" signals in the sense of Bauer et al. [19], [20], or "independent" signals in the sense of Scheiber [1], this relationship between recorded signals has no fundamental significance in four-speaker reproduction.

All other conformal transformations of the sphere may be represented as the effect of a rotation followed by a boost, or alternatively as the effect of a boost followed by a rotation. This provides us with a picture of the effect of all $2 \times 2$ matrixings on a two-channel signal.

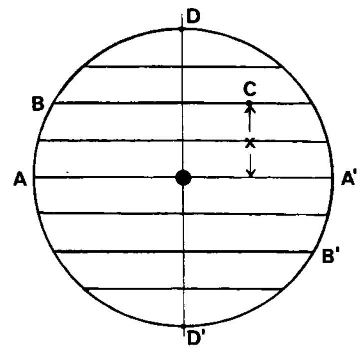
## III. 360° PAIRWISE PAN LOCI

The fact that conformal transformations preserve both circles and angles suggests that both of these geometric entities represent some important properties of matrix systems. Here we describe how circles and angles between circles characterize the effect of pan-potting sounds between pairs of channels that are subsequently matrixed into two channels. This will allow us to more completely specify the encoding matrix, explicitly when it is fed by pairwise pan-potted signals, and as a pictorial device otherwise. We can use this work to investigate directional ambiguities associated with pairwise pan-potting, a problem originally raised by Bauer [21].

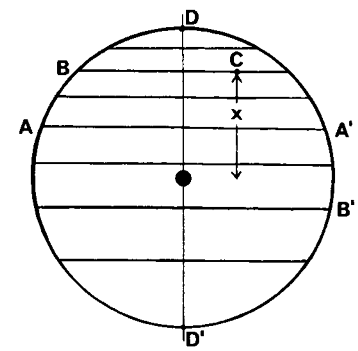
The encoding stage of a 4–2–4 system (i.e., the reduction from four channels to two) or an $n$–2–$m$ system is usually described by specifying how the four (or $n$) signals are recorded. On the energy sphere, and for independent channels, this corresponds to giving the four (or $n$) points corresponding to the stereo positions assigned to the input channels, along with associated energy gains. This, essentially, is the use made by Scheiber [1] and Eargle [22] of the sphere. However, in four-channel recordings, sounds are often assigned to positions between speakers by being pan-potted between the relevant channels. To each ratio of intensities on a pair of input channels corresponds a different energy sphere point on the encoded recording.

Suppose that a channel $U$ is encoded with $\alpha$ (phase shifted $\phi_L$) on $L$ and with $\beta$ (phase shifted $\phi_R$) on $R$, and that a channel $V$ is encoded with $\alpha'$ (phase shifted $\phi_L'$) on $L$ and with $\beta'$ (phase shifted $\phi_R'$) on $R$. Let u and v be the corresponding points on the energy sphere. It transpires (see Fig. 5) that sounds pan-potted between $U$ and $V$ are recorded in the encoded two-channel recording in stereo positions whose energy sphere points lie on a circular arc (not necessarily a great circle arc) joining $u$ and $v$ on the energy sphere.

The particular circular arc involved is determined entirely by the relative phases with which $U$ and $V$ are encoded. The following rule enables the relative phase of $U$ with respect to $V$ to be



a



b

Fig. 4. Effect of a "boost" transformation on energy sphere toward $D$ away from $D'$. Each horizontal line is a side view of a circle around the sphere. The "longitude" of points such as $C$ remains unchanged by the boost.

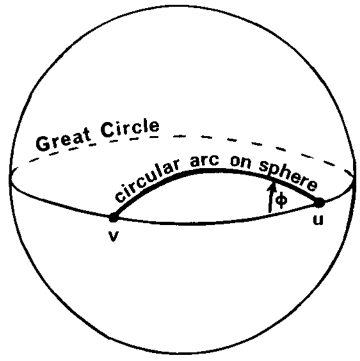



Fig. 5. Locus traced by energy sphere point of a sound pan-potted between channels which are encoded via points $u$ and $v$.

calculated. Although at first sight it might seem a little complicated, familiarity soon shows that it corresponds closely to intuitive ideas of what relative phase should mean. Compute the complex number

$$\alpha\alpha'[\cos(\phi_L-\phi_L') + j\sin(\phi_L-\phi_L')] + \beta\beta'[\cos(\phi_R-\phi_R') + j\sin(\phi_R-\phi_R')]$$

and write it in the form $k(\cos\phi + j\sin\phi)$, with $k$ being positive. Then $\phi$ is defined to be the relative phase of $U$ with respect to $V$. Thus, for example, the relative phase of two encoded signals is 0° (i.e., they are in-phase) if they are recorded without using any phase shifts and cause (when fed with identical signals separately) stylus motions within 90° of one another; otherwise if they are recorded without phase shifts, their relative phase is 180° (i.e., they are out of phase). In general, if $U$ is given an additional phase shift of $\Phi$, then the relative phase of $U$ with respect to $V$ is increased by $\Phi$.

