AudioAmateur

STALKING TRUE AMBIENCE

AMBISONICS COMES OF AGE

BY WILLIAM SOMMERWERCK

An Ingenious Man who had built a flying-machine invited a great concourse of people to see it go up. At the appointed moment, everything being ready, he boarded the car and turned on the power. The machine immediately broke through the massive substructure upon which it was builded, and sank out of sight into the earth, the aeronaut springing out barely in time to save himself.

"Well," said he, "I have done enough to demonstrate the correctness of my details. The defects," he added, with a look at the ruined brickwork, "are merely basic and fundamental."

On this assurance the people came forward with subscriptions to build a second machine.

—(the great American satirist and misanthrope) Ambrose Bierce

MR. BIERCE's observation sums up the history of quadraphonic sound. Manufacturers dumped millions of dollars into its development without any serious examination of its fundamental assumptions. Had anyone "in authority" thought twice about what was going on, *none*

of the proposed systems would have been used. All the discrete formats (Q-4 and Q-8 tapes, CD-4 records) and the two most popular matrix systems (SQ and QS) are based on incorrect assumptions about the way we hear directional effects.

Many people knew these things,

but none of them worked for the major record companies, and those who did speak up (e.g., John Eargle and Duane Cooper) were ignored. Surround reproduction ought to have become part of any good playback system and *especially* of perfectionist systems. The shortsightedness of en-
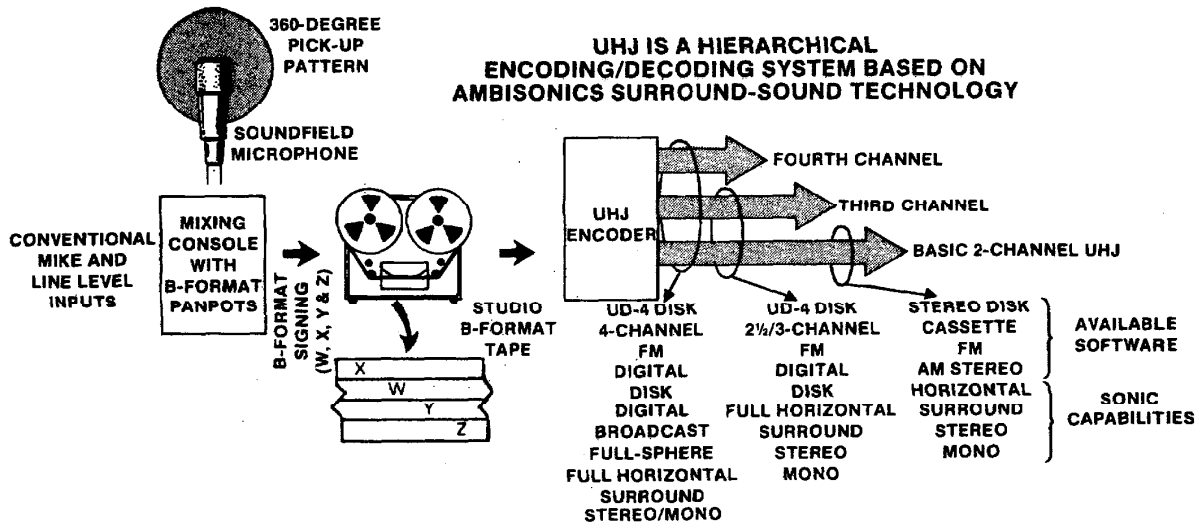
FIGURE 1: The UHJ hierarchical system of encoding and decoding directional sound information within the Ambisonics technology.

gineers and the greediness of recording companies guaranteed the slow, agonizing demise of quadraphonics.

If quad is dead, why am I writing an article about Ambisonics? Ambisonics is *not* a quadraphonic system. The only common characteristic is the use of at least four speakers. Ambisonics is a universal recording technology [*Fig. 1*]. I can't emphasize this too strongly: Ambisonics can create almost any sound effect imaginable with amazing directness and simplicity. These effects range from a highly accurate reproduction of a concert hall's ambience, including the vertical reverberation, to placing individual sound sources anywhere around the listener—even above and below—with any desired size, motion or apparent distance.

In the Beginning

To see why Ambisonics is right and the other systems are wrong requires more than a simple explanation of ambisonic technology. A brief discussion of the history of surround-sound technology will make it easier for you to understand where most of the designers went wrong.

The first surround-sound recording was the sound track of *Fantasia*, made in the early 1930s. This recording included seven audio tracks on a separate film reel, with additional control tracks to expand the dynamic range and to determine which speaker received a particular channel.

Where did the ''Fantasound'' system come from? Although there were some experiments with stereo and binaural techniques in the late 19th century, I have found no references to surround experiments. Even the Bell and Decca stereo patents of the early '30s make no mention of surround effects. (Alan Dower Blumlein, the scientist who filed Decca's patent, is generally given credit for ''inventing'' stereo—i.e., reducing it to a workable system. His patent, which is cited by almost every surround-sound patent, anticipates matrixing, with-height effects and even the 45-45 disk recording system. But nowhere does it say anything about horizontal surround.) It appears, then, that Disney Studios—or, more likely, RCA, who built the system—should receive the credit for surround sound.

After *Fantasia*, surround sound dis-

appeared for another quarter of a century, only to resurface in the early 1960s with the Revere-Wollensak tape-changer system. This system included a third track along the center of the tape to carry ambience. No one knows who developed the system, and there is no evidence of experimental three-channel recordings.

Let There Be Light

Not until 1969 did surround sound make its first ''modern'' appearance. At that time, Acoustic Research (AR) and Vanguard announced their experiments using two extra channels to convey the acoustic character of the performing space. The idea was to use four-channel, open-reel tape, with tracks 2 and 4 carrying the ambience. Two mikes at the back of the hall would pick up reverberation.

Sonic nirvana! Nonetheless, many cynics felt that quad had been developed just to sell more equipment. To use the system, a listener
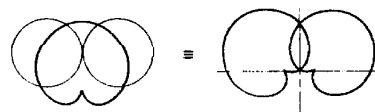

FIGURE 2: You can use the Middle-Side (M-S) technique to obtain conventional left and right channels.

had to buy a new tape deck *and* another amplifier *and* the special tapes from Vanguard.

Because nothing important was happening in audio at this time, audiophiles greeted this announcement from two of the industry's leaders with great interest. Thus the industry stumbled into the modern age of surround sound. ''Stumbled'' is the right word, for with one stroke, Vanguard and AR managed to create two myths about surround reproduction.

Surround-Sound Error #1

Ambience is to be recorded by placing the ambience microphones to the rear of the hall.

No matter how directional the rear mikes are, there is no way to prevent them from picking up some of the direct sound. The difference in acoustic path length between the

front and rear mikes delays this sound, so you hear the direct sound twice. This never occurs in real life because your head can be in only one place at a time (physically speaking). Such recordings have a fairly obvious slap echo.

It is surprising how many otherwise knowledgeable recording engineers made—and still make—recordings this way. The Ambiphon (Sonar) tapes engineered by Mitch Cotter use widely spaced omnis and have a slight, but noticeable, slap. Marc Aubort and Joanna Nickrenz, the producers of most of the Vox/Turnabout quad recordings, also use this technique. Again, you often hear a distinct front/back effect as the direct sound passes each set of mikes.

However pleasant the overall effect (it *does* enlarge the acoustic space), it is objectively incorrect because the listener is hearing the direct sound from one position and the reverberation from another. If the producer wants to use spaced omnis, he or she should place the ambience mikes below the front mikes (to prevent acoustical interference), facing into the hall. These mikes must be directional, and acoustical baffling should be used to block any direct sound that might reach them.

In one respect, it is easy to understand this kind of mistake. After all, conventional stereo consists of sampling the sound field at two points. What could be more natural than sampling it at four points for surround sound? This leads to the second surround-sound error.

Surround-Sound Error #2

If there are x loudspeakers, then there must be exactly x distinct channels of information feeding them.

This is a more subtle error. We are so brainwashed by 30 years of two-channel reproduction that we quite naturally assume that each speaker requires a discrete signal. A mathematical view of the subject suggests that we might be overlooking something. Two points define a line, and in regular stereo, the image is strung out in a line between the speakers. Only three points are needed to define a plane. Why, then, should we need four signals to position the sound around the listener? Wouldn't three be enough?

At first blush, you wouldn't think so. How do you get four speaker feeds out of three signals? You can understand the technique for doing this by examining how you get two speaker feeds out of two signals. Take stereo FM. Speaker-feed signals are not directly transmitted. Instead, a sum (L+R) signal in the baseband provides compatible reception for mono receivers, and a difference (L−R) signal in the subcarrier allows you to regenerate L and R by taking the sum and difference of these two signals.

A similar technique, called M-S (for Middle-Side), was common in the early days of stereo recording. One microphone, usually a cardioid, faces the orchestra to provide a mono pickup. A second mike, always a figure-8, is arranged sideways to specify the direction of the arriving sounds. By adding and subtracting these two mike outputs, you can obtain conventional left and right channels (Fig. 2). By simple extension, you could set a third figure-8 mike facing forward to obtain front-back directionality. Its output would be added and subtracted to each of the two signals created from the first pair of mike outputs.

This would create a total of four signals. Using this approach, each added transmission channel enables you to double the number of speaker feeds. This does not necessarily mean that these speaker signals will produce a correct image. It does show, however, that there might be better ways to transmit surround sound than by direct speaker feed.

To Market

There was just one catch to this "practical" system from AR and Vanguard: it worked only with open-reel tape, which has never been a mass-market item. If there were any profit (aha! the magic word) to be made in quad, it had to come from phonograph records. But no one had found a reliable way to record four channels on a disk.

JVC was experimenting with one method that added an ultrasonic carrier to provide an ambience channel. The AR-Vanguard demos encouraged JVC to speed up its research, and about a year later, they announced CD-4 (Compatible Discrete Four-Channel) records. The technology involved in producing these carrier disks—improved vinyl formulations, half-speed mastering, superior cutting and playback styli, wideband pickups—made a major contribution to disk recording and reproduction. It is a shame the same cannot be said for CD-4.

With most of the pickups designed for this system, CD-4 stood for "Continuous Distortion to the Fourth Power." Many recordings showed a kind of gurgling midrange mush, combined with severe high-frequency splatter and breakup. As one reviewer put it, you were constantly on the edge of your seat "waiting for something to happen." The disks varied widely in quality, with the earliest and latest giving the best performances. Even when they played properly, however, a high level of coloration falsified instrumental sound. In short, the sound quality of CD-4 was far below that of the best stereo records. Ironically, CD-4 paved the way for today's high technology disks.

The biggest breakthrough in the development of quadraphonic sound, however, came from Peter Scheiber, the musician-engineer who invented matrixing. Matrixing made it possible to put the four channels of the master tape on a conventional disk.

From an over-simplified mathematical point of view, a matrix transforms one set of discrete quantities into another set of quantities. In this case, the four original channels are transformed into two, which can be readily transmitted via any two-channel medium. The transformation is made in a controlled fashion, so it should be possible to reverse it.

Loss of Information

The catch is that we started with four pieces of information, but ended up with only two. It is a law of nature that we cannot unambiguously solve for four unknowns with only two equations. Why do I refer to audio signals as "unknowns"? Simply because they are just that. If a listener knew exactly what they were, he or she would not have to buy the recording. This is a fundamental principle of information theory. Because the intended receiver does not know the exact nature of the information being transmitted, such information is subject to noise and distortion. If the receiver knew exactly what would be sent, it could be received with perfect fidelity. But if he or she knew what would be sent, there would be no need to send it.

To see this, consider the SQ encoding equations. Lt and Rt signify the two transmitted signals, which are the left and right track, respectively. Lf, Rf, Lb and Rb should be self-explanatory.

$$Lt = Lf - (0.707 \times j \times Lb) + (0.707 \times Rb)$$

$$Rt = Rf + (0.707 \times j \times Rb) - (0.707 \times Lb)$$

where j is a 90-degree phase shift, which most matrix systems use to reduce encoding ambiguity.

