Design of Ambisonic Decoders for Multispeaker Surround Sound

Michael A. Gerzon
Mathematical Institute
University of Oxford
Oxford OX1 3LB , England

Abstract

We describe the design procedures used for psychoacoustically optimised decoders for multispeaker surround-sound reproduction, starting from the basic principles of human sound localisation expressed in a mathematical form. Included are data for phase shift networks, resistor matrix design, phase-compensated shelf filters and loudspeaker layout compensation, illustrated by circuits from a commercially available Ambisonic decoder.

Presented at the 58th Audio Engineering Society Convention, New York, 4th November 1977.

I. INTRODUCTION

The design of decoding apparatus for surround sound is determined both by the form of the encoded surround sound signals, and the particular design aims of the designer. The overall performance can only be assessed if all stages of the reproduction process are included, including the listener, since the decoder is not doing its job unless the illusion in the ears and mind of the listener has substantially the desired directional and ambient qualities. All high fidelity design involves some degree of subjective optimisation among mutually conflicting desirable qualities, and this can only be achieved by prolonged skillful listening. However, a tractable design procedure and usable mathematical description of the system is an aid to getting sufficiently close to the desired performance to enable the final subjective optimisation to be a process of refinement, rather than of radical alterations and guesswork.

Hitherto, the design of surround-sound equipment has involved large amounts of guesswork, since the requirements for a satisfying directional illusion were either unknown or known only in terms of complicated experimental or theoretical data not easily convertable into a workable design. In the absence of a precise knowledge of requirements, many designers have assumed that the "optimum" reproduction would simulate paiwise mixed reproduction, with non-corner sound directions being achieved by reproduction just from the adjacent pair of loudspeakers. Experimental data, especially but not only for nominally side localisations [1]-[4], showed that this approach actually gives a poor and uneven directional illusion which lends itself mainly to crude 4-point "gimmickry".

Over the last few years, a range of increasingly sophisticated signal-actuated matrix decoders (often termed "variomatrix" or "logic") have been designed to approximate pairwise mixed reproduction, which give poor phantom images, as well as an inevitable increase in listening fatigue under extended listening owing to the continual alteration of the signal parameters reaching the ears. All signal-actuated designs so far announced tend to cause the position of secondary sound sources to wander in the presence of louder primary sources, although some designs are much better than others in this respect. While further improvements in signal-actuated designs can be expected, it is extremely unlikely that forseeable technology will yield designs with low listening fatigue and which preserve the subtleties of ambient acoustics that can make surround-sound reproduction so worthwhile.

We here describe a systematic approach to decoder design that takes into account both the properties of the human listener and of the domestic contraints on the use of the equipment. While the design methods are mathematical (but no more so than, say, active filter design), it must be understood that the mathematical modelling of human hearing used below is an abstraction from the data of over seventy years of scientific research into the auditory localisation of sound. A series of papers in preparation will give the full theory that lies behind the models and results quoted, and give a much more extended discussion of its meaning and interpretation. Other mathematical models not quoted here also play an important part in a thorough design, since human hearing is a complex many-method sound analysing device. However, the models given here already take into account more aspects of human hearing than previous attempts. The main point to note is that the mathematics given here is cast into a form that allows relatively easy design of decoders, once one has the results of the various quoted theorems.

We then go through the various circuits used in practical consumer decoders designed using these procedures, including various means of achieving maximum circuit simplicity, and formulas or numerical data for realisation of the circuits. We describe the overall form of the decoder, and circuits for phase shifters, phase-compensated shelf filters, speaker layout adjustment controls and matrix circuits.

The decoders here described will give substantially correct sound localisation via a variety of different shapes of loudspeaker layout, and consequently, includes a control to adjust for different rectangular shapes, as well as for hexagonal layouts that are capable of improved subjective performance when circumstances permit their use. Also described is the use of delay-line compensation for irregular speaker layouts such as trapezia. We also discuss briefly the application of these design techniques to signal-actuated designs.

This work was done in connection with the Ambisonic surround-sound project of the British NRDC (National Research Development Corporation), and is the subject of a number of pending Patents. Further background that the reader may find helpful is contained in the series of articles [5]-[9]. In particular, the full circuit of a commercially available decoder designed by the methods given here is given in reference [9], which includes details of the decoding parameters that can be used for all the major 2-channel encoding systems, including mono, stereo, RM, BMX, SQ, matrix H, and system 45J.

II. MATHEMATICAL THEORY OF DIRECTIONAL HEARING

In this paper, we shall only consider sounds originating in the horizontal plane, although the following methods can be extended to the design of periphonic (with-height) [10] systems. Consider x,y axes pointing respectively forward and to the left (see figure 1), and consider, for the present, n loudspeakers situated on a circle around a central point in the azimuthal directions ϕ_i $(i=1,2,\ldots,n)$ measured anti-clockwise from the x-axis (due front). For simplicity, we also suppose that the speakers lie at a large distance from the central point, so that their sounds arrive at the listener as plane waves. If a given mono sound is fed to all loudspeakers, with the complex gain P_i to the i'th loudspeaker, then the following parameters influence the apparent psychoacoustic localisation of the sound in question.

1. Makita's localisation θ_V.

Describes the apparent localisation azimuth for listeners facing the apparent sound source at low frequencies less than about 1 kHz.

Calculate:

$$x = \mathrm{Re}\left\{\frac{\Sigma P_i \cos\phi_i}{\Sigma P_i}\right\},$$

$$y = \mathrm{Re}\left\{\frac{\Sigma P_i \sin\phi_i}{\Sigma P_i}\right\},$$

where Re means "real part of", and the sums are over $i = 1$ to n. Then Makita's localisation azimuth θ_V is given by

$$\left.\begin{aligned} x &= r_V\cos\theta_V \\ y &= r_V\sin\theta_V \end{aligned}\right\} \text{ with } r_V > 0.$$

2. Velocity vector magnitude r_V .

This is the quantity $r_V = (x^2+y^2)^{\frac{1}{2}}$ defined above. It describes the stability of sound localisation with head rotation at low frequencies less than about 1 kHz. For natural sounds, $r_V = 1$, and if r_V is much greater than 1 for a reproduced sound, one hears an "out-of-phase" effect. If r_V is close to zero, one hears an "in-the-head" or "close-to-the-head" effect, along with excessive image movement when the head moves.