The following rule chooses the circular arc between $u$ and $v$ on which sounds pan-potted between $U$ and $V$ lie (see Fig. 5): If $U$ and $V$ have 0° relative phase, then this arc is the shortest great circle path between $u$ and $v$; if the relative phase is 180°, then it is the longest great circle path between $u$ and $v$; if $U$ has a relative phase of $\phi$ with respect to $V$, then the arc is that circular arc on the surface of the sphere that starts off from $u$ toward $v$ at an angle $\phi$ clockwise from the shortest great circle path from $u$ to $v$.

If four channels I, II, III, and IV are encoded, with the channel pairs I, II and II, III and III, IV and IV, I forming adjacent pairs around the listener, then a pairwise pan-potted sound circling round the four speakers will be recorded in two channels in stereo positions corresponding to points on the sphere lying on four circular arcs forming a quadrilateral as in Fig. 6, known as the 360° pan locus of the system. Similarly, an $n$–2 system would have a 360° pan locus which is an $n$-sided circular-arc polygon on the sphere.

In order that sounds recovered from the two-channel encoded recording of a pairwise pan-pot master be reproduced from the correct position during playback, it is necessary that the 360° pan locus should not cross over itself (as otherwise two sounds assigned to different positions during recording would be encoded in precisely the same manner) and that it snould have no very sharp or cusp-like corners (as otherwise sounds recorded on the two sides of the corner would not be adequately distinguishable by any decoder design). Such effects causing a poor or incorrect location during playback are called directional ambiguities [21].

The main result about these is that for a 4-into-2 encoding, the 360° pan locus has its four corner angles add up to 360°. (For $n$-into-2 recordings, the angles of the 360° pan locus add up to $180n$-360°, as for a plane linear polygon). While we omit details,

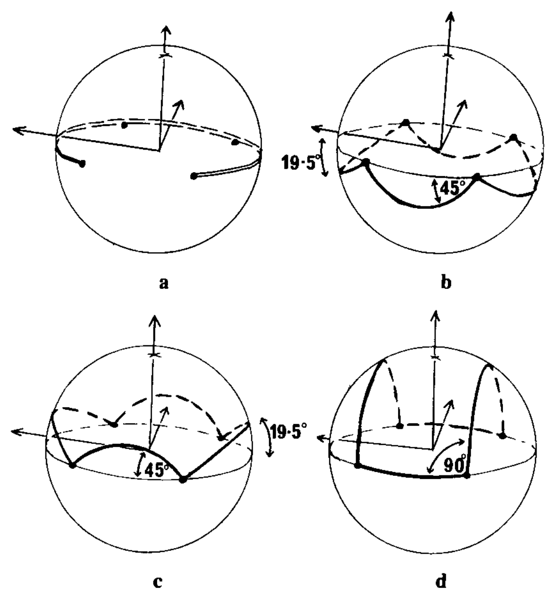


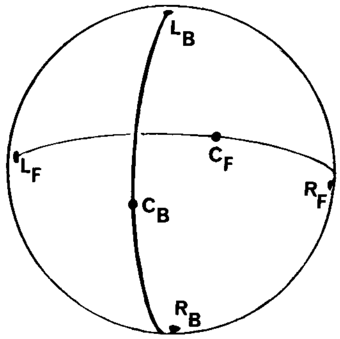

Fig. 6. 360° pairwise pan locus for 4–2 systems whose corner points are those of regular matrix system, in the cases when the relative phases between the $L_B$, $L_F$, $R_F$, $R_B$, and $L_B$ input channels are: **a.** 0°, 0°, 0°, 180°; **b.** −45°, −45°, −45°, −45°; **c.** 45°, 45°, 45°, 45°; **d.** 90°, 0°, 0°, 90°.

this may be proved by boosting the sphere strongly to a point inside the polygon (remembering that boosts do not affect angles), and altering all the relative phases (which cannot affect the sum of angles) until all but one side of the polygon approximates the straight side of a plane polygon. It is then not difficult to prove that the remaining side must also correspond to a relative phase of nearly 0°, so that the polygon approximates to an ordinary plane polygon for which the theorem is well known.

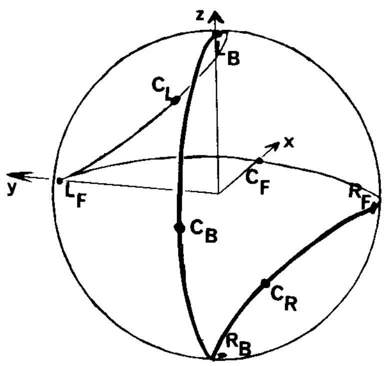
Ideally, in most cases one would like the arcs of the pan locus to be great circle arcs (as they are in the "optimal" encoding specification of the SQ, regular matrix, and BMX systems). However, the above pan locus theorem shows that the sum of angles of the pairwise pan locus is always less than that of the great-circle arc polygon with the same vertices (whose sum of angles in radians exceeds that of a plane polygon by the enclosed area in steradians, by a well-known result of spherical trigonometry). Thus a certain amount of "kinkiness" and loss of directional accuracy is inevitable with any system encoded from a pairwise mixed four-channel source. Presumably this is why all the "optimal" encoding specifications do not follow a pairwise 360° pan locus. Combined with the poor image quality of pairwise mixed four-channel tapes [14]–[16], this is a very powerful reason for not regarding pairwise mixed sources as a standard of reference for surround sound work.