You can play with these equations for the rest of your life, but you cannot manipulate them to get each channel back by itself. For example, if you try to solve for Lf, there will always be some Lb and Rb mixed in. This represents undesired crosstalk among the original four channels. A general analysis shows that, for any given channel, you can design a matrix to provide complete separation from one other channel, with the remaining two channels appearing as crosstalk, 3dB down. You can manipulate the matrix coefficients to reduce this crosstalk, at the expense of greater crosstalk in the direction of "full" separation. Unfortunately, the best you can do is about −7dB crosstalk, all around. This is not enough to give subjectively perfect separation. Because of this, virtually every technical advance in quadraphonics in the past decade has been directed at enhancing matrix separation, first with Vario-Matrix, then with Audionics' Shadow Vector and the Tate Directional Enhancement System for SQ.

Garbage In, Garbage Out

And *that* was the mistake. Channel separation has *nothing whatsoever* to do with good surround sound. The point of surround recording is accurate reproduction of recording site acoustics.

To put it a bit differently, there is only one criterion for a "good" surround-sound system: does the system allow the acoustical character of the performing space to be accurately reproduced? Because this requires highly accurate localization in all directions, a system that is good at recreating ambience will also excel

at creative or arbitrary effects. The discrete master tape, which CD-4 records and the output of advanced matrix decoders are supposed to mimic, is *not* a valid reference. Only live sound and the producer's intentions can be considered as such.

Many of you must be thinking that this is a lot of semantic hair-splitting. If separation of the original four channels on the master tape is maintained, won't the directionality be correct? The answer is no. That reasoning assumes that what is on the tape correctly reproduces the directional character of the sound source, real or imagined. But the tape does not provide accurate reproduction because "quad" recordings were always made with psychoacoustically incorrect techniques. This will make more sense when you examine the next surround-sound fallacy.

Surround-Sound Error #3

The proper way to encode directionality is to pan the signal between adjacent pairs of speakers.

For example, if a sound is to come to the immediate right of the listener, it should be fed at equal levels to the right front and right rear speakers. This technique is known as Pair-Wise Mixing (PWM). This mistake, more than any other, has blocked the proper development of surround sound. The PWM assumption is so dreadfully wrong that it is amazing so few people have blown the whistle on it. John Eargle's *Sound Recording* points out this error in his chapter on quadraphonics, and Katsumi Nakabayashi of the NHK mentions it in the 4/75 *JAES*.

Scientific American has recently published *The Science of Musical Sound*, by John R. Pierce. (Pierce spent many years at Bell Labs and is one of the great scientist/electrical engineers of this century.) In his chapter on sound reproduction, Pierce suggests that conventional quad systems are wrong and that the only "proper" approach is to extract the pressure and velocity components of the sound field and to adjust their levels at the listener's ear to match those in the live field. His suggested speaker layout places the speakers at the points of a tetrahedron, with one speaker directly above the listener and the other three below. As I will show later, this is the ambisonic ap-

proach, although its speaker layouts are more suited to domestic conditions. *Figure 3* shows a more practical arrangement.

We are accustomed to hearing PWM in almost all stereo recordings. A given sound is positioned between the speakers by adjusting its relative amplitude in each channel. Paradoxically, this technique is used both in
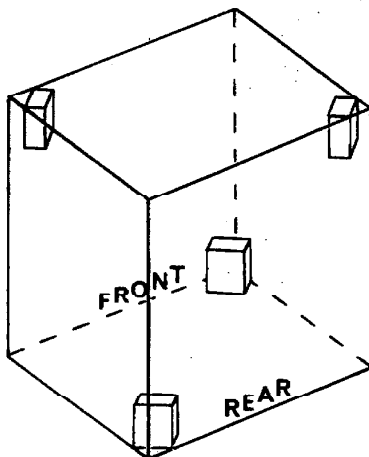
FIGURE 3: Tetrahedral loudspeaker layout, embedded in a cube.

gimmicky multitracked studio recordings and in single-point "purist" recordings. In studio recordings, each instrument or performer is assigned its own track, which results in dozens of little mono recordings. When the two-track cutting master is made, a pan pot adjusts the relative channel levels to position the sound where the producer wants it.

Oddly, the same effect results when making a single-point recording. With two closely spaced cardioid mikes, there is almost no separation to introduce phase or timing differences. Because the mikes are directional and are aimed in different directions, anything other than a central source will produce a different output level from each mike. This creates the amplitude differences needed for a directional effect.

In both cases, you use amplitude differences to position sound sources. This is called amplitude panning. Of

course, the ear and brain also use phase and arrival time differences for localization. Another form of purist recording technique, spaced omnis, uses these additional cues. As I have shown, however, this technique produces fundamental errors when applied to surround-sound recordings, so I will ignore it.

Amplitude Panning

Amplitude panning is used almost universally in conventional stereo. If someone asked you how to pan the sound around the listener, what would you suggest? The most obvious technique is to treat adjacent speakers as pairs and to pan the sound between them. Guess what? It doesn't work.

To see this for yourself, try the following experiment. Set your system on mono or play a mono recording or broadcast. Face the speakers. If your system is set up correctly, the sound should appear to be coming from dead center. Turn 90 degrees to the left or right. Now where does the sound come from? It is *not* from dead right or left. There is no clear-cut source—just an amorphous blob of sound—or it might appear to jump from one speaker to the other, particularly as you rotate the balance control to "pan" the sound. The distinction between face-on and sideways listening should be obvious, but if it is not, try adding a third speaker in the middle and use it as the reference. If you still cannot hear the difference, I would suggest selling one of your speakers and going back to mono.

If you own the Audionics Space & Image Composer, the Fosgate Tate II or a Sansui Vario-Matrix decoder, you might have noticed similar effects. Sitting sideways, you might hear instruments or performers clearly positioned between the (true) front and rear speakers. When you face forward again, the preciseness of their positioning disappears. Sometimes it becomes more difficult even to *hear* the instruments. Likewise, when expanding stereo recordings into a "horseshoe," you might have noticed that the sides are noticeably bereft of sound sources, unless you turn to face the sides. All this is due to the fact that SQ and QS recordings are pair-wise mixed, and the Tate and Vario-Matrix systems blindly follow this paradigm in their decoding action.

What about live recordings? What happens if you place four cardioid mikes in a square, at 90-degree angles to each other? Won't that work properly? No. You are still using pair-wise mixing. Reverberant sounds arriving from the sides are amplitude panned between the front and rear mikes on that side. Thus, their directionality is not correctly reproduced on playback. The side reverberation is critical to a proper appreciation of the exact sonic character of the performing space. I will discuss exactly how this psychoacoustic failure affects the reproduction in the section on miking technique.

It should now be apparent that every quadraphonic recording ever made has its intended directionality incorrectly encoded. *Every one*. Regardless of whether manufacturers used coincident mikes, spaced omnis or pan potting, they failed to recognize the fundamental engineering approach needed for good surround sound. That approach is not a question of how to transmit four sound channels, but rather of how to create an accurate, stable sound field in the listening room that closely mimics what the listener would have heard live or what the producer imagined. In other words, what technology gives accurate imaging? Instead of focusing on this problem, engineers focused on poor separation, an inherent weakness of the matrix system, which was the most practical commercialization of quad.

The problem of obtaining subjectively correct imaging is not new. It has been around for more than 50 years, since the advent of stereo. Few (if any) recording engineers can tell you what kind of mikes you should use and where you should place them so that the listener hears the performers at the correct angular position, at the right apparent distance and with the ambience in its correct relationship. The reason for their inability to do so is that no systematic study of stereo imaging or the technology required to produce accurate or arbitrary effects has ever been conducted.

Getting Their Act Together

The *raison d'etre* of surround sound is the ability to position a sound image at any point around the listener, including above and below. Any sur-

round-sound system must include a technology for controlled imaging. Otherwise, it is a waste of time, money and engineering effort. Unfortunately, all the proposed discrete and matrix systems *were* a waste because they were based on the wrong assumptions about the way the ear and brain determine directionality.

In approaching the design of a psychoacoustically correct surround-sound system, forget, for a moment, mono and stereo compatibility. Also ignore the idea that four discrete signals should feed four speakers. Instead, consider the recording and playback chain as a whole to see what technique of encoding and decoding directionality gives the best results.

First, let's define encoding and decoding directionality. The former refers to the way we specify the amplitude and phase of the signals on our transmission channels for any particular source direction. The latter refers to the signal processing necessary to produce speaker-feed signals that accurately reproduce that directionality.

Ideally, the encoding and decoding should be independent of each other. This means that in addition to providing a precise specification of directionality and using the least transmission space possible, the encoding should not limit the decoding process. That is, it should not restrict you to a fixed number of speakers or special speaker locations. The listener should have some freedom in speaker placement and the ability to add speakers for greater positional accuracy and a broadened listening area.

The encoding must not carry implicit assumptions about how it is supposed to be decoded. As we gain increased understanding of the hearing process, it might be possible to create improved decoders for superior imaging accuracy from all existing material. This approach is in sharp contrast to discrete quad, where exactly four speakers are arranged roughly in a square, and the encoding technique (PWM) cannot provide correct imaging.

It is important to point out that none of the quadraphonic "systems" are really systems at all, as they do not specify directional encoding in a way that can be correctly decoded. In essence, whether discrete or matrix,

they are simply ways of transmitting four uncorrelated channels of information, without any regard for positional accuracy.

Compare this with color TV, which *is* a true system. Broadcast standards focus on producing a signal that accurately represents the colors of the original scene, within the limits of existing technology. Because nothing is said about how this information is to be displayed on the home receiver, there is room for improvement at both ends. A modern Philips camera produces a noticeably more transparent and vivid picture than an early RCA, and this improvement is visible on any receiver. Likewise, the development of new phosphors has broadened the range of achievable colors, enhancing the reproduction of existing material.

The same general principles apply to surround sound. Certain classes of encoding schemes are fundamentally correct and will allow accurate directionality in playback. Similarly, the best playback techniques require encoding that is not limited by PWM or other erroneous techniques. Any optimized playback system should also allow use of as many speakers as you would like to improve the accuracy of the directional effect.

The Right Stuff

Not long after Scheiber's initial work, one of the seminal papers on surround sound appeared in the *Journal of the Audio Engineering Society* (Volume 20, Number 5, 1972, p. 346). The paper presented the Universal Matrix (UMX) system and was written by the widely known acoustical engineer Duane Cooper (then of the University of Illinois and best known for the Cooper Time Cube, an early digital delay system) and Takeo Shiga, of Nippon Columbia (Denon to us). It is important for two reasons: it introduced the first correct way to encode directionality, and it addressed the question of whether a given encoding/decoding produced the desired directional effects.

The encoding technique they used is known as azimuthal harmonic synthesis, which is a variation of Fourier analysis. Remember that Fourier showed how any repetitive wave motion could be analyzed into a fundamental frequency, plus harmonics. Each waveform has its own

characteristic pattern of harmonics, and the harmonics have distinctive phases. Each waveform also has an average, or DC, component.

If you surround a listener with performers or have the orchestra in front and the reverb all around, you will be able to recognize some pattern in the distribution of sound sources. Each time you trace a circle around the listener, the pattern will repeat. If you could analyze this repetitive waveform, you would have a useful representation of the position of all the sound sources, including ambience. How can you do this?

The easiest way to see this is in terms of microphone patterns. Think of an omnidirectional microphone placed at the listener's position. An omni mike has no directional preference: it responds identically to sounds arriving from any direction. Its output is independent of the direction of the arriving sound: it conveys no directional information. It simply presents the sum of direct and ambient sounds—a mono pickup. Walking around the mike's pattern, you can see that the output is constant, in the same way a DC signal is constant. To put it another way, the monophonic component of any sound field is equivalent to the DC component of a Fourier analysis of the field. But what about the AC components?