3. Phasiness q .

Apt for forward-facing listeners. It describes unpleasant "pressure on the ears" sensation and image blurring. It is given by

$$q = \mathrm{Im}\left\{\frac{\Sigma P_i \sin\phi_i}{\Sigma P_i}\right\},$$

where for real u,v, Im(u+iv) means v. Ideally, for natural sounds, one should have $q = 0$, but $|q| < 0.21$ is relatively innocuous, $|q| < 0.5$ is generally tolerable, and $|q| > 1$ is unacceptable.

4. Energy vector azimuth θ_E .

This describes the apparent image azimuth at high frequencies above about 500 Hz up to about 5 kHz, and is also apt at low frequencies for slightly off-centre listeners. Calculate:

$$x_E = \frac{\Sigma |P_i|^2 \cos\phi_i}{\Sigma |P_i|^2}$$

$$y_E = \frac{\Sigma |P_i|^2 \sin\phi_i}{\Sigma |P_i|^2}$$

and put

$$\left.\begin{aligned} x_E &= r_E \cos \theta_E \\ y_E &= r_E \sin \theta_E \end{aligned}\right\} \text{ with } \; r_E > 0 \; .$$

Then θ_E describes the apparent localisation, especially when the listener faces its apparent direction.

5. Energy vector magnitude r_E .

The quantity $r_E = (x_E^2 + y_E^2)^{\frac{1}{2}}$ defined above describes the stability of the sound image with movement of the head, especially at frequencies between about 500 and 5000 Hz. Ideally for natural sounds, $r_E = 1$, and for reproduced sounds r_E can never exceed one. In practice, $r_E = 0.7$ is excellent, $r_E = 0.5$ is quite acceptable, and $r_E = 0.35$ is tolerable for sounds to the rear of the listener.

III. THEOREMS USEFUL IN DESIGN

In the following, we use the same capital-letter symbols to indicate a signal and the complex gain with which a given mono sound is included within that signal. The following results can be proved about the various localisation parameters described above, and greatly simplify the task of decoder design.

Theorem 1. Rectangle decoder theorem.

The speaker feed signals LB, LF, RF, RB of a rectangle speaker layout (see figure 2) give reproduction whose Makita and energy vector azimuth localisations coincide provided that

$$Q = \tfrac{1}{2}(-LB+LF-RF+RB) = 0 \; ,$$

and in that case the speaker feed signals can be expressed in terms of three signals W, X, Y respectively representing the reproduced acoustic pressure, component of acoustic velocity along the x-axis, and acoustic velocity along the y-axis, as follows:

$$\begin{aligned} LB &= \tfrac{1}{2}(W - 2^{-\frac{1}{2}}X/\cos\phi \; + 2^{-\frac{1}{2}}Y/\sin\phi), \\ LF &= \tfrac{1}{2}(W + 2^{-\frac{1}{2}}X/\cos\phi \; + 2^{-\frac{1}{2}}Y/\sin\phi), \\ RF &= \tfrac{1}{2}(W + 2^{-\frac{1}{2}}X/\cos\phi \; - 2^{-\frac{1}{2}}Y/\sin\phi), \\ RB &= \tfrac{1}{2}(W - 2^{-\frac{1}{2}}X/\cos\phi \; - 2^{-\frac{1}{2}}Y/\sin\phi), \end{aligned}$$

where $180^\circ-\phi$, ϕ, $-\phi$ and $-180^\circ+\phi$ are the respective azimuth angles of the loudspeakers LB, LF, RF, RB in figure 2. For such a decoder, Makita's azimuth θ_V and the energy vector azimuth θ_E are both given by the quantity θ defined by either of the formulas:

$$\cos\theta : \sin\theta = \mathrm{Re}(XW^*) : \mathrm{Re}(YW^*)$$

and

$$\cos\theta : \sin\theta = \mathrm{Re}(X/W) : \mathrm{Re}(Y/W),$$

where * indicates complex conjugation.

Theorem 2. Regular polygon decoder theorem.

If $n \geq 4$ loudspeakers are positioned on a circle at azimuths ϕ_i ($i = 1$ to n) spaced apart by equal angles $360^\circ/n$, and for all i, the feed signal P_i to the loudspeaker at azimuth ϕ_i is given by:

$$P_i = W + (2^{\frac{1}{2}}\cos\phi_i)X + (2^{\frac{1}{2}}\sin\phi_i)Y ,$$

then the Makita and energy vector azimuths coincide, and their value $\theta = \theta_V = \theta_E$ is given by the same formulas for θ as in theorem 1.

Theorem 3. Decoder modification theorem.

In the decoders of theorems 1 and 2, if the signals W, X, Y are respectively replaced by signals of the form

$$W' = k_1 W$$

$$X' = k_2 X + jk_4 W ,$$

$$Y' = k_2 Y - jk_3 W ,$$

with $k_1 > 0$, $k_2 > 0$, and k_3 , k_4 real, then the Makita and energy vector azimuths remain unchanged, and the velocity vector magnitude r_V is changed to

$$r'_V = (k_2/k_1)r_V .$$

Alteration of k_3 changes the phasiness q of reproduction in a manner depending on the azimuth of the encoded sound.

Theorem 4. In the decoders of theorems 1 and 2, the velocity vector magnitude r_V and phasiness q are given by:

$$r_V^2 = \tfrac{1}{2}(\mathrm{Re}(X/W))^2 + \tfrac{1}{2}(\mathrm{Re}(Y/W))^2$$

and

$$q = 2^{-\frac{1}{2}} \mathrm{Im}(Y/W).$$

Comment. By these results, in order to design decoders that have correct Makita azimuth θ_V and velocity vector magnitude r_V at low frequencies, and correct energy vector azimuth θ_E at high frequencies, for all possible encoded sound directions, it is necessary that it be possible to matrix the original encoded signals into three signals W,X,Y whose gains satisfy for every encoded azimuth θ the equations

$$\begin{aligned} \mathrm{Re}(X/W) &= 2^{\frac{1}{2}} r_0 \cos\theta , \\ \mathrm{Re}(Y/W) &= 2^{\frac{1}{2}} r_0 \sin\theta , \end{aligned} \qquad (1)$$

where r_0 is a constant greater than zero independent of azimuth θ, which equals r_V if W,X,Y are used in the decoders of theorems 1 and 2 without modification, by theorem 4. We shall say that an encoding system is well-decodable if it is possible to derive such signals **W**, **X**, Y such that θ of equation 1 agrees with the encoded azimuth to within a few (2 or 3) degrees and if r_0 thus defined is constant (to within a few percent) with encoded azimuth. In order to design psychoacoustic decoders, it is important first of all to determine whether an encoding system is well-decodable, and if so, how to find the matrix that derives the important signals W, **X**, Y.