A widely used 4–2–4 proposal, whose front and back energy sphere points are as shown in Fig. 7**a**, cannot have a 360° pan locus not crossing over itself with angles adding up to 360° unless it is as shown in Fig. 7**b** (or its mirror image). Thus two out of the four angles of the illustrated system must equal 0°, and these cusp-like angles will cause a high degree of directional ambiguity near the $L_F$ (left front) and $R_B$ (right back) corners on pairwise mixed material. Because of small imperfections in practical decoders, encoders, and recording media, even the most sophisticated variable matrix decoders cannot resolve such ambiguities near these corners.

By recording the four channels in stereo positions corresponding to the vertices of a regular tetrahedron, a 4-into-2 encoding system can be obtained [1] that may be decoded with every channel cross talking at − 4.8 dB on every other. Unfortunately, such "tetrahedral" systems have 360° pairwise pan loci (for




a




b

Fig. 7. Pairwise pan locus of widely used matrix system. Angles at $L_F$ and $R_B$ are 0°; those at $R_F$ and $L_B$ are 180°.

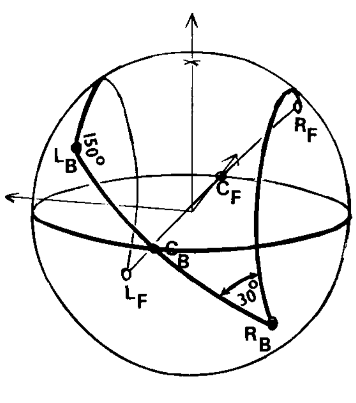



Fig. 8. Typical "tetrahedral" system with 0° phase differences at front and back, and 90° phase differences at sides.

example, as in Fig. 8) at least one of whose angles must be less than or equal to only 30°, as may be deduced from the pan locus theorem. This means that such a system would encode pairwise pan-potted material with rather large directional ambiguities near some of its corners. This contrasts with the 90° corner angles obtainable from pairwise mixed material for systems such as regular matrix, BMX, or QS (see Fig. **6b–d**).

The arc joining two points at an angle $2\theta$ apart on the sphere and having relative phase difference $\phi$, deviates from the "ideal" great circle path by an angle $\psi$ on the sphere (see Fig. 9) which spherical trigonometry shows is given by the formula

$$\cos\psi = (\cos\theta + \cos\phi)/(1 + \cos\theta\cos\phi).$$

For example, if $\theta = \phi = 45°$ (as in Fig. **6b** or **c**), then $\psi = 19.5°$. We note that this formula is the same in form as that describing boosts, with $\cos\theta$ and $\cos\phi$ replacing $x$ and $\lambda$.

All $n$-sided circular arc polygons on the sphere with angles adding up to $180n - 360°$ arise from $n$-into-2 encoding systems. Thus systems may be invented and, as we shall see, evaluated, simply by drawing such polygons on the sphere (or plane pictures of them), and using the formulas of Appendix I and the relative phase rules given above to get mathematical descriptions corresponding to the system pictured. This highlights the fact that the 360° pan locus is a useful pictorial device for visualizing interchannel phase treatment in a system, even if there is no serious intention of using the system with such an inadequate system of panning as pairwise pan-potting. In particular, any system whose pairwise pan locus shows a high degree of asymmetry is likely to exhibit phase asymmetry anomolies in decoding even if a symmetrical pan locus is chosen as "optimal."

## IV. ENERGY AND THE ENERGY SPHERE

Suppose a signal is recorded with $\alpha$ (phase shifted $\phi_L$) on $L$ and with $\beta$ (phase shifted $\phi_R$) on $R$. Then it is recorded with an energy gain $\alpha^2 + \beta^2$. If several independent signals are recorded at points $u_1, u_2, \cdots, u_n$ on the sphere with respective energy gains $g_1, g_2, \cdots, g_n$, and if the signals have respective energies $e_1, e_2, \cdots, e_n$, then the total energy of the resultant recorded signal is clearly $e_1 g_1 + e_2 g_2 + \cdots + e_n g_n$, and the distribution of energy among the two channels can be represented by that point inside the energy sphere placed at the center of mass of the points $u_i$, each weighted by its respective "mass" $e_i\, g_i$.

Although the general theory cannot be given here, this "energy point" of a recording turns out to have very great importance in the general theory of $n$-channel recording systems. In particular, the energy point of a signal represents essentially all the available information (within any given range of frequencies) about the distribution of signal energy among the (in this case two) channels of the recording. Thus it represents essentially all the information that can be used by "logic-type" nonlinear gain control or variable-matrix methods of playback, which increase the apparent separation between channels by varying the gains of the reproducing channels [19], [23]. Thus the three parameters describing the vector joining the center of the sphere to the energy point of a two-channel sound, called the "envelope vector," are all that are available within each frequency band for gain control.