Obviously, you need a pattern that varies with direction—but what kind of variation? We saw that an omni mike has no variation and that when a DC signal was bent into a circle, it matched the pickup pattern of an omni. Let's look at this idea more closely. The fundamental frequency of a Fourier analysis is a sine or cosine wave. What do you get when you bend a cosine wave into a circle? More accurately, what is the pickup pattern of a microphone whose response varies with the cosine of the angle of incidence? What you get is a figure-8 pattern, with the point of maximum sensitivity facing front (0 degrees).

If you look at a table of cosines, you will see that the cosine of any angle between 180 and 360 degrees is negative. Does that surprise you? It should not. After all, the output of a figure-8 mike is inverted in phase for rear sources. Remember that most figure-8s are ribbon mikes, in which
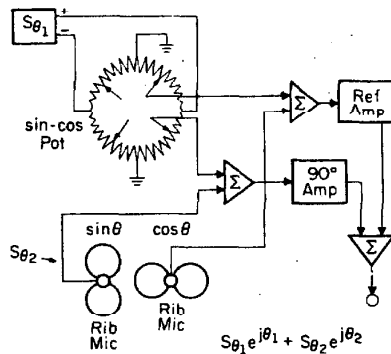
a thin foil ribbon is suspended in a powerful magnetic field. A membrane cannot respond to signals coming from the sides, but will give full output for sounds that are directly to the front. According to simple trigonometry, the effective pressure gradient at the diaphragm varies with the cosine of the angle of incidence—from 1.00 at 0 degrees to 0.707 at 45,

FIGURE 4: Azimuthal harmonic synthesis showing electrical and acoustical mixing.

0.5 at 45 and 0.0 at 90 degrees. Because sound sources at the rear move the ribbon in an opposite direction from those at the front, their outputs are inverted in polarity.

Remember from Fourier analysis that the fundamental and each harmonic have a specific phase angle. This angle is the arctan of the ratio of that harmonic's sine and cosine component amplitudes. To specify the phase angle of the azimuthal harmonics, therefore, you need a sine component, in addition to the cosine component from the figure-8 pickup. Sine(theta) equals cosine(theta − 90). In other words, there is a 90-degree phase angle between sine and cosine. Simply turning the figure-8 mike to the left or right by 90 degrees produces a sine-weighted output (Fig. 4).

Simplicity Is the Key

You can take a similar approach with higher azimuthal harmonics by weighting the sound sources according to the sine and cosine of two theta, three theta, and so on (Fig. 5). These are easily produced from mono sources by adjusting amplitude and polarity. They pose serious problems to live recording, however,
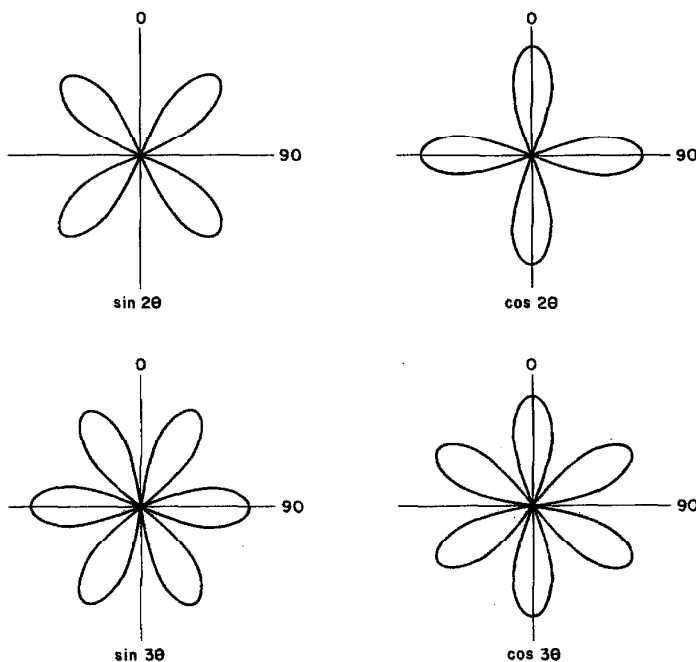
FIGURE 5: Higher-order azimuthal harmonics.

where avoiding multitrack-mono limitations would require a mike with these complex patterns. Consequently, you should not use anything higher than first-order harmonics. Because there are a finite number of channels, let's look more closely at what you can do with a mono component, plus the sine and cosine of the fundamental.

Clearly, while the mono component provides the sound, the figure-8 sine and cosine components indicate directionality. (If this is confusing, refer to the previous discussion of M-S miking. The principle is identical.) Are these three signals enough to specify directionality unambiguously? Let's see. Think of the four cardinal points of the compass. As a sound source moves from one point to the next, the relative outputs of the sine and cosine-weighted channels move in opposite directions, one rising while the other falls. The ratio of their amplitudes determines the relative angular position of the sound source within the quadrant.

Which quadrant is that? Here is where you use the relative polarity. In the left front quadrant, sine and cosine are positive. In the right front sector, cosine is positive and sine is negative. Each quadrant has a distinct pair of polarities. At least from an electrical point of view, it is possible to define the position of a sound source anywhere in a plane with only three channels of information. Any more information is redundant. Anyone who insists that you need four channels to preserve compatibility or to provide full artistic freedom is wrong.

We have now precisely encoded the directionality of any horizontally located sound source with only three signals and no reference to speaker-feed signals. This leaves the door open to using any number of speakers. This is in stark contrast to QUADraphonics, which assumes that acoustic space will be sampled at *four* points and that these *four* samples will be transmitted to the listener's *four* speakers via some *four*-channel medium. Since when is four a magic number?

Of course, the full ambisonic system *is* four channel, but the fourth signal carries up/down information. None of the channels is a speaker-feed signal. Furthermore,

you can show that there is a hierarchy of possible surround-sound systems, where the sound field is sampled with n^2 (1, 4, 9, 16, 25, and so on) patterns. It is a coincidence that the simplest practical ambisonic system needs four channels and that quad reproduction also has four. There is no fundamental similarity between quad and Ambisonics.
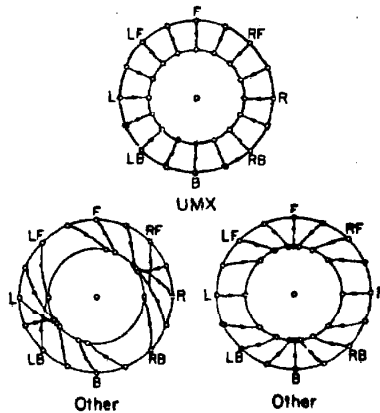
Reprinted, with permission, from the Journal of the Audio Engineering Society (Volume 20, Number 5), June 1972, p. 55.

FIGURE 6: The Cooper/Shiga Universal Matrix (UMX) system compares favorably with the "other two" systems. It satisfies the Makita localization for every position around the listener.

Delivering the Goods

How do you process these signals so that the listener hears correct directionality? Cooper and Shiga made their major contribution in this area. They suggested that these signals be decoded by placing at least four speakers in a circle around the listener. The speakers need not be "toed-in," just equidistant from some center point. The omni signal is fed to all the speakers with identical amplitude and polarity. The front/back figure-8 signal is also fed to all the speakers, except that it is weighted by the inverse of the cosine of the angle the speaker receiving it makes with center front. The same thing happens to the front/back signal, except it is weighted by the inverse of the sine of the angle. Notice that sine and cosine can be positive or negative, and this must be carried through in the weighting.

Therefore, the left/right signals fed to the right speakers will have a negative polarity with respect to the mono signals, since the sine of the angles between 180 and 360 degrees is negative. Similar reasoning shows that the front/back signals fed to the rear speakers will also have a negative polarity.

Basically, they applied the encoding process to decoding, but in an inverse fashion—i.e., division replaces multiplication, and the positions of the speakers replace the positions of the sources. In showing that this worked, Cooper applied a theory of directional hearing called Makita localization.

Developed by a Japanese acoustician of the same name, this theory allows for any number of sound sources, at any azimuthal positions, equidistant from the listener. They all carry the same signal, but with arbitrary phases and amplitudes. When you plug the amplitudes, phases and positions into a formula, an angle pops out. The sound appears to come from this direction when you turn to face it. (This means that the apparent angular position of the source might be slightly different from the actual angular position when you turn toward it. The same effect occurs with live sources.)

Cooper showed that his proposed encoding/decoding system, dubbed UMX, satisfied the Makita localization formula for every position around the listener. He also ran listening tests, which verified this. He displayed the results in three plots (Fig. 6), contrasting UMX with SQ and QS. Cooper refused to identify the other two systems, considering it a breach of professional ethics to do so. Frankly, the paper would have had more impact if he had. After all, he *was* telling the truth.

The importance of Cooper's work —a system that properly encoded directionality and the practical proof that such encoding could be used to create precise directional effects on playback—was ignored. Part of the problem was the opaqueness of the writing. Most technical papers are badly written, and Dr. Cooper's was no exception. The other part of the problem was that many AES members are just not interested in what their co-workers are doing.

"Showcase" piece *(TAA* 4/83, p. 53, *Photo 8)*, the word is used—improperly, I think—to indicate idling collector current.

In the dear old days of vacuum tubes, we used the potential drop across the cathode resistor to bias the tube into the proper operating condition, and the grid was tied to ground through a rather high-value resistor. Frequently, we increased the gain with a "bypass" capacitor. Sometimes we used a fixed negative voltage on the grid as a bias. We did all this to bias the characteristic of the tube into a more linear region of operation.

With transistors coming on the scene, the "Shea bias" was common. This is a resistive divider on the base that induces an idling current, putting the device into more linear operation. In addition, solid-state power amplifiers invariably have some circuit arrangement to "bias" the output devices into class B, AB or even A operation. Without such bias, the output would be near class C, and the crossover notch would be horrible. Now the "bias" can be adjusted to control the collector current of the output—either single ended, complementary push-pull or quasi-complementary.

What all this means is that Mr. Vikan has "biased" his output to run at an idling current of 400mA. In the future, I think we should be careful to use the term properly.

L. B. DALZELL
El Cajon, CA 92020

CONNECTION CORRECTION

I AM WRITING IN RESPONSE to Darin Ernst's letter in *TAA* 5/83 (p. 55). Mr. Ernst and I have had a very fruitful exchange of letters and telephone calls in an effort to solve his problem. He did not use the standard layout, but mounted Old Colony's boards in an elegant enclosure of his own design. We both considered exotic and obscure causes for his problem, but Mr. Ernst finally discovered the real cause himself—bad solder joints and bridges.

Let this be a lesson to other readers. If something does not work, check and recheck the connections and boards, then have someone else check them again. In my experience, if a transistor or IC is bad, it usually got that way as a result of one of my wiring errors.

BERNHARD F. MULLER
Milan, MI 48160

AMBISONICS

Badly written or not, the Cooper/Shiga paper is very important in the history of audio engineering because it was the first step in creating a rationalized, practical system of directional effects. Cooper's association with Shiga was also significant commercially, as Denon later released the system under the name UD-4. This was carrier-disk technology, with a significant difference. Unlike CD-4, the basebands were a matrixing of the four UMX signals, which could be decoded to give credible results by themselves or combined with signals from the carriers for uncompromised performance.

Cooper showed that the carrier signals did not need a bandwidth greater than 2 or 3kHz to provide good directional effects. (CD-4 disks modulated the carriers up to 15kHz.) Since the carrier's sidebands did not have to extend to as high a frequency, a quieter, lower-distortion, longer-wearing record was possible. It is difficult to understand why UD-4 failed and JVC's CD-4 succeeded, when UD-4 had less distortion and more accurate directional effects.

The Kernel
The choice of names also was detrimental to the Denon system. UMX stands for Universal Matrix, while UD-4 stands for Universal Discrete Four-Channel. UMX is neither a matrix nor a discrete system: it is a kernel system. Matrices are mathematical transformations, or mappings, of one set of variables into another. Usually, the transformed set has a smaller dimension than the original set. SQ and QS are examples of this, where four channels are transformed into two. The significant thing about matrix transformations is that they operate on discrete values.