Theorem 5. Pairwise suboptimality theorem.

No two-channel encoding originating from pairwise mixed material is well-decodable for all encoded directions, although it may be well-decodable for the four corner azimuths.

Theorem 6. SQ non-well-decodability theorem.

The **SQ** two-channel encoding system [11] is not well-decodable even for just the four corner azimuths.

Theorem 7. Three-channel decoding theorem.

Three-channel encoding systems with encoding channel gains of the form

$$W = 1 , \ X = 2^{\frac{1}{2}}\cos\theta , \ Y = 2^{\frac{1}{2}}\sin\theta ,$$

or any independent complex linear combinations thereof (e.g. UMX [12], System 45J [8]) are well-decodable, and if the signals W, X, Y are used without modification in the decoders of theorems 1 and 2, the reproduction will have zero phasiness and $r_V = 1$.

Theorem 8. Two-channel decoding theorem.

Two-channel surround-sound encoding systems whose transmission channels L and R have gains of the form (see also ref. [8]) :

$$L+R = a + c\cos\theta + ej\sin\theta ,$$
$$L-R = bj + dj\cos\theta + f\sin\theta ,$$

for all encoding azimuths θ, and with a,b,c,d,e,f real constants, are well-decodable provided that the quantities

$$u = \frac{cf+ed}{bc-ad} , \quad v = -\left(\frac{be+af}{bc-ad}\right)$$

are such that $|u| < 0.5$ and $0.5 < |v| < 2$ approximately. For such systems, the signals W, X and Y for the decoders of theorems 1 and 2 may be obtained as follows (see also ref. [8]):

Put

$$h = v^{-1}\left\{\frac{4+3u^2}{4-(u/v)^2}\right\}^{\frac{1}{2}} ,$$

$$g = \frac{u(1+3v^2)}{4+3u^2} \times \frac{h^2}{1+vh} ,$$

and compute the matrix

$$\begin{pmatrix} \alpha & j\beta & j\gamma \\ \delta & j\epsilon & j\xi \\ j\chi & \psi & \omega \end{pmatrix} = \begin{pmatrix} a & c & je \\ jb & jd & f \\ jg & jh & -1 \end{pmatrix}^{-1} .$$

Then

$$W = \alpha(L+R) + j\beta(L-R) ,$$
$$2^{-\frac{1}{2}}X = \delta(L+R) + j\epsilon(L-R) ,$$
$$2^{-\frac{1}{2}}Y = j\chi(L+R) + \psi(L-R) ,$$

and the value of r_0 for these signals is

$$r_0 = v/(v+h^{-1}) ,$$

which typically has a value near $\frac{1}{2}$.

Comment. As might be expected, the proof of theorem 8 involves very large amounts of messy matrix algebra. Although this theorem is for two-channel systems, it is in fact intimately related to the problem of adding a third channel of limited bandwidth to a two channel system, outlined in ref. [8]. Why this should be so is not clear, since the result is found to emerge only after lengthy analysis.

We are still left with the problem of choosing the gains k_1 and k_2 of theorem 3 at high frequencies so as to optimise energy vector magnitude r_E .

Theorem 9. Energy vector optimisation theorem.

Suppose that a regular polygon decoder of the form of theorem 2 has the signals W, X, Y modified as in theorem 3, with $k_3 = k_4 = 0$. Then the energy vector magnitude r_E is maximised in a given encoded sound direction if, for a sound in that direction we put:

$$k_2/k_1 = \frac{|W|}{(|X|^2 + |Y|^2)^{\frac{1}{2}}} ,$$

and for this sound direction, one then has $r'_E = r'_V$ for the modified decoder, i.e. the velocity vector magnitude and the energy vector magnitudes become equal. It is always true for the decoders of theorem 2 that $r_E \leq 2^{-\frac{1}{2}}$.

Comment. Having determined the optimum k_2/k_1 for energy vector localisation at high frequencies, one can use theorem 4 to compute the resulting values of r_V and r_E . For the 2-channel systems considered in theorem 8, it is found empirically that k_2/k_1 approximately equals 1 for the choice of W,X,Y given in that theorem for optimal energy vector localisation, whereas k_2/k_1 should equal r_0^{-1} for optimal low frequency localisation. Note that theorem 9 shows that it is never possible for regular polygon decoders of the type of theorem 2 to be simultaneously optimal for r_E and r_V, since ideally $r_V = 1$. For this reason, it is essential for best results to design a decoder so that the gains k_1 and k_2 are frequency-varying. Only in this way is it possible to satisfy both the low-frequency criteria 1 and 2 of section II and the high-frequency criteria 4 and 5. The fact that $\theta_V = \theta_E$ at all frequencies for the decoders of theorems 1 and 2 makes the transition between these two frequency regions subjectively smooth and without anomalies. The gain k_4 will be zero for left/right symmetric decoders, and k_3 is chosen so as to minimise the overall phasiness of reproduction. The overall gain at each frequency should be chosen so as to maintain the flattest possible frequency response for all directions, although clearly some encoding systems will allow more uniform frequency responses with encoded direction than others, with 2-channel "circle locus" systems (i.e. as in theorem 8 with $v^2 = 1+u^2$) allowing a flatter frequency response in all directions than "bent locus" systems such as the BBC matrix H system.

Suitable values of k_1, k_2, k_3 for various systems have been computed using the above theorems, and a list of some values is given in appendix A of ref.[8], and also as the shelf filter gains in ref. [9].