For example, the system of Fig. 7, with the $(x, y, z)$ axes shown in Fig. **7b** for the energy sphere, requires that the gain control should turn down the back channels if the envelope vector $(x, y, z)$ is such that $x \geqslant 0$, $z = 0$, and $x^2 + y^2 = 1$, and that the front channels should be turned down if $x \leqslant 0$, $y = 0$, and $x^2 + y^2 = 1$. If a large amount of sound energy is recorded at the side (i.e., if $z \neq 0$ and $x + y = \pm 1$), then the front and back channels should have equal gain. However, it will be noted from Fig. **7b**, that side-left ($C_L$) sounds are recorded within 45° on the sphere from the front-center ($C_F$) sounds, and that side-right ($C_R$)

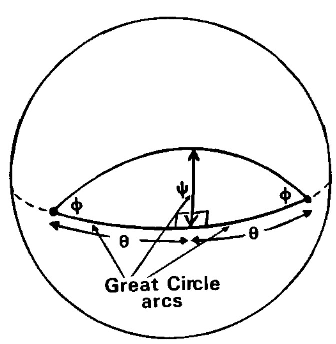



Fig. 9. Deviation $\psi$ of a circular arc from a great circle arc.

sounds are only 45° from back-center ($C_B$) sounds. Because of this, there will be a tendency for gain-control circuitry to turn down the back channels, even when there is quite a strong signal recorded on the left side when the predominant signal is at the front. Thus there will be a tendency for left-side signals to wander to the front with gain-control playback of the system of Fig. **7b**, and similarly for right-side signals to wander to the back. This illustrates the use of the energy sphere to determine gain-control behavior of $n$–2–$n$ systems without having to consider details of the circuitry used.

## V. PLAYBACK

So far, the process of playback or decoding has not been described in terms of the energy sphere. If a sound is derived (say for a speaker) from two channels with gain $\alpha$ from the left one (phase shifted by $-\phi_L$) and with gain $\beta$ from the right one (phase shifted by $-\phi_R$) by adding these two channel signals together, then this method of decoding or playing back a signal may be regarded as corresponding to the same point on the energy sphere as does a recording with $\alpha$ (phase shifted $\phi_L$) on $L$ and with $\beta$ (phase shifted $\phi_R$) on $R$. The reason for the minus signs on the phase shifts corresponding to an energy sphere point during playback is so as to undo the effects of the phase shifts during recording. The above method of playback is said to have an energy gain equal to $\alpha^2 + \beta^2$.

Thus any method of playback can partially be represented by four points on the energy sphere. If one desires to keep track of their relative phases, one can join the points with the relevant 360° pairwise pan locus, although the intermediate points of the locus have no direct meaning during playback. Many versions of $n$–2–$n$ systems (including versions of SQ [19], [20], earlier Sansui QS proposals [23], and a BMX proposal [17]) have their recording and playback energy sphere points precisely the same, and one reason for this property lies in the following result. If the center of mass of the $n$ points (weighted by their energy gains) of the encoding process of an $n$–2–$n$ system lies at the center of the energy sphere, then playback via $n$ speakers will assign sounds to their respective speakers with minimum total cross-talk energy onto the others if and only if the playback energy sphere points, and their energy gains, are precisely the same as those used for encoding.

If a recording does not have its center of mass at the center of the sphere, then a boost of suitable magnitude along the axis joining the center of mass to the center of the sphere will always convert it to such a form. The playback method that gives the same results after this boost as a given one before has its energy sphere points derived from the previous playback points by applying an equal boost in the opposite direction to that applied to the encoding points. While the property of encoded signals being "orthogonal" normally has no relevance to four-speaker playback, it thus becomes meaningful when systems are such that encode and playback points can be made to coincide on the sphere.

If $g_1$ is the energy gain of a channel $U$ recorded at the point $u$ on the sphere, and if it is played back via speaker $V$ with energy gain $g_2$ at a point on the sphere placed at an angle $\theta$ away from $u$ (see Fig. 10), then the energy gain with which channel $U$ is conveyed to speaker $V$ can be shown to be $\frac{1}{2}g_1g_2(1 + \cos\theta)$. The proof, in outline, is as follows. There is an energy-preserving $2 \times 2$ matrixing of the two channels which causes a rotation of the energy sphere such that $u$ and $v$ lie in the horizontal great circle. Thus (see Fig. 1) the energy gain of $U$ in $V$ is $g_1g_2$ times the square of the cosine of the angle between the corresponding directions of stylus motion. Thus the gain is $g_1g_2\cos^2\frac{1}{2}\theta$, which equals $\frac{1}{2}g_1g_2(1 + \cos\theta)$ as required.

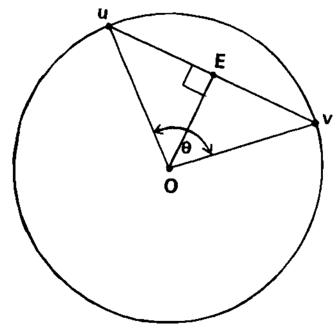



Fig. 10. Angle between points on a sphere. $0E$ has a length whose square is $\frac{1}{2}(1 + \cos\theta)$.

This cardioid-shaped energy[1] pick-up characteristic of playback is particularly useful in examining the monophonic compatibility of systems. For monophonic playback the point $v$ is at the front of the sphere, so that sounds recorded at points at an angle $\theta$ from the forward-pointing axis are attenuated by $10\log_{10}\frac{1}{2}(1 + \cos\theta)$ dB, e.g., the attenuation is $-3$ dB for sounds which are recorded on $L$ only, on $R$ only, or equally on $L$ and $R$ but with 90° relative phase between the two channels. Thus, for example, inspection of the pan locus of the system of Fig. **7b** shows that side-left sounds are reproduced in mono much louder than side right sounds. It is not difficult to establish that their respective angles from the $x$ axis are 45° and 135°, so that their gain differential is $10\log_{10}[(1 + \cos 45°)/(1 + \cos 135°)] = 7.7$ dB.