A kernel also transforms variables from one set of dimensions to another, but the transformation is continuous—that is, it is defined as a smoothly varying function, not as a fixed set of discrete coefficients. This distinction is critical. A kernel allows you to discard the idea that each transmission channel must correspond to a speaker-feed signal. Instead, you can implicitly specify directionality for all possible positions, without exactly referring to

---

how the directionality is to be presented.

On the other hand, in both discrete and matrix quadraphonics, the four speakers are treated as four separate sources. Sounds are panned between speaker pairs, and each speaker-feed signal is transmitted in such a way as to preserve its individuality. There is a one-to-one matrix relationship between the original four channels and the final four speaker feeds. In this sense, all discrete systems are actually matrix systems. This is unconsciously acknowledged in the 4-4-4 terminology used to distinguish discrete systems from 4-2-4 matrix systems.

For whatever reasons (probably the lack of commercial response to the UD-4), UMX never developed past this point. The British then took up where Cooper and Shiga left off.

Next time, Mr. Sommerwerck will continue his discussion of the development of Ambisonics.

POWER SUPPLIES

4. Boak, J., "Power Modifications for the ST-150 BJ-1," *TAA* 2/81, p. 40.

5. "Boak's Regulator Queried," *TAA* 1/79, p. 55.

6. Jung, W.G., "Build an Energy Storage Bank," *Audio*, August 1980.

7. Millman, J. and C.C. Halkias, *Integrated Electronics*, McGraw Hill (New York), 1972.

8. Sprague Electric Company *Engineering Bulletins*, Numbers 3431D and 3457B.

9. *Linear/Switchmode Voltage Regulator Handbook*, 2nd Edition, Motorola Inc., 1982.

10. Meyer, D., "Build a 4-Channel Power Amplifier," *Radio-Electronics*, March-April 1973.

11. "Equipment Review: Quad Model 405 Power Amplifier," *Audio*, April 1979.

12. McKen, Blair, "Build a High Quality Phono Preamp," *Audio Scene Canada*, November 1975 and January 1976.

13. Breakall, et al., op cit.

14. Motchenbacher, C.D. and F.C. Fitchen, *Low Noise Electronic Design*, John Wiley & Sons (New York), 1973.

15. Ott, Henry, "Ground: A Path for Current Flow," *EMC Technology*, January–March 1983.

16. Cherry, Edward M., "A New Distortion Mechanism in Class B Amplifiers," *JAES* (Volume 29, Number 5), May 1981.

17. Millman and Halkias, op cit.

18. Fairchild, *1982 Linear Databook*.

19. Siemens, *Analog IC's Databook*, 1981-82.

20. Fink, Donald G., *Electronics Engineers Handbook*, 1st Edition, McGraw Hill (New York), 1975.

21. Jung, W. and R.N. Marsh, "Picking Capacitors, Parts I and II," *Audio*, February and March 1980.

Part II

AMBISONICS COMES OF AGE

BY WILLIAM SOMMERWERCK

In TAA 3/84, Mr. Sommerwerck provided some background on Ambisonics. He will continue that discussion this time and give more details about the Ambisonic approach.

MICHAEL GERZON IS a mathematician at Oxford. In the early 1970s, with all the talk of quadraphonics going around, he realized that something was amiss. He saw the need for additional channels to convey the spatial character of live sound more accurately, but he also saw that the pair-wise mixing (PWM) techniques being promoted were incorrect. He was also disappointed with the poor performance of simple matrix decoders and thought that much more credible effects must be possible for listeners who had access only to two-channel program sources. With this in mind, he started theoretical and empirical research into the fundamental problems involved in reproducing directionality correctly. He was joined in this work by Peter Fellgett of the University of Reading. Their research was partially supported by the National Research Development Corporation and IMF Electronics.

Over a ten-year period, Gerzon and Fellgett produced a system that allows recording engineers to achieve almost any directional effect, with total confidence that the listener will hear it correctly. Effects can range from completely natural reproduction of live sound and its ambience to totally artificial studio effects, and these may be freely combined. The system is called Ambisonics.

Its starting point is the same as that for UMX. The sound field is analyzed into four components, one of which is an omni signal without directional information. Three figure-8 components represent front/back, left/right and up/down direc-

tionality. These are called the W, X, Y and Z signals, respectively. See *Fig.* 7. (Note that the fourth Ambisonic channel is *not* the "sin 2 theta" used in UMX. Gerzon showed that this signal actually degrades performance. Its presence gives the brain additional cues, which cause sounds *near* a given speaker to localize *at* that speaker. This is undesirable, since the goal is a smooth, continuous spread of sounds, without any loudspeaker awareness. When exag-
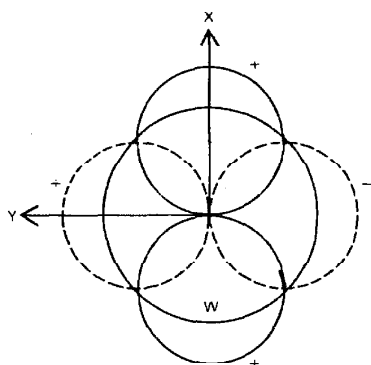
Reprinted with the permission of NRDC.

FIGURE 7: B-format signal arrangement. In this case, the Z signal points up.

gerated directionality is desired for dramatic emphasis, you can use the fourth channel for just that effect.)

The W, X, Y and Z signals can be assembled from live sounds by means of special microphone arrays or in the studio by simply adjusting the amplitude and polarity of the signals, as described earlier. Since the up/down channel is optional, only three channels are needed for normal surround sound.

Besides the formation of a viable

surround-sound system, Gerzon and Fellgett studied and applied existing mathematical models of directional hearing. These show that Ambisonic reproduction is correct, and they permit rational trade-offs among conflicting requirements in designing practical systems. Naturally, the researchers performed hundreds of hours of listening tests, but these were guided by mathematics and psychoacoustics. Little is "cut-and-try" in Ambisonic technology.

I have already discussed Makita localization, which is the most important directional cue. The success of any surround-sound system hangs on its proper implementation. Any system that does not satisfy Makita is inherently flawed and cannot accurately reproduce ambience or arbitrary effects. But other psychoacoustic laws must be obeyed, too.

One of these is described by the sound field's "velocity vector magnitude," which determines how stable the field is with respect to head rotation for frequencies below about 1kHz. Changes in this parameter can make the sound move toward, or even into, the listener's head. (Because PWM fails to satisfy this rule, it cannot create stable side images.)

Another rule is described by the "energy vector azimuth," which determines the apparent image location at high frequencies (about 0.5 to 5kHz). (Oddly, PWM *does* work in this range.) Just as the velocity vector magnitude defines image stability for Makita localization, a fourth parameter, called "energy vector magnitude," defines image stability for the energy vector azimuth.

And then there is "phasiness." One example of phasiness occurs when your speakers are connected out of phase. There is no clear center image, and other sounds are bloated and hard to localize. Another symptom is an uncomfortable "pressure-

in-the-ears" effect, which might range from a slight blurring of image location to the sense that fingers are pushing against your eardrums. Phasiness is related to the amount of 90-degree (quadrature) component in the signal. Only small amounts are acceptable, and the effect appears most pronounced in the 300 to 1,000Hz range.

These effects are described in a group of equations that you can use to analyze any sound-reproduction system, regardless of the number of transmission channels or speakers. I decided that I could not be an Ambisonic "maven" without going through the "maths," so I applied them to the three-channel (W, X, Y) version. They all came out on the button. (The equations, along with an explanation of how to apply them, appear in the sidebar on page 41.) When you analyze the equations, you will see that they account for amplitude and phase differences, but they do not mention arrival time differences. These differences are a significant source of directional information for the ear and brain, especially for the initial transient of any waveform. Why have they been left out of Ambisonics?

You Can't Have Everything

The two reasons for this are simplicity and compatibility. The ear and brain use at least three mechanisms to localize sound—amplitude, phase and arrival time differences. Although the arrival differences *do* make a major contribution, if amplitude and phase are well implemented, the need for delay is less important.

Adding those delays introduces serious problems. You must either use a *kunstkopf* (dummy head—literally, "art head") for live recordings or introduce actual delay for studio recording, via (expensive) delay lines. If you do not want to use headphones in playback, you must supply a crosstalk canceller for loudspeaker listening. This limits the useful "window" to a space suitable for only two people—if they are Siamese twins! When you mix down to mono for broadcasting, the delayed components will "comb" with the undelayed to create frequency-response aberrations. (This might not be significant to those hearing the program over cheap mono players or receivers, but it *is* a potential problem.) The

worst part is that you must either wear headphones or install a crosstalk canceller. You cannot sit down and listen to your system in its conventional state.

JVC has combined all the principles of directional hearing, including interaural delay, into a surround system called Q-Biphonics. It works beautifully, but its complexity and
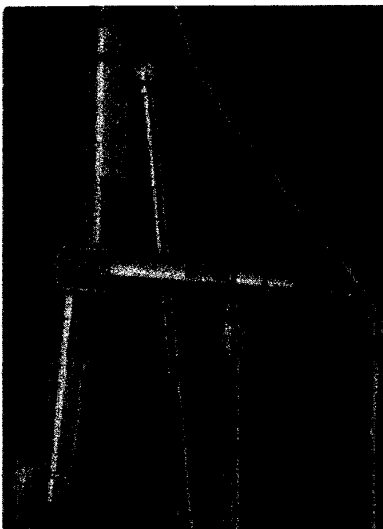
FIGURE 8: The author's home-brewed Ambisonic recording setup, using variable-pattern SoundField microphones.

incompatibility with existing systems and listening habits will, I believe, prevent it from ever becoming widely accepted. That leaves us with Ambisonics.

The Ambisonic Ideal

After all this build-up, just how good is Ambisonics? The three-channel version (horizontal only, no height) is amazing. The sound field has a remarkable coherence that you rarely, if ever, hear in conventional quadraphonic recordings. The ambience and the direct sounds hang together in a way that is difficult to describe, except to say that it sounds almost exactly like the hall in which the recording was made. This is in contrast to quadraphonics, where the ambience is enhanced, but the orchestra sounds too far away, and the reverberation is overdone. Perhaps an example will clarify this point.

I did some of my earliest quad recordings with the Orchestra Society of Philadelphia, an "amateur" group. (A lot of "professional" orchestras should play this well.) I set

up four cardioid mikes in prescribed "purist" fashion, arranging them as nearly coincident as possible, with two of them facing the rear. But the recordings still sounded like conventional stereo. Although the mikes were 6 feet behind the conductor, the recording sounded as though I was sitting in row M. And although there was a pleasant increase in ambience, it just did not relate to the direct sounds. If I raised the rear channels to the point where they made a noticeable contribution, they drew attention to themselves. If I reduced their level, the ambience disappeared.

Something was wrong, and now it is obvious what it was. As I explained before, all the side sounds (the significant ambience) were being directionally encoded by amplitude panning between the front and rear mikes—acoustical PWM, if you like. No wonder the ambience was displayed incorrectly in playback!

Legal Home-Brew

After getting the Integrex decoder, I tried a home-brewed SoundField setup (Fig. 8). (The SoundField mike is manufactured by Calrec and distributed by Audio + Design, Calrec, E4480 Hwy. 302, Belfair, WA 98528.) Fortunately, my mikes are variable-pattern, so all I had to do was rig up the necessary stand/boom arrangement to bring an omni and two crossed figure-8s into close proximity.

The recordings were a revelation. Now it seemed as though I was about 10 feet behind the conductor, the reverberation was no long excessive (in fact, the playback was a bit drier than the hall), and everything hung together beautifully. Furthermore, these qualities were audible over a wide listening area. Moving close to one speaker did *not* cause the sound field to collapse in that direction. And although I could not hear specific sounds from the sides, they no longer seemed empty. The sound of the hall at Drexel University filled the listening room, without drawing any attention to itself, just as it would in real life.

