The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound

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

Tuesday 4 March 1975

So-called "coincident" microphone arrays are often used for recording stereo or surround sound. Experience has shown that two of the main causes of poor image localisation and of spurious secondary images are: (i) the usual capsule spacing of 3 to 10 cm, and (ii) poor polar diagrams and polar phase responses in the treble. These defects also cause a significant degradation in the tonal quality if a stereo or surround sound recording is mixed down to mono or matrixed either to modify the recording's stereo effect or for 2-channel quadraphonic encoding.

Because there are practical limits on the smallness of low-noise directional microphone capsules, it has not hitherto been possible to design precisely coincident microphone arrays. Such precise coincidence (ideally to better than 5 mm) is necessary whenever signals are to be mixed or matrixed, in order to avoid high-frequency cancellation effects. We describe the use of a tetrahedral array of capsules with electronic spacing compensation to ensure effective coincidence of all outputs.

Let the polar diagram of an individual capsule in an array (with respect to the capsule's notional centre) be $f(x,y,z)$, where (x,y,z) are the direction cosines of the direction from which a sound arrives, with the x-axis forward, y to the left side, and z upward, and with $x^2+y^2+z^2=1$. It is not widely realised that in absolute physical terms, there is no such thing as the "position" of a capsule. If a capsule has its notional centre at the coordinates (u,v,w), then its polar diagram relative to the origin of coordinates at angular frequency ω is

$$f(x,y,z)\exp(j(ux+vy+wz)\omega/c) \qquad (1)$$

where c is the speed of sound (340 m/s) and the factor multiplying $f(x,y,z)$ is simply a time delay due to the extra distance $-(ux+vy+wz)/c$ travelled by the sound.

Thus if we have an array of n capsules with polar diagrams $f_i(x,y,z)$ relative to their notional centres (for $i = 1,2,\ldots n$), and whose notional centres have co-ordinates (u_i, v_i, w_i), then we may compute the effective polar diagram given by an output consisting of the various individual capsule outputs added together with respective frequency-dependent gains $g_i(\omega)$. This polar diagram is given by (1) to be:

$$\sum g_i(\omega)\, f_i(x,y,z) exp\left[j(u_i x + v_i y + w_i z)\omega/c\right] \qquad (2)$$

In principle, one may use (2) to design arrays and frequency-dependent matrixing circuits to give outputs with any desired coincident polar diagrams $F_(x,y,z)$.

In practice, the following problems arise: (i) For a 4-capsule 4-output array, one has to design 16 different frequency responses $g_i(\omega)$. (ii) Even if the actual polar responses $f_i(x,y,z)$ of capsules close together were known exactly, the approximation given by (2) to the desired polar responses from so few capsules may even at best be a poor approximation unless the shape of the array is carefully chosen. (iii) The polar diagram $f_i(x,y,z)$ will depart from its theoretical form in a not entirely predictable manner, and one then has the impossible task of 'tweaking' 16 complex (in both senses) frequency responses to get the (unknown) best match to the desired polar diagrams.

We have solved these problems as follows. We assume that the capsules are axially symmetric, and lie on a sphere with the axes pointing outward from the centre. We first consider an idealised "spherical microphone" uniformly covered with many capsules. Each spherical harmonic polar diagram $\phi(x,y,z)$ is derived by adding the outputs of all capsules with respective gains $\phi(X,Y,Z)$ for the capsule in the direction (X,Y,Z). Using Schur's lemma from the theory of group representations [1], one may show that for ideal spherical microphones, no matter what their radius or capsule polar diagram, there is *no* cross-talk between different spherical harmonic outputs thus derived, and that all spherical harmonic outputs of the same order require precisely the same frequency response correction for a flat response. Thus, when only omnidirectional and figure-of-eight outputs (i.e. 0^{th} and 1^{st} order) are derived, only two frequency responses, one for each order, require empirical adjustment to account for departures from theoretical behaviour, rather than 16.

One may show that the best practical approximation to a "uniform covering" of the sphere with few capsules is obtained by placing them at the points of an "efficient" numerical integration rule for the sphere surface [2], [3], giving them respective gains equal to the "weights" of the relevant rule. The simplest such array is a regular tetrahedron, giving omnidirectional and three orthogonal figure-of-eight outputs. The general scheme of the array is shown in fig. 1. Analysis shows that the omni output is contaminated by spurious polar responses of order ≥ 3, and the figures-of-eight by spurious responses of order ≥ 2. Denoting n^{th} order spherical Bessel function [4] by $J_n(x)$, we have from [5] the following mathematical result:

$$\exp(jkx) = \sum_{n=0}^{\infty} j^n (2n+1) P_n(x) J_n(k) \tag{3}$$

Since the n^{th} degree Legendre polynomial $P_n(x)$ is, as a function of (x, y, z) an n^{th} order spherical harmonic, we may substitute (3) into (2) to compute that the frequency responses of the 0^{th} and 1^{st} harmonic components of the nominally zero and first harmonic outputs of a tetrahedral array of cardioids are respectively:

$$\left.\begin{array}{ll} J_0(\omega r/c) + jJ_1(\omega r/c) & \text{(zero'th order)} \\ \frac{1}{3}J_0(\omega r/c) + jJ_1(\omega r/c) - \frac{2}{3}J_2(\omega r/c) & \text{(first order)} \end{array}\right\} \tag{4}$$

where r is the distance of the effective centres of the cardioid capsules from the array centre.