IV. DECODER ARCHITECTURE

An overall decoder architecture based on the results just described is shown in figure 3. The phase-amplitude matrix at the input derives the signals **W**, **X**, Y from the encoded input signals, along with a fourth signal -j**W** to enable control of phasiness as described by theorems 3 and 4. The shelf filters 1 to 3 have the gains k_1 to k_3 respectively of theorem 3, and in order that these frequency-dependent gains suffer from no unwanted relative phase shifts, these filters are "phase compensated", i.e. are designed so that at all frequencies they all cause substantially the same phase shift. The distance-compensation RC high-pass filters modify the velocity signals to compensate for the effects of finite loudspeaker distance. The output amplitude matrix derives the linear combinations of W',X',Y' required to feed the particular loudspeaker layout in use according to theorems 1 and 2, and in the case of a rectangular layout include a variable potentiometer adjustment to compensate for the angle ϕ of figure 2.

VI. SHELF FILTERS

In order to satisfy the requirement $r_V = 1$ at low frequencies and that r_E be maximal at high frequencies, the gains k_1, k_2 and k_3 must be implemented using shelf filters having substantially identical phase responses so that k_2/k_1 and k_3/k_1 are real. This is most easily done by using shelf filters of RC "all-pass" type such that the frequency at which the output leads the input by 90° in phase is the same "centre frequency" F for all filters. As yet there is no exhaustive body of experimental data on the optimal centre frequency for best subjective results (and indeed this may vary according to the system being decoded), but F = 400 Hz is a typical choice.

If each filter has low frequency gain k_L and high frequency gain k_H, then the complex frequency response of such a filter is

$$\frac{-k_L + jk_H\,\omega\tau'}{1 + j\,\omega\tau'} \quad , \qquad (2)$$

where the time constant τ' is related to the centre frequency F via the formula

$$\tau' = (k_L/k_H)^{\frac{1}{2}}/(2\pi F) \ . \qquad (3)$$

In practice, for reasons given below, it is neccesary to design the filters with an additional overall gain by fixing the value of k_L (say around 3), but retaining the desired ratio of k_H/k_L. The attenuation required to overcome this additional gain may be designed into the preceeding resistor matrix.

Figure 5 shows a practical op-amp implementation of such shelf filters according to equation (2). This has the useful feature of having a restive input impedance with resistance R. This means that the matrix resistors of the preceeding matrix circuit can be designed to feed this resistive load without any need for buffer amplifiers, provided that the shelf filter has gain to spare. The circuit shown also includes provision (the switch S) for obtaining more than one shelf filter characteristic for multisystem decoders without unduly complex circuitry or switching. Interestingly, this does not appreciably alter the centre frequency F of the filter. The low frequency gain k_L is not changed by the switch S, meaning that any change of k_L has to be designed into the appropriate resistor matrix.

Suppose that the desired low frequency gain is k_L and that the high frequency gains required are k_H (with the switch "out", i.e. R_2 in circuit) and k_H' (with the switch "in", i.e. R_2 out of circuit). It is necessary to choose $k_H \leqslant k_H'$ and $k_H' > 1$. Choose the desired input impedance R. Then the component values in figure 5 are given by the formulas:

$$R_3 = k_L R \ ,$$

$$R_1 = R_3/(k_H' - 1) \ ,$$

$$R_2 = R(k_H' - k_H)/(k_L + k_H) \ ,$$

$$C = (R + R_2)^{-1} \tau' \ ,$$

where the time constant τ' in the last formula is that computed for the filter with high frequency gain k_H in equ. (3).

Note that the circuit of figure 5 with $R_1 = \infty$, $R_2 = 0$ and $R_3 = R$ implements (with a phase inversion) a single pole unity gain RC all-pass phase shifter, with time constant $(2\pi F)^{-1} = RC$ and a resistive input impedance R. It can therefore be used to implement one or more of the poles of the phase shifters described earlier.

VII. RESISTOR MATRIX DESIGN .

The $\psi+0^\circ$ and $\psi+90^\circ$ phase shifters in the phase-amplitude matrix, along with inverter stages to obtain $\psi+180^\circ$ or $\psi-90^\circ$ shifts, can directly feed a resistor matrix to obtain the required signals W, X, Y, $-j$W. For virtual earth mixing, such resistor matrix design is too well known to require comment, as long as one remembers that the matrix coefficients should be chosen to give an overall matrix gain equal to $k_L^{(req)}/k_L$ times the theoretical coefficients computed via theorems 7 and 8, where k_L is the low frequency gain of the shelf filter in the relevant signal path, and $k_L^{(req)}$ is the low frequency gain actually required according to the design procedures of section III.

If the resistor matrix feeds the resistive input impedance R of the shelf filters directly, then the design procedure is not quite so straightforward. The effective load resistance R will cause attenuation of the matrixed signal, which is one reason for choosing an excessive gain $k_L > k_L^{(req)}$ in the shelf filters.

If the resistor matrix mixes n different source signals $S_1, \ldots, S_n$ from the phase shifters and inverters with intended overall low frequency gains (including the effect of the shelf filter) $g_1, \ldots, g_n$, yielding a signal $g_1 S_1 + \ldots + g_n S_n$, then the n resistors R_1' to R_n' feeding the input of the shelf filter from the sources (see figure 6) can be shown to have values given by the formula:

$$R_i' = R(1 - \sum_{i=1}^{n} G_i)/G_i$$

where

$$G_i = g_i \, k_L^{(req)}/k_L \; .$$

It is, of course, necessary to choose the low frequency gain k_L of the shelf filter so that $\Sigma G_i < 1$.

We comment that in the design of the phase-amplitude matrix, it proves to be simpler to design the matrix as a preliminary sum-and-difference matrix, followed by phase shifters acting on the sum and difference signals, followed by a resistor matrix. This is generally preferable to designs without an initial sum-and-difference matrix since it reduces the gain requirements of the shelf filters, reduces the number of matrix resistors and their required precision, and does not increase the number of active stages since it turns out that some of the phase inverters following the phase shifters are no longer needed.