Perhaps the most remarkable and unexpected improvement was the significant reduction in coloration. It was similar to going from Bextrene to polypropylene drivers. A layer of excessively liquid sweetness was stripped away. Woodwinds were

noticeably "reedier," violins started sounding more like rosin on gut, and brass instruments had more snap and attack.

This effect seems related to the more correct presentation of ambience that Ambisonics provides. As I have said, the most important ambience is the side delays. In regular stereo recording, these are folded into the direct sound. The brain is incapable of separating them under these conditions, and they "comb" with the direct sound to alter frequency response (and, therefore, instrumental timbre). In Ambisonics, the directional cues necessary for proper localization of the side delays are present, and the combing does not occur.

I made another interesting recording at a neighborhood church's choir rehearsal. The choir was divided, the sections facing each other across the loft. The organ pipes were behind one group, so it was quite natural to have them and one choir to the front, with the other choir "behind" the mike array. The result was marvelous. I could close my eyes and hear the acoustics of the church. (Again, the recording was a bit drier than the real thing.) The speakers disappeared: the recording produced no sense of four sound sources, just a coherent, continuous sound field, no matter which way I turned. It was especially interesting to turn around and pick out individual voices in the rear choir.

Oddly, the organ appeared to be above the front choir, as it was in the church, but I was not using a fourth channel for height, so I chalked this up to autosuggestion. (A British recording engineer told me that although height information is not explicitly encoded, it might be there implicitly. The brain might judge the height of a sound source by noting the arrival time difference between direct sounds and those bouncing off the floor or stage. Perhaps this is so, but I do not see how the brain can distinguish between a reflection from the floor and one from the side wall.)

Of course, it will be some time before three and four-channel Ambisonic recordings become available. For better or worse, they will have to await true digital recording techniques. In the meantime, Fellgett, Gerzon and others have developed a two-channel system, UHJ. ("U" for
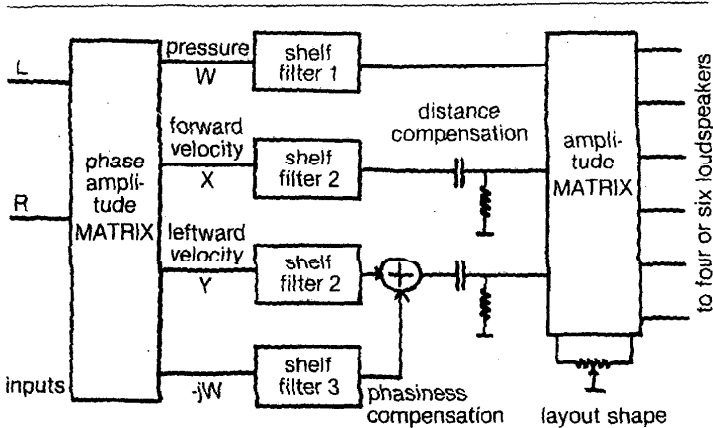
Reprinted with the permission of IMF Electronics.

FIGURE 9: Overall UHJ decoder circuitry. In a VDP decoder, the −jW signal would be dynamically varied by logic circuitry to control the direction of lowest phasiness.

Universal, "H" for Matrix H (hearing), a BBC system whose patents are now part of the Ambisonic pool, and "J" for 45), the original two-channel Ambisonic system. It is sometimes called BHJ to acknowledge the Cooper-Shiga BMX system, which is also part of this pool.) This system audibly outperforms any four-channel discrete system.

Two Is Greater Than Four

I have already asked you to swallow quite a bit by saying that a three-channel system is superior to any discrete technique. A careful examination of the psychoacoustics involved and a brief listen will confirm this. But to say that a mere two channels can provide exceptional performance really strains credulity. After all, simple two-channel matrices were awful, having to await such developments as the QS Vario-Matrix and the Tate Directional Enhancement System for SQ. Certainly, UHJ must need some kind of logic enhancement? No way! It beats SQ and QS (not to mention discrete) at their own game with simple, linear circuitry (Fig. 9).

Before I explain how this is possible, I ought to describe what you can expect to hear from UHJ playback. It is so good that many listeners will not be able to tell it from the full three-channel, B-format system. The sound field is equally coherent. You have no sense of four distinct sound

sources, but instead perceive a continuous field. Moving close to one speaker does not degrade the image as badly as it does in quad or regular stereo. In addition, it is possible to turn sideways and hear the front sounds to one side, the rear to the other. As you sidle up to the front, the front sounds predominate, and vice-versa. The listening window is wider and deeper than that for regular stereo, even extending outside the speaker boundaries.

These effects are not possible with discrete quadraphonics, let alone matrix systems. A minute or so of listening will sell even the most jaded audiophile on two-channel Ambisonic sound. Why, then, outside of up/down effects, do we even consider three-channel Ambisonic systems? The reason has to do with the inevitable trade-offs involved in reducing the signal from three to two channels.

The W, X and Y signals are mixed to produce a compatible stereo signal. (The stereo is actually slightly better than average, since sounds at the far left or right will appear outside the speakers.) UHJ is a "kernel" system, so the encoding for any position is exactly specified. After dematrixing (to produce approximations to the original W, X and Y signals), the signal must again go through the speaker-feed matrix. The amazing thing is that it is possible to arrange this processing so that the exact

Makita localization is retained for *all* directions. This is why UHJ has such remarkably good performance.

The trade-off is a significant increase in phasiness. Sounds take on a slightly bloated quality and come forward from the speakers. A subjective increase in coloration, caused by the quadrature phasiness components' altering of the frequency response, is also noticeable. Fortunately, there is a solution for this.

The ear and brain are much more sensitive to phasiness to the front than to the rear. You can adjust the decoding so that there is virtually no phasiness in the front, at the expense of high phasiness to the rear. For a forward-facing listener, however, the effect is one of a large improvement at the front, with little degradation to the rear. (Of course, if you turn around, you can hear the phasiness.) It is possible to dynamically alter the phasiness reduction so that the loudest sound is the least phasey. A special decoder using this Variable Directional Preference (VDP) technique might be available soon, but such a system enhances full surround recordings the most. VDP produces little improvement with ambience-only recordings, and none is really needed.

Optimized Two-Channel Encoding

You can see the advantages of UHJ encoding by manipulating the encode/decode equations, plugging the results into the speaker-feed equations, and relating all this to the psychoacoustic localization formulas mentioned earlier. But there are other ways of explaining what happens in a two-channel surround system. The Scheiber Sphere is one option.

In any two-channel surround-sound system, we are interested in three parameters—the intended direction of the sound source, the relative amplitude of the signals in the two channels and the relative phase of the channels. (The last two parameters, of course, represent the way we encode the first.) The usual way of representing the relationships of these parameters graphically is by plotting the encoding locus on the surface of a sphere. In the US, it is

---

The Psychoacoustics of Directional Hearing

The following approach considers only sounds originating in the horizontal plane, although you may extend the methods to periphonic (with-height) systems. Consider the X and Y axes as pointing forward and to the left, respectively. Assume that N loudspeakers are situated on a circle in the azimuthal directions phi-sub-i (i = 1,2,...N) measured counterclockwise from the X axis (due front). For simplicity, assume that all sources lie at a long distance from the center so that sounds arrive as plane waves. If a given mono sound is fed to all speakers, with the complex gain phi-sub-i to the i'th speaker, the following parameters influence the localization of that sound.

1. *Makita's localization* describes the localization azimuth for listeners facing the apparent sound source at low frequencies (less than 1kHz). Calculate x and y as follows:

$$x = \text{Re}\left[\frac{(\Sigma P_i)(\cos\phi_i)}{\Sigma P_i}\right]$$

$$y = \text{Re}\left[\frac{(\Sigma P_i)(\sin\phi_i)}{\Sigma P_i}\right]$$

where Re means "real part of," and the sums are over i = 1 to N.

Makita's localization azimuth (θ_v) is given by the following equation:

$$x = (r_v)(\cos\theta_v)$$
$$y = (r_v)(\sin\theta_v)$$

where the velocity vector magnitude

(r_v) is greater than zero.

2. The *velocity vector magnitude* (r_v) equals

$$\sqrt{x^2 + y^2}$$

where x and y are as defined above. It describes the stability of sound localization with head rotation at low frequencies. For "natural" sounds, r_v equals 1. If r_v is much greater than 1 for a reproduced sound, you can hear an out-of-phase effect. If r_v is close to zero, you hear an "in-the-head" or "close-to-the-head" effect, along with excessive image movement when you move your head. (This effect is under complete control in Ambisonics, allowing the producer to move the sound toward or away from the listener.)

3. *Phasiness* (q) most often affects forward-facing listeners. It is an unpleasant "pressure-on-the-ears" sensation and might include image blurring. Its value is given by the following formula:

$$q = \text{Im}\left[\frac{(\Sigma P_i)(\sin\phi_i)}{\Sigma P_i}\right]$$

where for real u, v, Im(u + iv) means v. Ideally, for natural sounds, q should equal zero, but |q| less than 0.21 is relatively innocuous, |q| less than 0.5 is generally tolerable, and |q| greater than 1 is unacceptable.

4. The *energy vector azimuth* describes the image azimuth at high frequencies (500Hz to 5kHz) and also appears at low frequencies for slightly off-center listeners. Calcu-

late it as follows:

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

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

where

$$x_E = (r_E)(\cos\theta_E)$$
$$y_E = (r_E)(\sin\theta_E)$$

and r_E is the energy vector magnitude.
This describes the apparent localization, especially when the listener faces the apparent direction.

5. The *energy vector magnitude* (r_E) equals

$$\sqrt{x_E^2 + y_E^2}$$

where x and y equal the values defined above. This quantity describes the stability of the sound image with head movement, especially at frequencies between about 500 and 5,000Hz. Ideally for "natural" sounds, r_E equals 1, and for reproduced sounds r_E can never exceed 1. In practice, an r_E of 0.7 is excellent, 0.5 is quite acceptable, and 0.35 is tolerable for sounds to the rear of the listener.

When the encode/decode equations for B-format Ambisonics are used to derive speaker-feed signals and these signals are plugged into the directional equations given above, all the equations are solved *exactly* for all directions. □
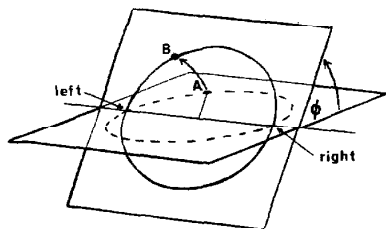
FIGURE 10a: Representation of a stereo position on a sphere. Point A in the horizontal plane represents a sound recorded with a gain of α on L and β on R. Point B represents a sound recorded with α (phase shifted ϕ) on L and with β on R.

called the Scheiber Sphere, after the first person to describe the concept. The British, in dogged refusal to acknowledge Peter Scheiber's contributions to surround sound, insist on calling it the Energy Sphere or a Stokes-Poincaire plot.

The plotting goes like this. The azimuth (that is, the angular position in the horizontal plane) equals [2 × arctan(L/R)], where L and R are the amplitude (including signal polarity) of the left and right transmission channels. Zero degrees is considered to be due right on the sphere. A signal on only the right channel will appear due right. Signals of equal strength on both channels will point forward (90 degrees), while signals of equal strength but opposite phase will appear at −90 degrees. A left-only signal will appear due left at 180 degrees. If this is not clear, work through the trig on your calculator.

Phase differences between the channels are shown by angles of elevation and depression. L is considered to "lead" R, so if L's phase is ahead of R's, the point is plotted above the "equator," and vice-versa. Angles vary from 0 degrees (at the equator) to ± 90 degrees at the poles. See Figs. 10–12.