While the energy sphere has proved useful for studying mono playback, it has not, in the author's experience, proved to be of great use in designing linear four-speaker decoders. Two-speaker stereo playback and its subjective effects are discussed in Appendix II.

## CONCLUSIONS

While the energy sphere is not always the best model for studying two-channel systems, it does seem to be the most generally useful. The scope of the model and its utility are enhanced when one has the various theorems about pairwise pan loci, a knowledge of the effects of conformal transformations such as boosts as well as rotations, and interior sphere points to represent full two-channel signals (which really goes back to Stokes [4]). It has been the object of this paper to present these results as geometrically as possible, with the minimum of mathematical formulas and we have only outlined the proofs of results in cases where the idea is primarily one of geometry. However, some formulas are useful for using the sphere to derive detailed answers, and Appendix I gives some of the more basic ones.

Inevitably, we have had to leave out many applications, and just give individual examples of others (e.g., of the study of variable-matrix decoders, directional ambiguity and mono compatibility). In Appendix I, we derive the two-channel periphonic (i.e., with-height) system discovered independently in [1] and [2] [25], and in Appendix II we study stereo localization using the sphere, by way of an example of a non-four-speaker use. Since the sphere has also had application to other phenomena such as polarized light [4], [5] and radio waves [6] [18], it is to be anticipated that other two-channel information applications will be found in the future not involving surround sound.

It may also be asked whether any of the above work may be

[1] Ordinary cardioid microphones have cardioid amplitude pick-up characteristics.

extended to more than two information channels. The answer is certainly yes, but it must be remarked that invariably the extended model becomes much less intuitive. One generalization is to spin spherical harmonic systems on a sphere [25], [7], another distinct one represents $n$-channel stereo "positions" as being points in $(n-1)$-dimensional (real or complex) elliptic space, rather than the sphere, and a third looks at the convex set of positive trace-one (i.e., of total energy one) "coherency matrices" [26], [27] of an $n$-channel system. All of these generalizations do prove useful in tackling particular multichannel problems, but do not have the breadth of application that the energy sphere has in two-channel problems.

Thus, in practice, the energy sphere remains a uniquely simple way of cutting through a jungle of matrix formulas, and a working knowledge of its properties is as useful to those who design two-channel processing equipment as more traditional mathematical tools.

## APPENDIX I

## STEREO POSITIONS ON THE SPHERE

Let a point on the energy sphere have rectangular coordinates $(x, y, z)$, where $x^2 + y^2 + z^2 = 1$, where $x$ points forward, $y$ to the left, and $z$ upward as, for example, in Fig. 7**b**. Let the corresponding signal be recorded with $\alpha$ (phase shifted $\phi_L$) on $L$ and with $\beta$ (phase shifted $\phi_R$) on $R$. We may consider the left channel as having a complex gain $a = \alpha \exp(j\phi_L)$ (where $j = \sqrt{(-1)}$ represents a 90° phase shift) and the right as having a channel gain of $b = \beta \exp(j\phi_R)$. Then applying geometrical arguments to Figs. 1 and 2, it is not hard to show that

$$\begin{aligned} y &= (\alpha^2 - \beta^2)/(\alpha^2 + \beta^2) \\ &= (|a|^2 - |b|^2)/(|a|^2 + |b|^2) \end{aligned}$$

and

$$\begin{aligned} x + jz &= 2\alpha\beta e^{j(\phi_L - \phi_R)}/(\alpha^2 + \beta^2) \\ &= 2ab^*/(|a|^2 + |b|^2) \end{aligned}$$

where the asterisk denotes complex conjugation.

Conversely, given $x$, $y$, and $z$, the values of $a$ and $b$ can be determined only to within multiplication by some number, as $(x,y,z)$ gives no indication of the gain or absolute phase of the recording. Thus $a$ and $b$ are given, to within multiplication by a common factor, by any of the following pairs of formulas (which may be obtained from one another by multiplying both $a$ and $b$ by a suitably chosen number):

$$a = (1 + y)^{\frac{1}{2}}, \qquad b = (x - jz)(1 + y)^{-\frac{1}{2}}$$

or else

$$a = (x + jz)(1 - y)^{-\frac{1}{2}}, \qquad b = (1 - y)^{\frac{1}{2}}$$

or else, for any chosen value of the number $\theta$,

$$a = (1 - je^{j\theta}x + y + e^{j\theta}z)(1 + x \sin\theta + z \cos\theta)^{-\frac{1}{2}}$$

and

$$b = (-je^{j\theta} + x + je^{j\theta}y - jz)(1 + x \sin\theta + z \cos\theta)^{-\frac{1}{2}}.$$

We note that these results and other related ones may be found in several texts on the rotation and Lorentz groups, such as [29] and [24].

A sound recording system that assigns sounds coming from the direction of $(x,y,z)$ from the origin to the two channels with the gain $a$ on $L$ and with $b$ on $R$ as above will have encoded a full three-dimensional directionality ("periphony" [25]) into two channels (see also [1]). Of course, these two channels may be matrixed into the actual two channels of recording or transmission media in any way that gives adequate mono and stereo compatibility. Such a recording may be played back via, say, a regular tetrahedron, octahedron, or cube of loudspeakers [25] by feeding a speaker in the direction of $(x,y,z)$ with the two channels mixed together with respective gains $a^*$ and $b^*$. Due to phase anomalies [25], this system can never reproduce low-frequency directionality correctly. [16].