VIII. LOUDSPEAKER LAYOUT CONTROL

According to theorem 1, it is neccesary to modify the gains of the X and Y signal paths by $2^{-\frac{1}{2}}/\cos\phi$ and $2^{-\frac{1}{2}}/\sin\phi$ respectively for the non-square rectangle layout of figure 2. In practice, for $25^{\circ} < \phi < 65^{\circ}$ approximately (beyond which localisation becomes very poor since the minimum value of r_E becomes excessively small), these gains may be approximated by $2^{\frac{1}{2}}\sin\phi$ and $2^{\frac{1}{2}}\cos\phi$ respectively. A suitable control circuit, using just a single potentiometer, for such "loudspeaker layout control" is shown in figure 7, where L is the load impedance of the output amplitude matrix circuit, R is the resistance of the potentiometer (including any padding resistors at the two ends to limit the range of the control to between 25° and 65°), and T is a fixed impedance chosen to achieve a suitable control law. To obtain a gain proportional to $\sin\phi$ in the X signal path and $\cos\phi$ in the Y signal path, one chooses the values of the load resistance L and the total potentiometer resistance R such that $L > 2^{-\frac{1}{2}}R$, and then the resistors T have the value:

$$T = LR/(\sqrt{2}\,L - R) ,$$

and the overall circuit gain (in the central square $\phi = 45^{\circ}$ setting) is

$$\frac{LR}{(2L+R)T + LR} .$$

It is necessary to design the output amplitude matrix to recover this lost gain, although it is also possible to recover it by putting gain into earlier stages of the circuit. In order to avoid crosstalk between X and Y, the wiper resistance should be less than 1% of R.

IX. REMAINING CIRCUITS

The high pass filters in the X and Y signal paths (see figure 3) should have time constant

$$RC = \tau = 2.94d \text{ msec.},$$

where d is the distance in metres of the loudspeakers from a central listener. This compensation for the effects of sound field curvature is only audible in its effect for special types of program (e.g. a double bass from behind), and is thus not very critical. A suitable choice for most uses is d = 3 metres, and it is not necessary to use better than 10 % tolerance capacitors in this circuit.

We shall not comment here on the design of the output amplitude matrix, since this is straightforward given the results of theorems 1 and 2 on deriving speaker feed signals from W, X, Y. Also, suitable designs have been published in references [7] and [9] for four and six loudspeakers. However, we note that in reference [9], we describe the use of speaker matrix techniques to reduce the cost by reducing the number of power amplifiers without altering the desired reproduced directional effect.

X. FURTHER REFINEMENTS

For trapezium-shaped speaker layouts (see figure 8), the decoders described above are no longer optimal. For a listener at or near position O in figure 8, it is possible to use a rectangle decoder with layout control set to the angle ϕ illustrated, provided that the alterations in the time of arrival of sounds and their intensity at the listener is compensated for the fact that some loudspeakers are closer than others. This is done by putting a delay in the signals fed to the closer speakers, as well as a gain (of less than 1) as illustrated for the layout of figure 8 in figure 9. In this particular case, the delay τ of the front speaker feeds is set equal to $2.94(d_B - d_F)$ msec. approximately, where d_B and d_F are the respective distances in metres from the point O of the back and front speakers. The amplitude gain of the front speakers is set equal to d_F/d_B so as to compensate for the inverse square law of sound intensity. Extensions of this delay compensation technique to more complex irregular layouts not equidistant from the listener are obvious.

The psychoacoustic design techniques of this paper can also be applied to optimise the subjective performance of variable or signal-actuated parameter (SAP) decoders. This can be done by making the gains k_3 and k_4 of theorem 3 dependent on the instantaneous localisation of dominant sounds, so as to reduce phasiness and maximise r_E in that sound's direction. Such "variable preference" decoders have the same basic block diagram as figure 3, except that the gain k_3 in the $-jW$ signal path is made variable, and that $-jW$ is subjected to a second variable gain k_4 and added to the X signal path also. By making the gains k_3 and k_4 apply to high-pass filtered signals (via double RC high-pass filters phase matched to the shelf filters k_1 and k_2), it is possible to arrange for the variable preference to operate over one or more frequency bands independently, without employing cross-over networks except in the $-jW$ signal path. SAP decoders employing these techniques are unique in giving correct Makita and energy vector azimuths for <u>all</u> sounds, not just those in the instantaneous dominant direction. As a result, the degree of image "pumping" is inherently smaller than for non-psychoacoustic SAP designs, although the listening fatigue problem remains.

REFERENCES

[1]. P.A. Ratliff, "Properties of Hearing Related to Quadraphonic Reproduction", BBC Research Department Report, BBC RD 1974/38 (Nov. 1974)

[2]. G. Thiele and G. Plenge, "Localization of Lateral Phantom Sources", J. Audio Eng. Soc., vol. 25, pp. 196-200 (April 1977)

[3]. P. Damaske and Y. Ando, "Interaural Crosscorrelation for Multichannel Loudspeaker Reproduction", Acustica, vol. 27, pp. 232-238 (1972)

[4]. K. de Boer, "A Remarkable Phenomenon with Stereophonic Sound Reproduction", Philips Tech. Rev., vol. 9, pp. 9-38 (1947)

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

[6]. M.A. Gerzon, "Criteria for Evaluating Surround-Sound Systems", J. Audio Eng. Soc., vol. 25, pp. 400-408 (June 1977)

[7]. M.A. Gerzon, "Ambisonics. Part two: Studio Techniques", Studio Sound, vol. 17 (8), pp. 24, 26, 28-30 (Aug. 1975) Correction: ibid, vol. 17 (10), p. 60 (Oct. 1975)

[8]. M.A. Gerzon, "The Optimal Choice of Surround-Sound Encoding Specification", Preprint no. 1199 (A-5), 56th Audio Engineering Society Convention, Paris, 1 March 1977

[9]. M.A. Gerzon, "Multi-System Ambisonic Decoder. 1- Basic Design Philosophy. 2- Main Decoder Circuits", Wireless World, vol. 83, pp. 43-47 (July 1977), pp. 69-73 (Aug. 1977)

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

[11]. 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)

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

Design 1. 6 poles.

Frequency range ($90^\circ \pm 2.4^\circ$): 149.9 Hz to 5003 Hz.
($90^\circ \pm 5^\circ$) : 48.8 Hz to 15358 Hz.

Pole frequencies.