As I mentioned before, do not confuse polarity inversion with phase shift. If one of the encoded channels is of opposite polarity with respect to the other, then its position along the locus will be on the equator, not above it, since the calculation of azimuth includes signal polarity. [For
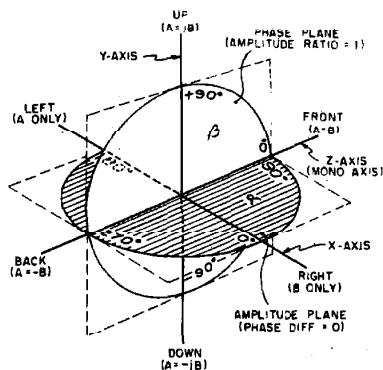
FIGURE 10b: Spherical α,β (amplitude ratio, phase difference) coordinates showing amplitude and phase planes. The positive Z axis is designated the mono axis because the level with which any matrixed channel of information is reproduced in monaural playback depends on the proximity of its α,β coordinates to those of the mono axis (90.0).
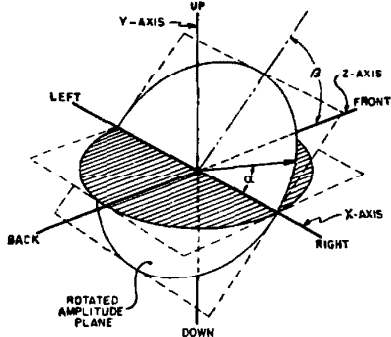
FIGURE 10c: The plane of the amplitude ratio angle, rotated by the plane difference parameter β.

the arctan of an angle to fall between 90 and 270 degrees, either the sine or cosine of the angle (but not both) must be negative.]

Most people immediately grasp these concepts, then get bogged down in a very common and understandable mistake. They think that the azimuth of the locus necessarily corresponds to the intended direction of the sound source. It may, but it does not have to. (As we will see later, RM shows exact correspon-
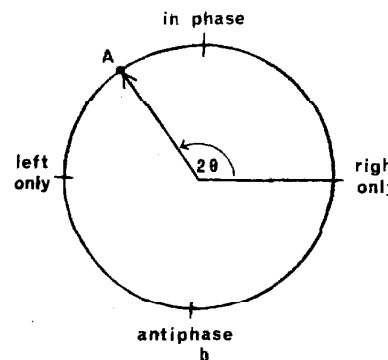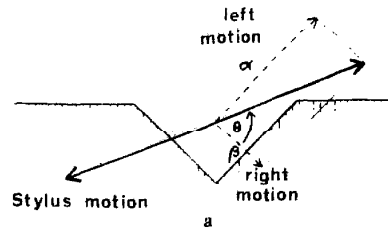
FIGURE 11: Point A at an angle 2θ from the right point of b represents a direction of stylus motion at an angle θ from the right wall's direction of motion as in a.
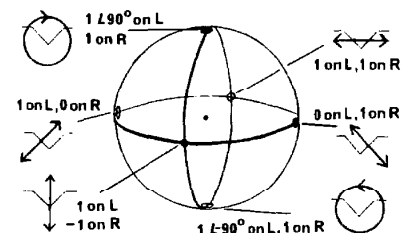
FIGURE 12: Stereo positions and stylus motions corresponding to various points on the energy sphere.

dence, whereas UHJ does not.) To avoid confusion, the intended direction should be marked on the locus. A simple example will illustrate this point.

Forget about surround sound for a moment, and suppose that we have plotted regular stereo on the sphere. A source that is on only the left channel will be plotted due left on the sphere (180 degrees). Our intuition suggests that this represents due left (or center-left). But in the listening
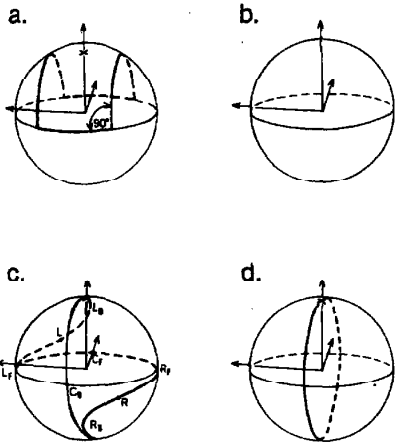
a.　　b.

c.　　d.

FIGURE 13a: Scheiber Sphere plots of some representative horizontal pan-loci: a is the Sansui QS, b is the RM, c is the CBS SQ, and d is a great-circle locus consistent with Ambisonic encoding. In sphere c, note the cusps and left-right asymmetry due to the choice of front-sector mapping and the limitations of encoding from four pair-wise blended channels. CF is the center front, CB is the center back, L is full left, R is full right, and LF is left front.
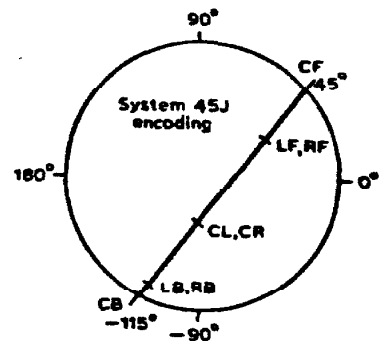
FIGURE 13b: Scheiber Sphere plot of two-channel UHJ encoding.

room, this signal will appear at the left speaker, which is usually about 30 degrees left of center-front, *not* due left.

Let's examine the Scheiber Sphere plots for RM, QS, SQ and UHJ (*Figs. 13a* and *13b*). One thing not immediately obvious is that RM and UHJ are kernel systems. Their system-encoding specifications describe *exactly* how each direction is to be encoded. On the other hand, QS and SQ are matrices, which are specified for only the four corner positions. The intermediate positions shown have been computed by assuming PWM between adjacent channels. When Sansui says that QS is "derived from" RM, it is true only in that the four corner positions have the same encoding in each system. QS assumes, like all other matrix/discrete systems, that signals are to be pan-potted between the channels.

It should be obvious that SQ and QS have serious problems. The worst of these is that the encoding loci are not smooth. As you can see, both have distinct cusps, with SQ the worse of the two. This means that points that are relatively far apart in "real" space will be relatively close together in "encoding" space. That's a no-no! Both SQ and QS depend on sophisticated "logic" decoders for acceptable performance. These en-

FIGURE 14: This circle shows the optimized nonsymmetric distribution of different encoded directions within the circle "pan-locus." CF is center front, RB is right back, CL is center left, and so on. This is not a picture of the energy sphere.
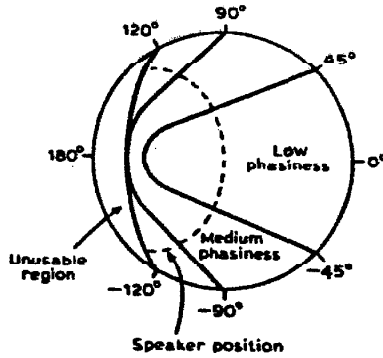
Speaker position

FIGURE 15a: The quality of mono and stereo reproduction shown on the energy sphere as viewed from the right side. The speaker position curve indicates that the signal appears to come from only one speaker in stereo.
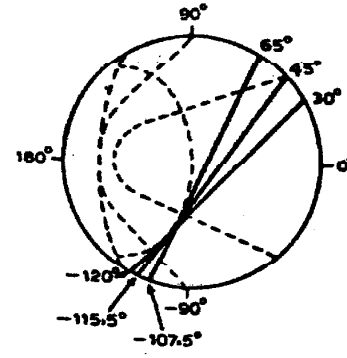
FIGURE 15b: Three possible choices of the two-channel encoding system having optimized mono and stereo reproduction.

coding cusps make it more difficult for the logic circuitry to determine the intended directionality. When UHJ decoders use logic circuitry to reduce phasiness, this kind of problem cannot occur because UHJ has a smooth locus.

Another problem with SQ is its lack of symmetry. Some abstruse mathematics in Peter Fellgett's article "Surround Sound '76" (Hi-Fi Sound Annual 1976) shows that a nonsymmetrical matrix cannot be decoded (that is, manipulated to produce speaker-feed signals) in such a way that all the psychoacoustic criteria governing low-frequency directionality are satisfied simultaneously. Before the QS supporters start snickering, it turns out that the side signals in QS cannot be decoded correctly either. Although both systems are bad, SQ is worse than QS.

Both RM and UHJ allow correct decoding for all signals. Why doesn't UHJ use the same encoding as RM? Why does the locus have that odd tilt to it? And why are the sources spread out all over the place? Let's look at the last question first.

The ear and brain do a much better job of localizing sounds to the front than to the rear. Logic suggests we arrange the encoding so that frontal sounds take up as much of the encoding space as possible. This ensures that any transmission or decoding errors will have minimal effect on the perceived positions of frontal sounds. Of course, the rear sounds

then get the worst of it, but it does not matter because the brain does not localize them as well anyway. This is the philosophy of UHJ, where the front half-circle of "real" space gets 240 degrees of the encoding locus. RM is fully symmetrical, with front and back taking up equal encoding space (Fig. 14).

Now for the tilt. It was introduced to improve mono and stereo compatibility and to reduce the phasiness of rear sources as heard by stereo listeners. Let's examine compatibility first (Figs. 15a and 15b).

Compatibility problems are not new: they arrived with stereo. The most common is the build-up of the center signal when a stereo recording is played in mono. The channels add, giving a 6dB boost to a central performer. Sounds not quite center are increased proportionately less, while sources to the far left or right get no increase at all. The soloist becomes unduly prominent, a problem with all types of music. When matrix quad came in, a new incompatibility appeared. Sounds at or near center-back were recorded anti-phase (QS, RM, SQ), which caused them to cancel or be severely attenuated in mono. (Stereo listeners also perceive phasiness and lack of localization.)

Therefore, the word went out: thou shalt not encode center-rear signals in matrix quad. But why should the producer be prohibited from placing the sounds to the rear so that mono listeners will not miss

anything, when the raison d'etre of surround sound is the ability to place sounds anywhere around the listener?

One goal of Ambisonics was to allow any type of recording to be produced, without favoring one technique over another. This means that a record producer should be free to place the performers anywhere around the listener, without losing anything in mono or producing irritating aural quirks in stereo. That's what compatibility is all about.

Ambisonics solves the compatibility problem (as well as reducing phasiness for rear signals) by the simple expedient of tilting the encoding locus. To see how this works, look at Fig. 15a. It shows the Scheiber Sphere from the right side. The encoding space has been marked off to show the degree of phasiness perceived by stereo listeners. As sources move closer to center-rear, the encoded signals are less and less in phase with each other, especially for encoding loci that are strictly on the equator, as in RM. Sounds encoded close to center-rear but not dead-on approach full inversion less closely and show correspondingly less phasiness. Figure 15a reflects phasiness caused by polarity differences, as well as interchannel phase shift. Notice how, as the angle of elevation (i.e., interchannel phase shift) increases, the area of unacceptable phasiness moves closer to the front of the sphere.

Lowering and tilting the locus gives us what we need. (See *Fig. 15b* for the variety of possible center-front phase shifts.) All rear sources are now moved out of the positions of unacceptable phasiness. This comes at the expense of *increased* phasiness for most frontal signals, but, paradoxically, this is desirable! For one thing, the extra phase shift causes sounds encoded at the sides to move out beyond the loudspeakers.

We also get the desired enhancement of compatibility. Center-front sounds are now 45 degrees out of phase, limiting their mono build-up to only 3dB. (The maximum amount of stereo phase shift tolerable to critical listeners is 45 degrees.) With the shift reduced to 115 degrees or less, center-rear signals are no longer anti-phase, giving a maximum mono attenuation of less than 2dB.

Going Around in Circles

(You will notice that the encoding locus for UHJ is a circle. Is this optimum? Yes. We have seen that a smooth locus, free of kinks and cusps, is desirable, and a circle meets this requirement perfectly. Furthermore, the circle should be a "great" cir-

cle—that is, its center should pass through the center of a sphere. A great circle has the largest possible circumference of any circle that can be drawn on a sphere, so we are assured of using the maximum amount of encoding space. The UHJ locus sits a bit below center, so it is not *quite* a great circle. This allows the encoding locus to touch the locus of apparent single-speaker sources, which is required to ensure a good spread of sound between the speakers.

