Ambisonics

Ambisonics is a technique for recording and reproducing spatial information captured by an array of microphones. The typical arrangement of transducers is tetrahedral, using four cardioid capsules arranged on the faces of a pyramid. By mathematically combining the individual signals, the original sound field can be recreated by an array of loudspeakers. The concept is similar to the operation of the mid-side microphone only in three dimensions, however the actual implementation is far more complicated if the desire is to reproduce exactly the original sound field so as to be perceived precisely the same way.

The direct outputs from the four capsules are called A-format and are not directly useful. A combination of the individual signals is used most commonly, the B-format which combines all elements into a summed signal that represents the total sound pressure as if the array was a single pressure microphone, termed W. Also sent in B-format are X (equivalent to a pure gradient figure-8 facing forward), Y (a figure-8 facing left) and Z (a figure-8 facing upward.) The B-format outputs are also corrected by high-pass filtering to reduce low frequency sensitivity. Further correction is employed because the capsules are not identical and at higher frequencies the capsules cannot be considered as coincident. By combining the B-format elements, phantom microphones can be simulated and oriented after the fact, much like the stereo separation is adjustable with mid-side systems.

Mathematically, a simple B-format signal encoder (panner) can be represented as:

$$W = \frac{S}{\sqrt{2}}$$
$$X = S \cdot \cos\theta \cos\phi$$
$$Y = S \cdot \sin\theta \cos\phi$$
$$Z = S \cdot \sin\phi$$

where S is the source signal, θ is the horizontal angle (azimuth) and ϕ is the vertical angle (elevation.) This representation is called first-order Ambisonics and in practice produces somewhat indistinct phantom images of sound sources. By increasing the complexity of the system, second and third order systems can be generated that produce more convincing imaging. Theoretically, even third order Ambisonics cannot reproduce exactly the full 3-D sound field at higher frequencies. In order to perfectly simulate the incident sound field up to 20 kHz in a space the size of a human head, an order of 32 is required which would require a 1000 loudspeaker array. Fortunately, with proper DSP corrections a third order array is capable of convincing spatial placement for the frequency range important to music.

It is possible to set up symmetrical pairs of loudspeakers in arbitrary arrangements that may be driven by appropriately decoding the B-format signals to suit the geometric

arrangement of speakers. The same recording can be reproduced in any such setup by using an appropriate decoder, allowing a recording to be played similarly on many different speaker arrays. Besides native and converted B-format recordings, Ambisonic encoders (panners) can create Ambisonic playback from traditional mono recordings like multi-track recordings. This allows three-dimensional mixes to be created from any traditional multi-channel recording and greatly expands the options for creative mixes.