$\psi+0^\circ$ lag network: 33.47, 499.7, 4770 Hz.

$\psi+90^\circ$ lead network: 157.2, 1501, 22409 Hz.

Design 2. 8 poles.

Frequency range ($90^\circ \pm 1\frac{1}{3}^\circ$): 30 Hz to 16 kHz.

Pole frequencies.

$\psi+0^\circ$ lag network: 14.90, 163.1, 1120, 8069 Hz.

$\psi+90^\circ$ lead network: 59.49, 428.5, 2943, 32210 Hz.

Design 3. 7 poles

Frequency range ($90^\circ \pm 2.5^\circ$): 34 to 17700 Hz.

Pole frequencies.

$\psi+0^\circ$ lag network: 19.43, 258.8, 2319, 30874 Hz.

$\psi+90^\circ$ lead network: 83.88, 774.6, 7153 Hz.

Table 1. Pole frequencies of various designs of relative 90° phase shifts, with frequency ranges over which various tolerances are maintained assuming ideal component values.

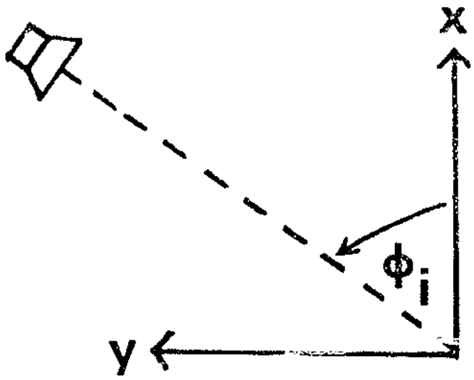

Figure 1. x,y coordinates and azimuth angle used in this paper.

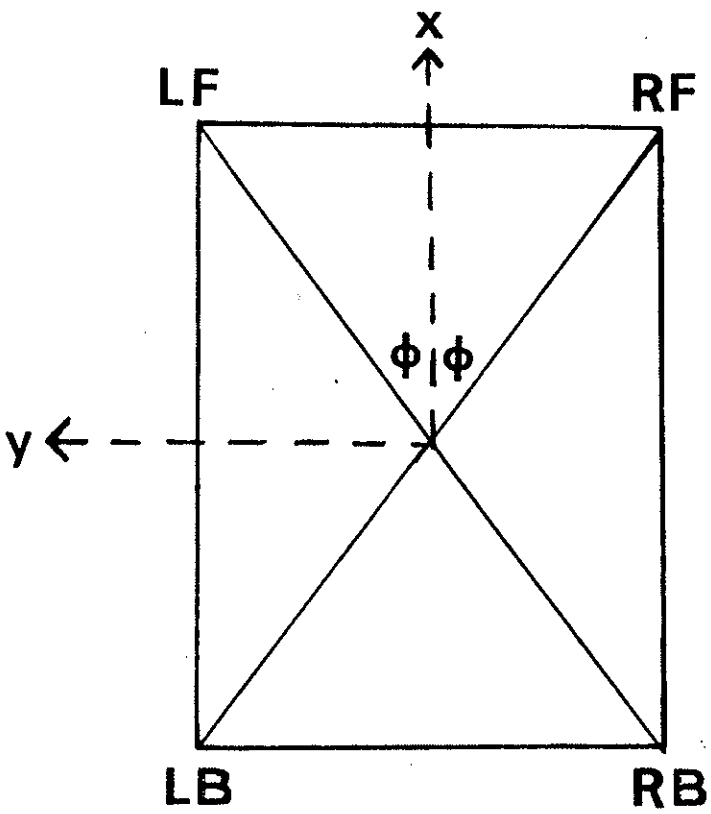

Figure 2. Rectangle decoder loudspeaker layout of theorem 1.

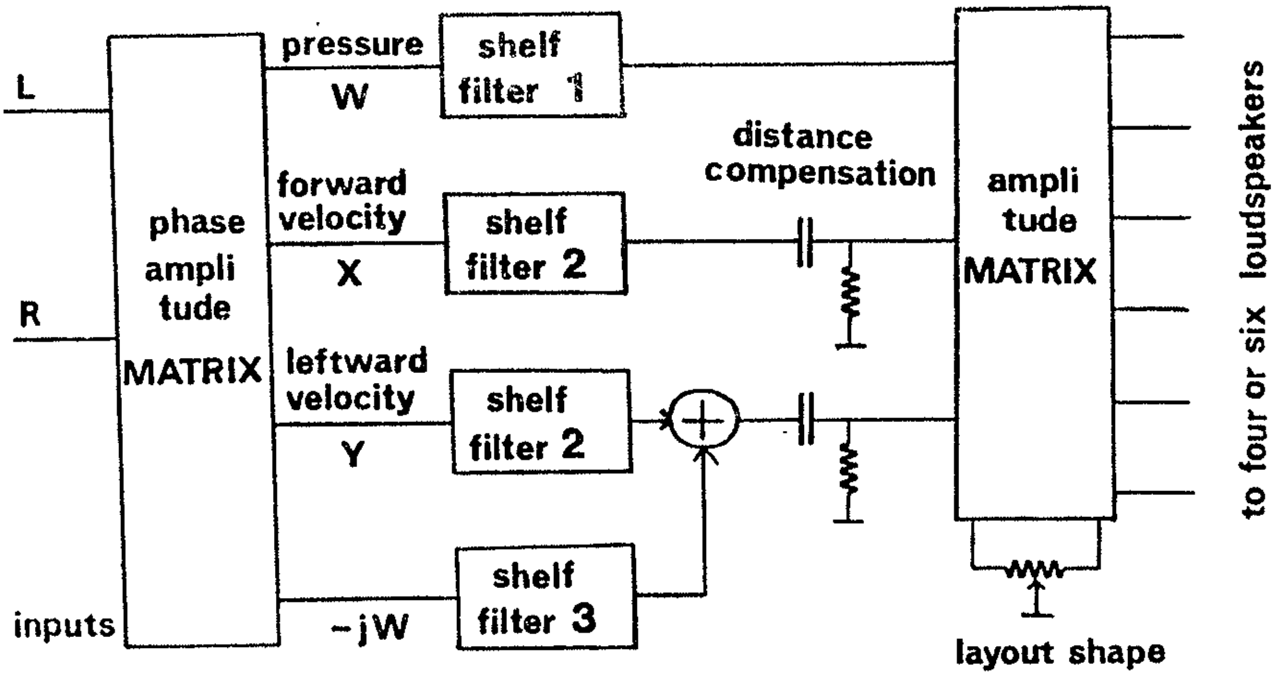

Figure 3. Overall decoder architecture.