Recordings for this system may be made using an arrangement of several coincident conventional pressure and pressure gradient microphones [25], [2] having an amplitude sensitivity in the direction $(x,y,z)$ of $1 - je^{j\theta}x. + y + e^{j\theta}z$ on $L$ and $-je^{j\theta} + x + je^{j\theta}y - jz$ on $R$, for arbitrary $\theta$. It is interesting to note that this periphonic system is, most remarkably, implicit in the work of Blumlein [28] in 1931.

## APPENDIX II

## TWO-SPEAKER STEREO LOCALIZATION

The energy sphere is particularly useful for discussing two-speaker reproduction of signals having interchannel phase differences. If the signals on the $L$ and $R$ channels have respective complex gains $a$ and $b$, then the sound pressure information at the listener has gain proportional to $a + b$, the forward-pointing velocity has gain also proportional to $a + b$, and the leftward-pointing component of velocity has gain proportional to $a - b$. Thus, at low frequencies the only sound localization information available to the ears is the complex number $(a - b)/(a + b)$, both for theories supposing the head to be fixed and for those involving information obtained from head movements [16].

It is normally assumed at low frequencies that interaural phase information is important, but not interaural amplitude ratios (which are close to 1). It may be shown that this assumption is equivalent to saying that all localization information is contained in the quantity

$$P = \operatorname{Re}\left(\frac{a - b}{a + b}\right) \tag{1}$$

and that the information

$$Q = \operatorname{Im}\left(\frac{a - b}{a + b}\right) \tag{2}$$

is not used for localization (where Re and Im are defined by $\operatorname{Re}(x + jy) = x$ and $\operatorname{Im}(x + jy) = y$).

Such theories are apt only at low frequencies (below about 700 Hz [16]), and one therefore has reason to suspect that in practice the quantity $Q$ will start contributing to the apparent sound position at higher frequencies. Since variation of sound position with frequency will cause image blurring, the quantity $Q$ (which we shall term "phasiness") is a measure of image blurring caused by interspeaker phase differences. At high frequencies other sound localization theories have to be used (involving the ear using higher order harmonic components of the directional sound field), so we must regard $P$ and $Q$ as descriptions of "position" and "phasiness" information only to a first approximation. Also, because there are several theories of low-frequency localization [16], a given value of $P$ corresponds to different stereo positions in different theories.

We can, however, get useful information by asking what, in this first approximation, are the points on the energy sphere corresponding to a constant $P$ (whatever actual localization this may correspond to). Similarly, we can ask how we can measure the phasiness $Q$ of a sound from its energy sphere representative point.

From Appendix I we have

$$P = \mathrm{Re}\left[\frac{(1+y)^{\frac{1}{2}} - (x - jz)(1+y)^{-\frac{1}{2}}}{(1+y)^{\frac{1}{2}} + (x - jz)(1+y)^{-\frac{1}{2}}}\right]$$

$$= \mathrm{Re}\left[\frac{1 + y - x + jz}{1 + y + x - jz}\right]$$

$$= \frac{(1 + y - x)(1 + y + x) - z^2}{(1 + y + x)^2 + z^2}$$

$$= \frac{1 + y^2 - x^2 + 2y - z^2}{2 + 2x + 2y + 2xy} = \frac{2y + 2y^2}{2 + 2x + 2y + 2xy}$$

making use of $x^2 + y^2 + z^2 = 1$. Thus,

$$P = y/(1 + x). \tag{3}$$

The locus of constant position is thus the intersection of a plane $y = P(1 + x)$ and the sphere $x^2 + y^2 + z^2 = 1$. This plane passes through the antiphase sphere point $(x,y,z) = (-1,0,0)$ and is perpendicular to the "horizontal" plane $z = 0$. Thus the circle of intersection is a vertical circle on the surface of the sphere passing through the antiphase point. Thus we can announce the following theorem.

## LOCALIZATION THEOREM

In all low-frequency interaural phase theories of sound localization, points on the energy sphere describing a sound fed to two speakers all correspond to the same image location if and only if they lie on a vertical circle on the sphere surface passing through the antiphase point.

In particular, when viewed from directly above the sphere, and looking down onto the "horizontal" circle, the loci of constant two-speaker position look like straight lines, as in Fig. 11. Since $y = P$ when $x = 0$ by eq. (3), the intersection of these "lines" (as they seem when viewed from above) with the diameter through the left and right points is exactly the predicted position on the line joining the speakers of a sound image predicted by the theories of Leakey [30], Makita [31], and Tager [32]. $P$ given by eq. (1) is this interspeaker location in these theories by the "stereophonic law of tangents" for images given by them. In practice, of course, the images become hazy and unstable toward the antiphase point, as the different loci converge close toward one another.