Why is a great circle optimum? Why not use an elongated, twining locus, like the stitching on a baseball? Wouldn't that use much more of the available encoding space? Yes, it would, but it would be worse in other respects. Notice that walking around the sound field is the same thing as traversing the encoding locus. If you follow this on your baseball, you will see that encoding points that are well separated along the length of the locus are nonetheless relatively close together on the surface of the sphere (i.e., close in encoding space). A "serpentine" locus would, therefore, defeat the reason for enlarging the encoding space.

One other point is worth address-

ing. Some of you might be wondering whether Ambisonics is "holographic." (Its supporters sometimes describe it as "quasi-holographic" because its performance is so remarkable.) An optical holograph presents to the eye essentially the same waveform it would receive if it were actually looking at the object. For a comparable sonic result, it would be necessary to duplicate the sound field over a very large listening area. Currently, this is impossible, so Ambisonics takes a different tack.

As we have seen, the most important parts of the sound field for determining directionality are the zero and first-order components. Ambisonics extracts these and presents them to the listener in a psychoacoustically optimum way. This produces accurate and stable imaging over a wide listening area. It is holographic only in the sense that it properly reproduces those parts of the sound field that are required for good localization. No other so-called system does this.

Next time, Mr. Sommerwerck will give construction and recording details for the Ambisonic system.

AMBISONICS COMES OF AGE

BY WILLIAM SOMMERWERCK

In the first two parts of this series, the author explained the background and logistics of Ambisonic technology. This time, he concludes with a description of his home-brewed system and offers construction and recording tips.

PARDON THE HYPERBOLE, but Ambisonics is the most important development in sound recording and reproduction since the advent of electrical recording—and for some reasons you might not have considered.

The importance of electrical recording lies not so much in its elimination of the gross mechanical limitations and colorations of acoustical recording, but rather in its development methodology, which rationalized the audio design process. Maxfield and Harrison, the Bell Labs engineers who invented the electrical recording process, did not "tinker" it together. A rigorous mathematical analysis of the mechanical and electrical aspects of the design guided them. Since then, science has had some influence on the art of sound reproduction.

Likewise, the importance of Ambisonics lies in its rationalization of recording technique. The engineer can now produce almost any arbitrary effect imaginable or create a nearly perfect reproduction of a recording site's ambience. The mind boggles (and the ears twitch) to think what John Culshaw would have done with his classic production of Wagner's "Ring" if Ambisonic technology had been available in 1958.

Constructing the System

There is no practical way to duplicate the SoundField system. The micro-
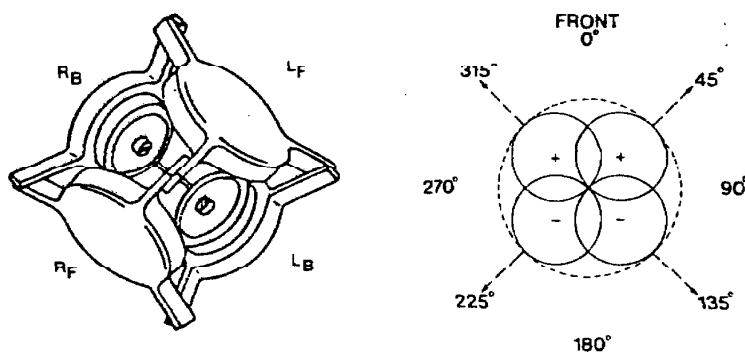
FIGURE 16: The SoundField microphone head contains four tiny hypercardioid capsules positioned as if they were on the surfaces of a regular tetrahedron.

phone head contains four tiny hypercardioid capsules *(Fig. 16)*, positioned as if they were on the surfaces of a regular tetrahedron. [The *exact* pattern does not matter, as long as it is a reasonable mixture of zero-order (omni) and first-order (figure-8) components. The common cardioid has these components in equal amounts; the hypercardioid has more figure-8s than omnis and thus shows a slight anti-phase back lobe. I assume Calrec chose this capsule because it had the smoothest response and/or the most consistent pattern from sample to sample.] The capsules' outputs are mixed to produce an omni (W) and three figure-8 patterns (X, Y, Z).

For this process (akin to solving four simultaneous equations for four unknowns) to work well, the capsules must be matched in amplitude and phase response, as well as have

identical patterns. (Hypercardioid is made of omni and figure-8 mixed, so a pattern change represents a change in the relative level of these components, upsetting the process of extracting the correct W, X, Y and Z signals.) If you have the time, money and anechoic chamber required to hand-select the elements, you probably also have the $8,000 needed to buy a SoundField mike.

Clearly, the only practical approach for poverty-stricken amateur recordists is to generate the three (or four) signals of the B format directly. This means using an omni and two (or three) figure-8s pointed in the right directions. All you have to do is crank up the old four-channel recorder, and away you go.

There is a little more to it than that, though. First, you cannot use just any mikes. Not only do mikes

from different companies differ in amplitude and phase response, but they also differ in overall coloration. (Imagine how a recording would sound with a Neuman on one channel and an AKG on the other.) This suggests that you should use variable-pattern mikes from the same company.

I was lucky. When I bought my mikes, I realized that no single pattern could do justice to all types of recording situations, so I bought PML (Pearl Microphone Labs) TC-4v variable-pattern mikes. As you can see from *Fig. 8* in Part II of the series, it was not too hard to find an arrangement that would hold them in the right position. The omni (W) is on the bottom, the left/right (Y) is in the middle, and the front/back (X) is on top. I thought that this would produce the least acoustical interaction among the mikes. (The first Ambisonic recordings were made with four cardioid mikes in a tetrahedral arrangement. This was quickly abandoned when listening showed just how much coloration acoustical interference introduced. In the setups I will discuss, the mikes are smaller and are not so intimately intertwined.) The major problem lies in tightening up everything so that the mikes do not shift or flop over.

Using variable-pattern mikes is not the only possible approach. Another, which I jokingly call "gluing Schoeps capsules together," is also available. Schoeps' "Colette" series uses interchangeable capsules built around the same fabrication technology. This means that, frequency response differences aside, they will have similar colorations. (Schoeps would say "lack of coloration," of course.) This makes it possible to combine them into a coherent-sounding array. You can separate the capsules from their electronics in the mike body with a special cable, allowing you to place the body in any position on the stand, with the capsules "upstairs." This reduces acoustical interaction. To glue the elements together, you can use cyanoacrylate glue, which hardens in a few seconds and can be dissolved with a bit of acetone.

What is the best approach for you? Variable-pattern mikes are pretty much out of the question. They are generally expensive and so bulky

that it is hard to position them well. (I bought mine when the dollar was worth more and the kroner less. As you can see from *Fig. 8* (Part II), they are reasonably thin in comparison with with Neuman or AKG mikes.) The Schoeps mikes are more practical, and they form the heart of a complete miking system. An omni and two figure-8s, plus power supplies, will cost about $2,000. (For more information about Schoeps mikes, contact Posthorn Recordings, 142 W. 26th St., New York, NY 10001.)

Schoeps is not the only name brand system you can use. AKG's 450 series of microphone preamps and capsules includes the omni and figure-8s needed for a B-format pickup. Prices are comparable to the Schoeps mikes, but the AKGs include no provision for extending the capsule from the preamp. (Contact AKG Acoustics at 77 Selleck St., Stamford, CT 06902.)

At a lower price, I believe that Beyer Dynamic (5-05 Burns Ave.,

Hicksville, NY 11801) makes a series of omni, cardioid and figure-8 mikes, all using the same ribbon element. I have not been able to find any information on them, however. Can anyone help?

Perverse Suggestion

Of course, although these approaches are a lot less expensive than buying the SoundField microphone, they are still quite costly. Even if you already own AKG or Schoeps mikes, you are probably using cardioids and will have to buy another preamp and three more capsules. Never one to be intimidated, however, I came up with a brilliant alternative that is rather—uh—perverse.

How about using $3 condenser mike capsules from Radio Shack? It should be possible to put together a credible SoundField mike for less than $20. Doing so is actually rather simple. Suppose you placed two cardioid elements back-to-back—i.e., very close to each other and 180 degrees apart. Now, think of a cardioid

---

Ambisonic Recordings

No B-format recordings are currently available. The following are all UHJ disks and cassette tapes. Nimbus has already released UHJ recordings on compact disk, and Unicorn will if it has not already done so. This list is not comprehensive, but is simply an indication of the kinds of recordings available.

Nimbus was, perhaps, the first firm to use the SoundField mike and has the most UHJ disks available. Most of its UHJ disks are of solo keyboard and chamber music, ranging from Baroque to Late Romantic. The high point is a complete Beethoven sonata cycle (and the Diabelli Variations), all on direct disk. The artist is Bernard Roberts, and if the few pieces I have heard are a true indication, his is one of the great interpretations. Also available are some vocals and nonmusical recordings.

On the other hand, *Unicorn* probably has the greatest variety of offerings. These range from movie music (the sound track for *North-by-Northwest* is the best-known example) to what appears to be becoming a complete traversal of all of Messiaen's organ works. Also available

are chamber music, some of Panufnik's orchestral music, a fair amount of Delius, band and college-orchestra music, and some of Peter Maxwell Davies' more recent works.
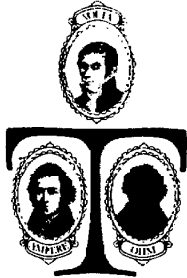
All of *Libra Realsound's* recordings are on cassette. They include unusual works, such as pieces by Loeillet, Neopolitan madrigals and 18th-century English organ music.

IMF Electronics has concentrated on "popular" music, recording live performances of groups such as The New Reformation Dixieland Band, The Reef Petroleum Big Band and Loose Caboose. Also available is a disk of the Albion Ensemble, a wind chamber group.

Music From York recordings are released on disk and cassette and are principally of church and choral music. The most infamous of the latter is "The Teddy Bears' Picnic," an album of children's songs.

Several other firms offer one or two UHJ recordings. In all, at least 120 UHJ recordings are available, a surprising number when you consider that the system has no support from any major record company. □

as a figure-8 added to an equal-strength omni. With the mikes facing in opposite directions, the figure-8 components are reversed in polarity with respect to each other. (Of course, the omni components have no directionality, so their phasing is meaningless in this context.) If the outputs of the two mikes are added, then the figure-8 components will cancel and the omni components add. If they are subtracted, the figure-8 components will be retained at the expense of the omni.

There are two ways to get these sum and difference signals. The first is simply to wire the capsule outputs together, in series. With the hot of one feeding the ground of the other, they will be in phase, and the output will be omni. Connected hot-to-hot, they will be anti-phase, and the output will be figure-8.

This application requires six capsules: four to create the X and Y signals and another two to make the omni. (Yes, you could use an omni capsule, but remember that you want to match the phase, frequency response and overall colorations.) The capsules would be glued together like that old toy "Lincoln Logs."

With a few op amps, you can reduce the mike count to four. You can add or subtract one pair (or even all four) to produce the omni and one figure-8, with the remaining figure-8 created by the other two mikes. The use of op amps would also allow you to adjust gains to compensate for differences in capsule sensitivity. The later addition of a more complex mixing circuit would allow independent selection of the mike pattern and angular separation for regular stereo recording.

Even when using op amps to adjust level, it will still be necessary to match the capsules for pickup pattern and amplitude/phase response. This is done by summing the outputs of the two capsules under test, out of phase. Using a speaker as a sound source, adjust the capsule levels to get the best null. Then, by noting what sort of residual signal is left and how the residual varies as the capsules are rotated, you can get a good idea of how closely the capsules are matched in phase and pattern.

The absurdity of this idea derives not from the physical principles in-

volved (which, after all, are sound), but rather from the use of cheap microphone elements to build a state-of-the-art microphone. If I could find high-quality capsules, with reasonably consistent performance, the idea would not be so absurd. I think I could design a complete B-format SoundField mike system kit, including a UHJ encoder/decoder and a variable pattern/microphone angle control for about 5 percent of the cost of the Calrec mike.