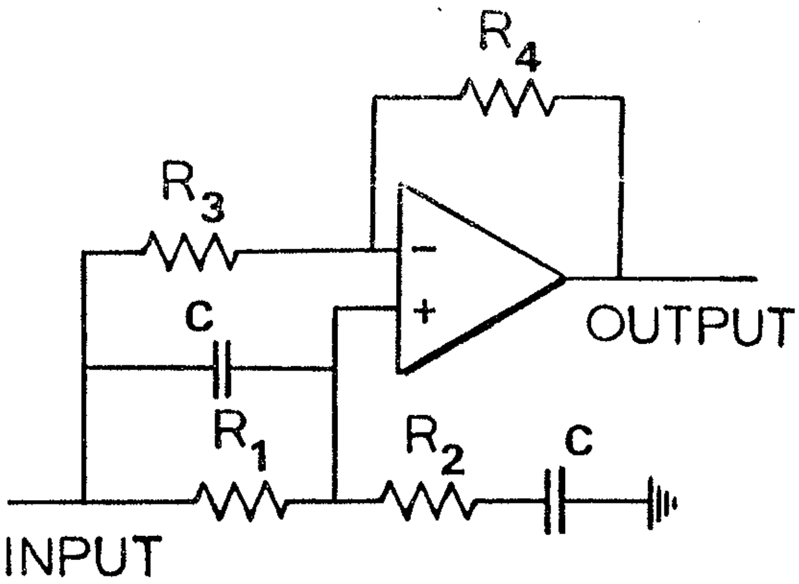

Figure 4. Two-pole all-pass phase shifter circuit.

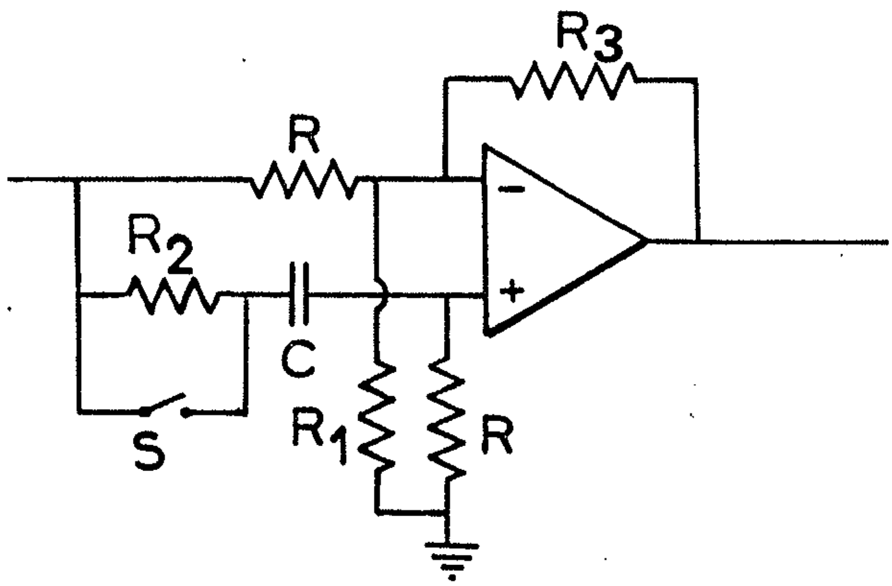

Figure 5. Phase-compensated shelf filter, with switchable shelf characteristic and resistive input impedance.

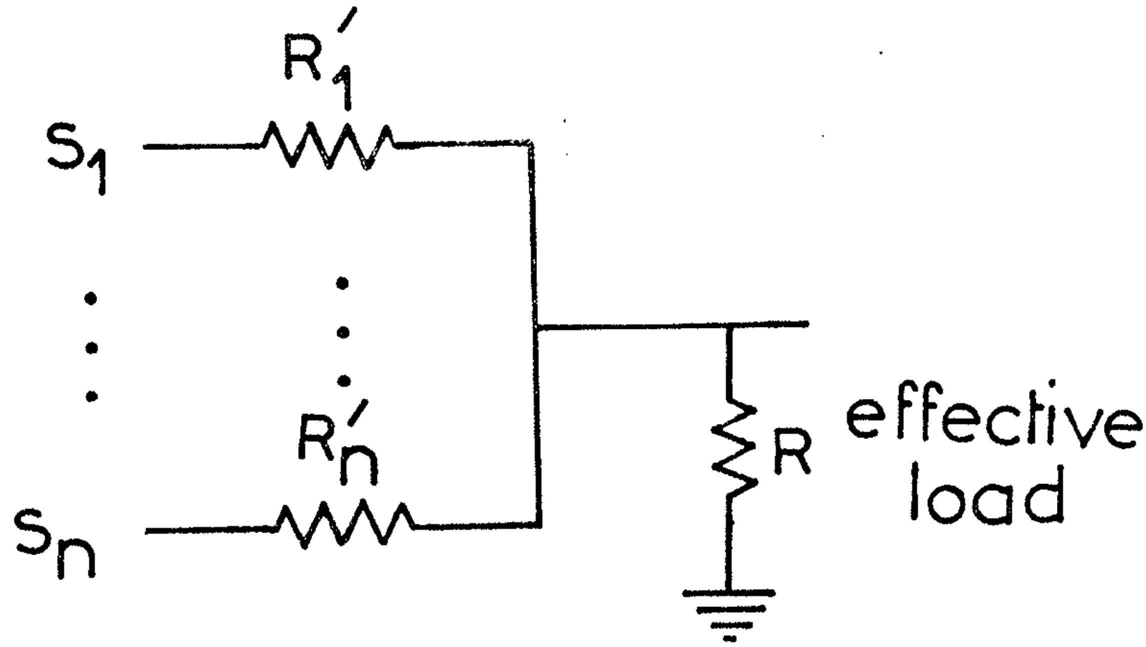

Figure 6. Resistor matrix loaded by shelf filters.