There is an equally attractive picture for phasiness $Q$. From eq. (2) and Appendix I we have

$$Q = \mathrm{Im}\left[\frac{1 + y - x + jz}{1 + y + x - jz}\right] = \frac{z(1 + y - x) + z(1 + y + x)}{(1 + y + x)^2 + z^2}$$

$$= \frac{2z(1 + y)}{2(1 + x)(1 + y)} = z/(1 + x).$$

Thus by arguments similar to before, we announce the following theorem.

## PHASINESS THEOREM

Suppose that, to a first approximation the phasiness $Q$ describes image degradation caused by phase differences between a pair of stereo speakers. Then $Q$ is the height above the energy sphere center of the intersection of the vertical axis and that plane through the energy sphere point of the sound that passes also through the antiphase point of the sphere and which meets the horizontal plane there along a tangent to the sphere.

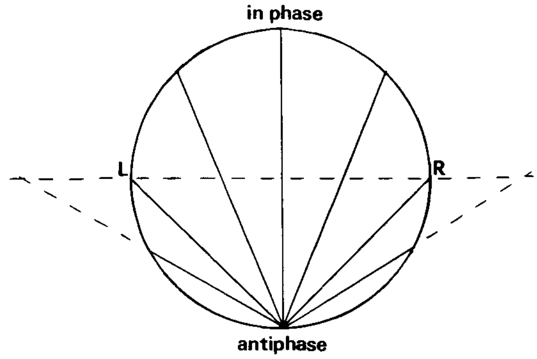



Fig. 11. Loci of constant two-speaker stereo position on energy sphere, viewed from above.

Thus taking now a side view of the sphere, looking along the left/right axis, the phasiness corresponding to a point $C$ is given as in Fig. 12 if the sphere is taken to be of unit radius. There is some experimental disagreement as to what values of $Q$ are "acceptable." Based on results and impressions from central images (which correspond to points on the circumference of Fig. 12), the permissible range of $Q$ varies between

$$-0.1 < Q < 0.1$$

for some experienced and critical recording engineers,

$$-0.25 < Q < 0.25$$

for those who tolerate up to 30° phase difference (a figure often quoted informally),

$$-0.4 < Q < 0.4$$

according to some recent research by the BBC Research Department, and

$$-1 < Q < 1$$

seems to be the maximum range that is often considered to be tolerable.

We comment that, as with position, values for phasiness $Q$ obtained for sphere points anywhere near the antiphase point are likely to be misleading, since the images there are highly unstable and the "constant phasiness" loci become packed closely together there.

## ACKNOWLEDGMENT

We would like to thank Dr. Duane Cooper for many helpful comments on an earlier draft of this paper.

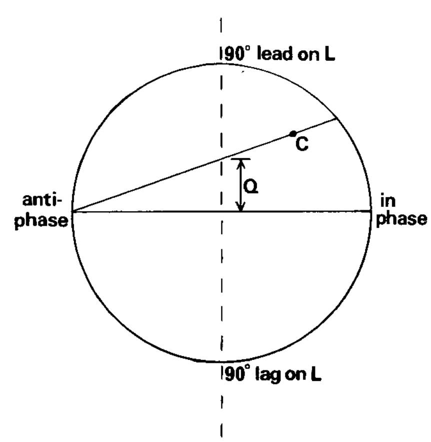



Fig. 12. "Phasiness" $Q$ of encoded signal $C$ on energy sphere, viewed from side looking along left/right axis.

## REFERENCES

[1] P. Scheiber, "Analyzing Phase-Amplitude Matrices," *J. Audio Eng. Soc.*, vol. 19, pp.835-839 (Nov. 1971).

[2] M. A. Gerzon, "Matrix Systems for Four-Speaker Stereo I," (Sept. 1971).

[3] P. B. Fellgett, "Perspectives for Surround Sound," in *Hi-Fi Sound Annual 1974,* (Haymarket Publ., London, 1973), pp. 56-67.

[4] G. G. Stokes, "On the Composition and Resolution of Streams of Polarised Light from Different Sources," *Trans. Camb. Phil. Soc.*, vol. 9, p. 399 (1852); reprinted in G. G. Stokes, *Mathematical and Physical Papers,* vol. III (Cambridge University Press, Cambridge, 1901), pp. 233-258.

[5] H. Poincaré, *Théorie Mathématique de la Lumière II* (Georges Carré, Paris, 1892), pp. 282-285, 302.

[6] G. A. Deschamps, "Geometrical Representation of the Polarization of a Plane Electromagnetic Wave," *Proc. IRE,* vol. 39, pp. 540-544 (1951).

[7] A. Held, E. T. Newman, and R. Posadas, "The Lorentz Group and the Sphere," *J. Math. Phys.*, vol. 11, pp. 3145-3154 (Nov. 1970).

[8] M. A. Gerzon, "The J.O.K.E. System," *Hi-Fi News,* vol. 15, pp. 843, 847 (June 1970).

[9] L. Feldman, "On Matrix Quadraphonic Systems," *Audio,* vol. 55, pp. 20-27 (Oct. 1971).

[10] C. Stark, "One Approach to Four-Channel Sound—Dynaco/Gately," *Audio,* vol. 55, pp. 30-32 (Oct. 1971).

[11] "Regular Matrixed Disk Records," Engineering Committee Standard, Japan Phonograph Record Assoc., Tokyo (Mar. 23, 1972).