Above a limiting frequency $F \approx c/\pi r = 10.8/r$ kHz (r in cm), the polar diagrams become severely contaminated by higher order spherical harmonics, and it is found best to equalise the nominal omni and figure-of-eight outputs for an approximately flat response to homogeneous random sound fields. Calculations show that before any equalisation, the omni and figure of eight outputs from a regular tetrahedral array of cardioids with $r = 1.47$ cm have frequency responses in random sound fields as shown in fig. 2. This should convince sceptics of the fallacy of assuming that "nearly coincident" is good enough. In order to ensure perfect coincidence below the limiting frequency, the phase responses should also compensate for those in equ. (4).

Designs involving all capsules being in one plane [6] cannot give good polar diagrams in other than horizontal directions. Since most of the incident sound will be non-horizontal, this would tend to give a colored sound quality in the reverberant surroundings in which a coincident microphone is likely to be used.

Prototype microphones have been constructed as a part of the development program for the N.R.D.C. ambisonic system of surround sound. The microphones are effectively coincident up to about 7 kHz and are subjectively well behaved above that. Trial recordings show a stability and lack of ambiguity of 4-speaker sound localisation much superior to the best stereo, even for non-central listeners and for sounds at the sides. A control unit for mixing microphone outputs down to mono, stereo or quadraphonics has been designed. This permits one to record the information from the microphone onto 4-channel tape, and to select any desired stereo coincident microphone technique (including adjustable angle of vertical tilt) at any later time. For the first time, this allows a full mixdown capability off coincident microphones. This, of course, is only practical because the microphones are, in effect, precisely coincident.

Acknowledgments. I would like to thank Peter Craven, who coinvented the spherical microphone principle, Calrec Audio, both for their constructional work and their excellent cardioid capsules, Wilf Jones for some of the circuit design, and Peter Fellgett for his encouragement.

References

1. I.M. Gel'fand, R.A. Minlos, Z. Ya. Shapiro, "Representations of the Rotation and Lorentz Groups and their Applications" (Pergamon Press, Oxford, 1963) pp. 177,178,36-47

2. A.D. McLaren, "Optimal Numerical Integration on a Sphere", Math. Comp., vol.17, pp. 361–383 (1963)

3. A.H. Stroud, "Approximate Calculation of Multiple Integrals" (Prentice Hall, Englewood Cliffs NY, 1971)

4. M. Abramowitz, I.A. Stegun, "Handbook of Mathematical Functions" (Dover Inc., N.Y., 1965), Ch.10 pp. 435–465

5. G.N. Watson, "A Treatise on the Theory of Bessel Functions" (2nd edn., Camb. Univ. Press, Cambridge, 1944) p.128

6. T.Yamamoto, "Quadraphonic One Point Pickup Microphone", J. Audio Eng. Soc., vol. 21, pp. 256-261 (May 1973)

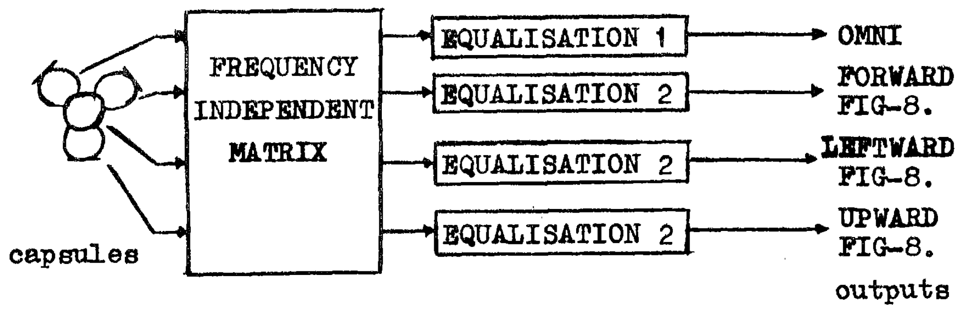

Figure 1: Schematic of tetrahedral microphone array with precisely coincident outputs.

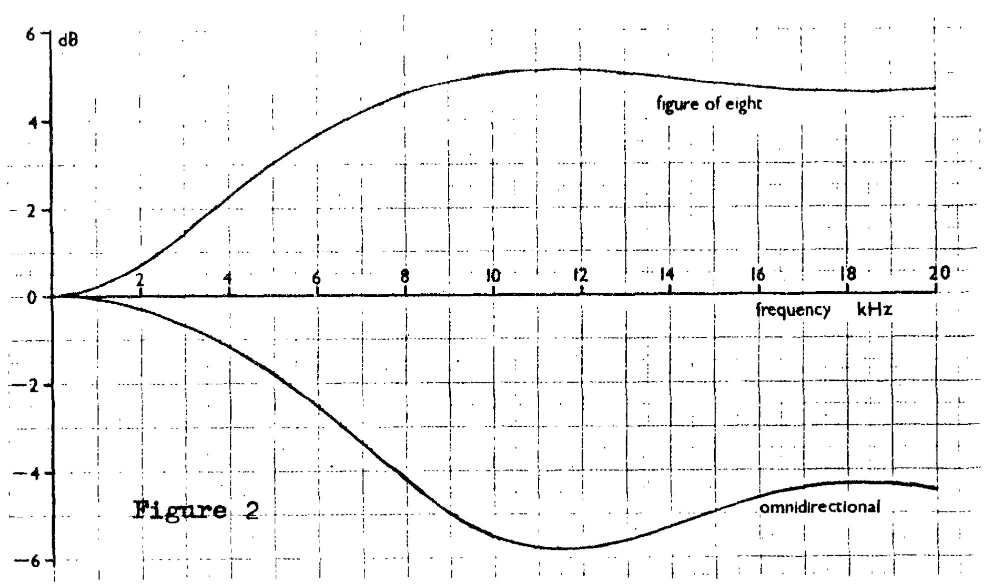

Figure 2: Random-incidence frequency response of the omni and figure-of-eight output from a regular tetrahedral array of cardioids with $r = 1.47$ cm.