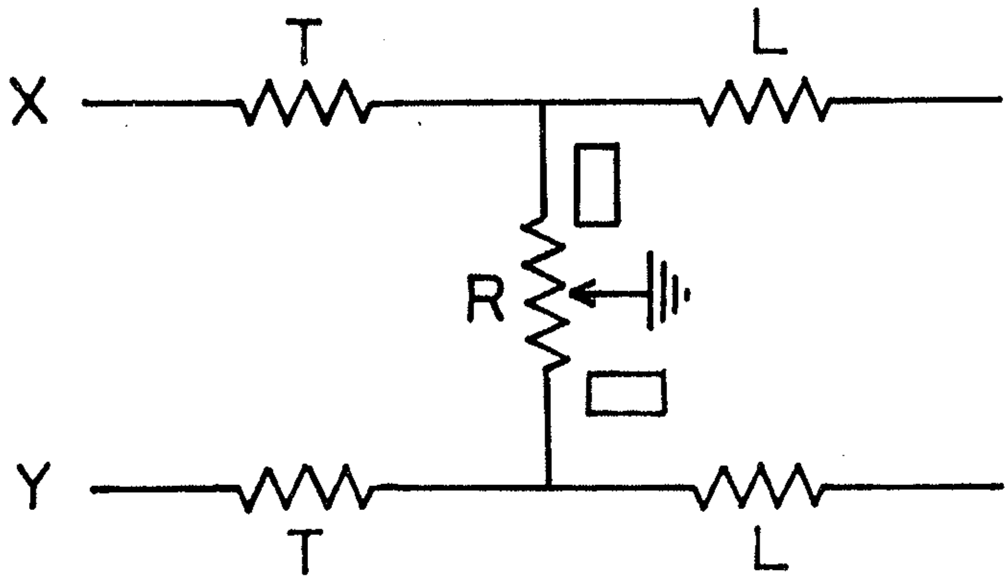

Figure 7. Loudspeaker layout control for rectangle layouts.

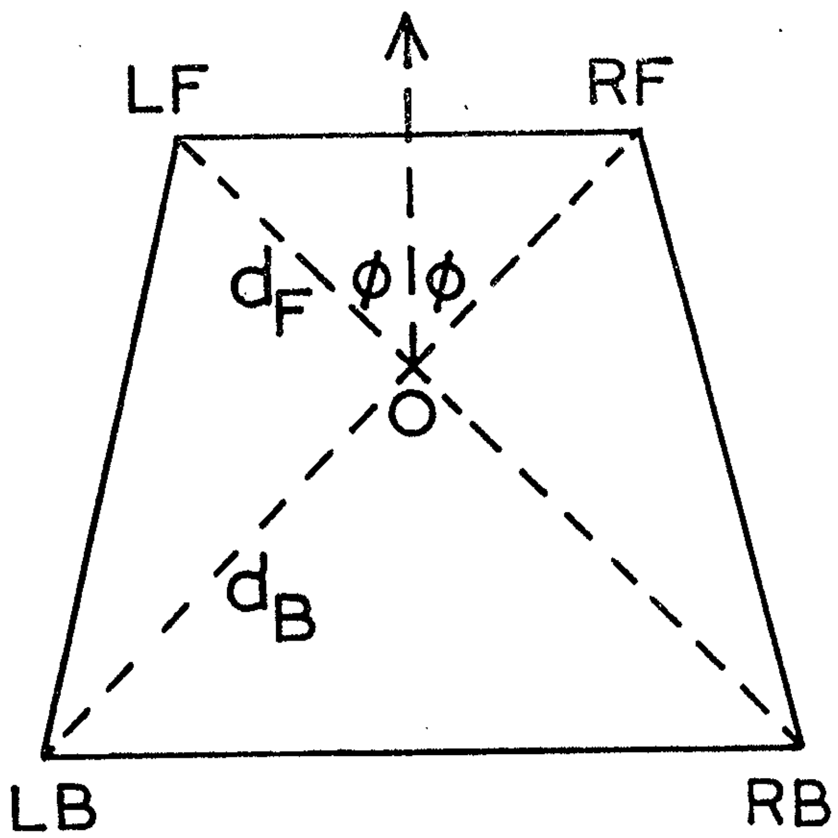

Figure 8. Trapezium loudspeaker layout

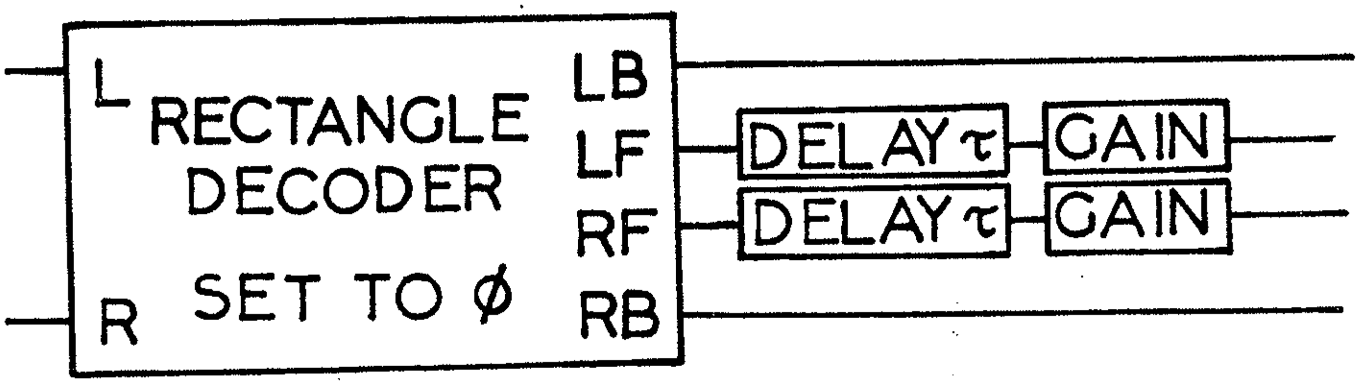

Figure 9. Delay compensation of rectangle decoder for trapezium speaker layout of figure 8.