[12] P. B. Fellgett, "The Japanese Regular Matrix," *Hi-Fi News,* vol. 17 pp. 2393-2397 (Dec. 1972).

[13] W. R. Albersheim and F. R. Shirley, "Computation Methods for Broad-Band 90° Phase-Difference Networks," *IEEE Trans. Circuit Theory,* vol. CT-16, pp. 189-196 (May 1969).

[14] P. B. Fellgett, "Directional Information in Reproduced Sound," *Wireless World,* vol. 78, pp. 413-417 (Sept. 1972).

[15] O. Kohsaka, E. Satoh, and T. Nakayama, "Sound-Image Localization in Multichannel Matrix Reproduction," *J. Audio Eng. Soc.*, vol. 20, pp. 542-548 (Sept. 1972).

[16] M. A. Gerzon, "Surround Sound Psychoacoustics," *Wireless World,* vol. 80, pp. 483-486 (Dec. 1974); also published in French; "Critères Psychoacoustiques Relatifs à la Conception des Systèmes Matriciels et Discrets en Tétraphonie," in *Conf. des Journées d' Études, Festival International du Son 1974* (Éditions Radio, Paris, 1974), pp. 145-155.

[17] D. H. Cooper and T. Shiga, "Discrete-Matrix Multichannel Stereo," *J. Audio Eng. Soc.*, vol. 20, pp. 346-360 (June 1972).

[18] T. S. Chu, "Restoring the Orthogonality of Two Polarizations in Radio Communications Systems II," *Bell Syst. Tech. J.*, vol. 52, pp. 319-327 (Mar. 1973).

[19] B. B. Bauer, D. W. Gravereaux, and A. J. Gust, "A Compatible Stereo-Quadraphonic (SQ) Record System," *J. Audio Eng. Soc.*, vol. 19, pp. 638-646 (Sept. 1971).

[20] B. B. Bauer, "The Compatible Stereo-Quadraphonic 'SQ' Record," *Audio,* vol. 55, pp. 34-40 (Oct. 1971).

[21] B. B. Bauer, "Directional Ambiguity of Quadruphonic Matrices," *J. Audio Eng. Soc.* (Letter), vol. 19, pp. 315-316 (Apr. 1971).

[22] J. M. Eargle, "4–2–4 Matrix Systems: Standards, Practice, and Interchangeability," *J. Audio Eng. Soc.*, vol. 20, pp. 809-815 (Dec. 1972).

[23] R. Itoh and S. Takahashi, "The Sansui QS Coding System and a New Technique to Improve its Inter-Channel Separation Characteristic," presented at the 42nd Convention of the Audio Engineering Society, Los Angeles, May 2-5, 1972, preprint F-6.

[24] T. Kahan, *Theory of Groups in Classical and Quantum Physics, Vol. I* (Oliver and Boyd, Edinburgh, Scotland, 1965), appendix to ch. 2, pp. 49-60; ch. 3, pp. 61-97.

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

[26] N. Wiener, "Coherency Matrices and Quantum Theory," *J. Math. Phys. M. I. T.*, vol. 7, pp. 109-125 (1928).

[27] N. Wiener, "Generalized Harmonic Analysis," *Acta Math.*, vol. 55, pp. 117-258 (1930), sec. 9, pp. 182-195; also in *Selected Papers of Norbert Wiener* (M. I. T. Press, Cambridge, Mass., 1965); N. Wiener, *Generalized Harmonic Analysis and Tauberian Theorems* (M. I. T. Press, Cambridge, Mass., 1966).

[28] A. D. Blumlein, U. K. Patent 394, 325 (1931), page 5, lines 70-106, and page 15, lines 17-89.

[29] I. M. Gel'fand, R. A. Minlos, and Z. Ya. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon Press, Oxford, 1963), pt. I, ch. 1, sec. 1, pp. 8-13; pt. II, ch. 1, sec. 1, pp. 166-172.

[30] D. M. Leakey, "Some Measurements on the Effects of Interchannel Intensity and Time Difference in Two Channel Sound Systems," *J. Acoust. Soc. Am.*, vol. 31, pp. 977-987 (1959).

[31] Y. Makita, "On the Directional Localization of Sound in the Stereophonic Sound Field," *E. B. U. Rev.*, pt. A, no. 73, pp. 102-108 (June 1962).

[32] P. G. Tager, "Some Features of Physical Structure of Acoustic Fields of Stereophonic Systems," *J. Soc. Motion Picture Eng.*, vol. 76, pp. 105-110 (Feb. 1967).

## THE AUTHOR

Michael A. Gerzon received an M.A. degree in Mathematics at Oxford University in 1967. Since completing postgraduate research in axiomatic quantum theory at Oxford University, he has been involved in research into the properties of linear and nonlinear multichannel systems, including analytic system theory.

One aspect of this work has been a study of surround sound recording and reproduction systems. This includes the abstract mathematical theory of these systems, the study of mathematical models for non-directional and directional psychoacoustics, and the design of microphone arrays, studio processing equipment and decoders based on these studies. Since the start of 1974, much of this work has been carried out in connection with the British N.R.D.C. (National Research Development Corporation) and its 'ambisonic' systems of surround sound.

Mr. Gerzon has published over 15 papers and articles on the theory and practice of surround sound systems, as well as papers on a number of other audio topics, and is a member of the Audio Engineering Society.