Make Your Own Recordings

Once your microphone system is ready, you are prepared to make a recording. The first consideration is determining which signal goes into which deck input.

I assume you know that the four tracks on a quarter-track machine are numbered from 1 to 4, top to bottom. Tracks 1 and 3 record left and right, respectively, with 4 and 2 taking the same channels when the tape is flipped. Tracks 2 and 4 carry the left rear and right rear channels in quad recordings. Except for the mono (W) channel, you cannot play Ambisonic tapes directly on existing systems. Therefore, you need not be concerned about compatibility, and technically, the tracks could go anywhere. There is, however, a standard that you should follow.

No matter how accurately the tape is slit or how well the transport is designed, the tape never moves perfectly parallel to the heads. There is always some skew, which continually varies, producing interchannel phase jitter. The closer any two channels are to each other, the less the relative skew and the lower the jitter. The Y signal carries the left/right information, and the ear is more sensitive to variation in its positioning than it is to front/back or up/down. This suggests that you should place the W (omni) and Y tracks next to each other and in a position that has the least skew. You do not have to be Albert Einstein to see that this means the two center tracks. W goes on track 2 (left rear) and Y on track 3 (right front). Front/back accuracy is more important than up/down, so the X signal should be as close as possible to W. Therefore, it is recorded on track 1 (left front), and Z winds up on track 4 (right rear).

Now you are ready to set recording

levels. A rather odd standard appears to have come about because the Ambisonic engineers do not like noise-reduction systems, since tracking errors appear as shifts in image position. This standard dictates that you record the difference channels 3dB higher than the W (omni) channel. Since there is no easy way to calibrate your mikes, you need not follow this standard. A figure-8 pattern picks up less total energy than an omni, so those tracks are recorded at a higher level to get a better S/N ratio. To compensate for this, you adjust the gains in the decoder. If, as I have suggested, you are using un-calibrated mikes, you will have to adjust playback pretty much by listening, so record at whatever level you like.

By the way, feel free to use any noise-reduction system you wish. It is true that tracking errors will cause image shifts, but they appear as a subtle expansion or contraction of the width or depth of the sound field, a most innocuous error. I use dbx II, and I have never heard any side-effects.

I have one final piece of advice on recording: if you are using variable-pattern mikes, set them all to omni and record a minute or two of some disposable part of the performance. This will be useful in adjusting playback levels, as I will explain later.

Because it is not practical to make a calibrated recording, you must set the X and Y playback levels by ear. This takes a bit of practice, but really is not difficult to do. If you have followed me thus far, you should not have much trouble understanding how to get properly balanced playback.

Keep Your Balance

First, set your Ambisonic decoder for B-format reproduction, and connect the appropriate tape deck outputs to the correct B-format inputs. If you use noise reduction, it must go between the tape deck and the decoder. Be sure the forward preference control is turned off. Set the layout control to correspond with the speaker positioning.

Second, turn the output level controls on your tape deck all the way down. Set the gain controls on your amplifiers, front and back, to where you would normally set them for tape playback. Play the tape and advance the W output (track 2, left rear) to the point where the sound is slightly softer than what you would consider a normal playback level. Make sure that all amplifier channels play at exactly the same *electrical* (*not* acoustical) level. Once the channels are balanced, leave the amplifiers' gain control alone.

Third, turn off the rear speakers. Increase the Y level (channel 3, right front) until the sound goes beyond the boundaries of the speakers. The sound might be a little "spacy" or have a hole in the middle, but you will fix this later with the X signal. It is important to get this adjustment right. The decoding equations in the sidebar below show that the difference signals (X, Y, Z) always appear in the outputs at a higher level than the omni. This produces the exaggerated effect you are hearing with only the front speakers and no X channel. If you ignore this consideration and go for a sound that matches regular stereo, you will not achieve the full benefits of Ambisonic reproduction.

Fourth, turn on the rear speakers. Bring up the level of the X channel (track 1, left front) until the performing group is clearly in the front and the ambience is properly disposed around you. If you recorded the X and

---

Ambisonic Encoding and Decoding

Assume that zero degrees is "straight ahead," that angles advance in a counterclockwise direction, that theta is the azimuthal angle, and that phi is the angle of elevation or depression. Any sound source is encoded onto the four Ambisonic B-format signals by multiplying it by the following weighting factors, sign included: W = 1 (i.e, no weighting); X = $\cos\theta$; Y = $\sin\theta$; Z = $\sin\phi$.

You can achieve this weighting in a studio by a simple adjustment of level and polarity. For live sounds, you can use either a combination of omni and figure-8 mikes or (preferably) the SoundField microphone. You may also use regular miking, with encoded directions conforming to the "position" of the microphone, although this is less than ideal. The earliest Nimbus recordings use this technique, and it works amazingly well.

The decoding process exactly reverses the encoding process. Here, however, the positions of interest are the speaker locations, not the sound sources. No matter how many speakers you use, *all* the speakers receive *all* the signals. Each component is weighted by dividing it by the sine or cosine of the angle at which that speaker is positioned:

W = 1 (i.e., no weighting)
X = $1/\cos\theta_s$ (where S = Speaker)
Y = $1/\sin\theta_s$
Z = $1/\sin\phi_s$

Of course, if you drop the up/down information, you do not use the fourth channel and the extra speakers.

As you can see, you can realize this playback "matrix" very simply with inexpensive op amp circuits. If the speakers are arranged in a rectangle, then the sines and cosines will have the same magnitude for all speaker positions, differing only in polarity. Therefore, the X, Y and Z signals need be weighted only once. (A trapezoidal layout requires the rear speakers to be weighted differently, since front/back symmetry has been destroyed.)

I have not shown the encoding or decoding for UHJ because I think the concepts associated with the Scheiber Sphere are more important. Suffice it so say that the left and right UHJ signals are manipulated to produce close approximations of the original W, X and Y signals. These W', X' and Y' components are applied to the speaker-feed matrix as the pure B-format signals would be. Some additional response contouring is used to make a smooth transition between the various "regions" where the differing psychoacoustic principles apply and to reduce subjective coloration. □

Y signals at the same level, the settings of the output level controls should be very close, too. This is the hardest part of the adjustment. Be patient, listen carefully, and you will get it right. I find this easiest to judge when I am sitting somewhat off-center.

Finally, sit back and enjoy the most realistic reproduction you have ever heard from your system. Note particularly the more accurate reproduction of instrumental timbre, reduced dynamic compression at high volume levels, the way the ambience "relates" to the direct sound of the instruments and performers, and the almost total elimination of the speakers as apparent sound sources. (Although you might question the presence of reduced dynamic compression, the effect *is* audible. Stereo recordings made under identical conditions show a certain amount of dynamic "gagging" at high volume levels. The effect is almost certainly psychoacoustic and is not understood at this time. Perhaps the ears rebel at hearing loud sounds reproduced from only one direction.)

If you are able to record some music with all mikes set to mono, you can speed up the calibration process. This technique assumes that your variable-pattern mikes have the same on-axis output when set for figure-8 as they do in omni. In any case, it will quickly get you in the ballpark. Follow the above instruc-

---

IF YOU SHOULD HAVE A TECHNICAL QUERY...

. . . please drop us a note explaining precisely what information you need. We will answer your question ourselves or forward your letter to someone with expertise in that area. Make sure to enclose a self-addressed, stamped envelope for our reply, and address your letter to. *Audio Amateur*, Technical Dept., PO Box 576, Peterborough, NH 03458.

Help us out by *not* calling in your question. We have neither the staff nor the time to respond to each query over the phone.

---

tions up to where you set the overall W level, then take these steps:

1. Turn the layout control to the "square" position (straight up). Turn off all the speakers *except* the right front one. Play the part of the tape with all the mikes set to mono.

2. Increase the Y signal (track 3, right front) until you get the best possible null at low frequencies (below 1kHz).

3. Turn off the right front speaker so that neither speaker is playing. Pull the W cable out of the decoder. Turn on the right front speaker again. Adjust the X signal (track 1, left front) to get the best low-frequency null. Turn off the speaker again and reconnect the W cable to the decoder. Set the layout control to the correct position and enjoy.

Extended listening might indicate the need for slight adjustments. Always leave the W signal alone: it is your point of reference. If you change its level, you will have to start the calibration over from scratch. Also, if the front or back seems too prominent, do *not* change the amplifiers' gains more than a decibel or so. Remember, the speakers are cooperating to produce the sound field, and changing the levels too much might make them audible as separate sources. Instead, move your seat forward or back.

After you have had some time to evaluate Ambisonic reproduction, I have no doubt that you will agree with Michael Gerzon, one of the system's designers: "As it (Ambisonics) is the first system design based on a complete mathematical analysis of both system theory and human psychoacoustics, it is possible to say with some confidence that no system appreciably different from it can exceed its performance in optimal surround-sound decoding, so that no further system change is ever likely to be needed for encoding." □

SOURCES

1. Cooper, Duane and Takeo Shiga, "Discrete-Matrix Multichannel Stereo," *JAES*, September 1971. Where it all started.

2. Cooper, Duane, "QFMX—Quadruplex FM Transmission Using the 4-4-4 QMX Matrix System," *JAES*, March 1974.

3. Cooper, Duane and Takeo Shiga, "QMX Carrier-Channel Disk," *JAES*, October 1973.

4. Elen, Richard, "Ambisonics—Sound-Field Experience," *Studio Sound*, October 1979. The practical side of using the Sound-Field microphone.

5. Fellgett, Peter, "Ambisonic Reproduction of Directionality in Surround-Sound Systems," *Nature* (Volume 252, Number 5484), December 13, 1974, pp. 534–538. A brief, nontechnical introduction.

6. Fellgett, Peter, "Ambisonics: Part I—General System Description," *Studio Sound*, August 1975.

7. Fellgett, Peter, "Ambisonics: Part II—Studio Techniques," *Studio Sound*, August 1975. How the producer can create almost any effect, most of which are not available to those using discrete technology.

8. Fellgett, Peter, "Surround Sound '76," *Hi-Fi Sound Annual 1976*. Not what the title says: actually a not-so-subtle attempt to show why RM, UMX and Ambisonics are correct and all the other systems are wrong. A good summation, but difficult for the beginner.

9. Hope, Adrian (pseudonym, Barry Fox), "Ambisonics—Theory and Patents," *Studio Sound*, October 1979.

10. Gerzon, Michael, "Criteria for Evaluating Surround-Sound Systems," *JAES*, June 1977. A thorough discussion of what makes a surround-sound system "right" or "wrong." A good explanation of kernel systems.

11. Gerzon, Michael, "A Geometric Model for Two-Channel Four-Speaker Matrix Stereo Systems," *JAES*, March 1975. A detailed explanation of the Scheiber Sphere.

12. Gerzon, Michael, "Multi-System Ambisonic Decoder, Parts I & II," *Wireless World*, July & August 1977. Part I—Basic Design Philosophy; Part II—Main Decoder Circuits. A DIY article.

13. Gerzon, Michael, "The Optimal Choice of Surround-Sound Encoding Specification," *Preprint No. 1199* (A-5), 56th AES Convention.

14. Gerzon, Michael, "Periphony: With-Height Sound Reproduction," *JAES*, January 1973. The sword of Damocles comes to surround sound.

15. Gerzon, Michael, "Surround Sound Psychoacoustics," *Wireless World*, December 1974.

16. Smith, J. Howard, "Ambisonics—The Calrec SoundField Microphone," *Studio Sound*, October 1979. The design of the mike, from the company that makes it.

NOTE: All the *JAES* articles cited appear in *Quadraphony*, an AES publication covering some of the articles from 1969 through 1975. This only scratches the surface of Ambisonic literature. Except for the few *JAES* articles, everything ever written on Ambisonics up to this point has appeared only in British publications. You might find it fun to grab a stack of *Hi-Fi News & Record Review*, starting around 1970, and work your way through it.