Understanding the Scheiber Sphere

by William Sommerwerck

At the risk of asking our readers to do the mental equivalent of patting their heads and rubbing their tummies at the same time, I'm going to start off this article by asking a question. (We'll see how it relates to our topic later on.) What (if anything) is the factor which distinguishes a good surround-sound system from a poor one? (Don't consider "practicality" — it goes without saying. Besides, as technology advances, what is wildly impractical one year becomes trivially easy the next.) You may want to stop and think about it now, or chew on it while you're reading the article. But be prepared to defend your position!

The Scheiber Sphere is a means for graphically visualizing the amplitude and phase relationships in any two-channel surround-sound system. It is usually applied to systems which are "planar", i.e., not including height information, although an extension could be made. We'll consider the planar version only.

It is named, of course, after Peter Scheiber, the American engineer/musician who introduced the concept of quadraphonic matrixing. (The British steadfastly refuse to acknowledge his contributions and call this presentation the "Stokes-Poincaire Sphere", or the "energy sphere". Rather ungracious, I think.)
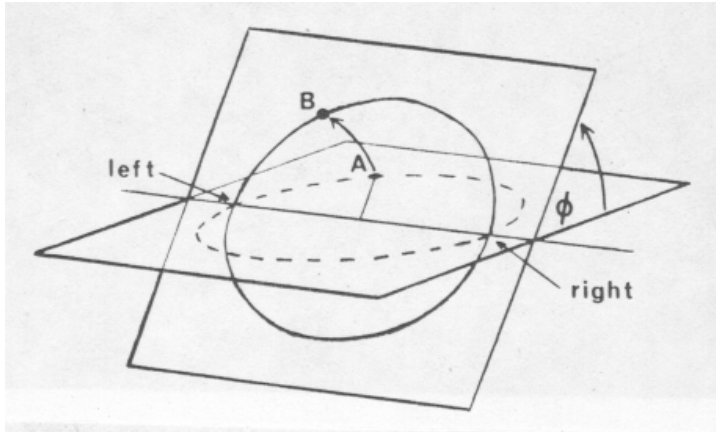

Figure 1. *A sound recorded on the two groove walls of a stereo disc is represented by point A on the horizontal circle (cf. Figure 4). All in-phase and out-of-phase lines of stylus motion can be mapped to points around such a circle. Point B represents a phase shift of the left channel by angle ø and is obtained by rotating the circle about its left/right axis by the angle ø.*
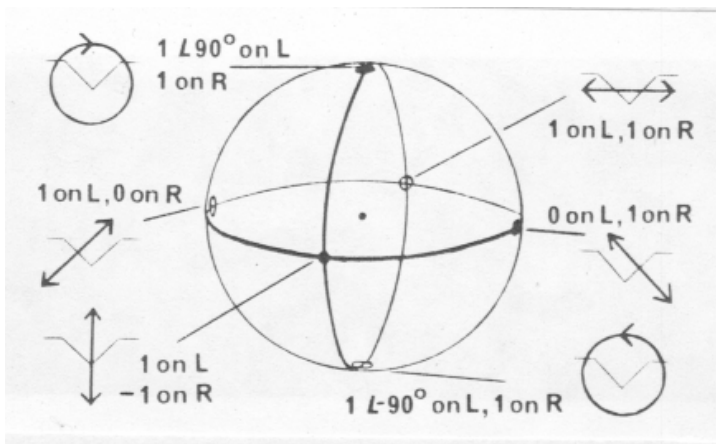

Figure 2 *shows this concept extended into a full sphere. Stylus motions and left and right channel phases and gains corresponding to various points on the Scheiber Sphere.*

The use of a sphere, rather than some other presentation, stems from the need to display three pieces of information at once. In any two-channel surround system, there are three parameters. The first of these is the intended direction of the sound source [for example, 20 degrees left of center front, or 127.2 degrees counterclockwise (CCW) from right center]. This direction is encoded as phase and amplitude differences, which then become the other two parameters. With three variables, we naturally need a three-dimensional space[1] in which to graph them. Ergo, a sphere.

Unlike the earth, there are no markings on the sphere to tell us where we are. The convention is to arrange things so that the encoding position for center-front points "forward" (at least as we define "forward" in a perspective drawing). It is also helpful to think of the Scheiber Sphere as having poles and an equator, as the earth.

How do we represent the phase and amplitude differences between the channels on the sphere? It's fairly simple; phase is easiest to understand. The "equator" is the reference. If a point is above the equator, then the phase on the left channel leads the phase of the right. And vice-versa. The greatest phase difference that can be portrayed is +/-90 degrees (since in traversing the arc from equator to pole, we go through 90 degrees).

Representing amplitude differences is a bit harder. All points are plotted on the surface of the sphere, so all lines from any point on the surface to the center of the sphere would be the same length. Therefore, there is no *direct* way to show amplitude differences. The solution is to do it indirectly. Just as we represented phase differences by angles *above* and *below* the equator, we can show amplitude differences as angular displacements *around* the equator.

"Wait!", you say, "Amplitudes aren't angles!" No, they're not. But we can turn them into angles, through the miracle of trigonometry. We take the ratio of left to right amplitude (ignoring phase, but including polarity — more on this in a moment) and compute twice the inverse tangent[2]:

$$\text{ø} = 2\tan^{-1}|L|/|R|$$

For all possible amplitude values (including zero in either channel), this will give us a full range of angles, from 0 to +/-180 degrees.

Angles are not absolute; they must be measured from some reference. On the Scheiber Sphere, the 0-degree reference is "due right", so-to-speak. You would logically think that this reference should match the center-front reference chosen earlier, especially since many surround systems compute angular offset from center-front. Well, the reason for picking this seemingly odd starting point is the result of the way we chose to plot amplitude differences as angles. One or two examples will make it easier.

The most obvious case is a center-front source. Every two-channel surround system encodes it with equal amplitude on both tracks. The ratio of left to right amplitude is thus 1, the arctan of 1 is 45 degrees, and twice that is 90 degrees. If we rotate CCW 90 degrees (the convention is that + angles go CCW), lo and behold, we are at center front, the same position as our source!

Let's try another one. All two-channel systems encode center-back sources with equal amplitude, but out of phase. Thus our ratio becomes -1, the arctan of whIch is -45 degrees, doubling to -90 degrees. That brings us around to the "rear" of the sphere, again mimicking our encoding position.

An "obvious" pattern is emerging: the longitude of any point on the Scheiber Sphere is the same as the intended angular position of the sound source. Right? Wrong! These things don't have anything to do with each other!

Yes, I know this is confusing. We've just seen two cases where the angular position of the source and its encoding on the sphere *did* match. But they don't have to! In principle, *any* sound source position could be encoded with *any* combination of amplitude ratios and phase differences. There is no law of nature, or any mathematical principle, which *requires* a two-channel surround system to have its Scheiber Sphere positioning match the positions of the sound sources. It is this single point that is the most confusing thing about the Scheiber Sphere. Understand this and you'll have the Sphere "licked". (The Cube, of course, is another matter.)[3]

Of course, what we *can* do and what we *should* do are not necessarily the same thing. We wouldn't want to have source positions encoded in a completely arbitrary fashion, with encoding points peppering the surface of the Scheiber Sphere as stars are flung across the night sky. No. There has to be some order. To cite the most obvious example, if we wished to maintain stereo compatibility, the front sources would have to be encoded in much the same way as they would be in regular stereo. This means their encoding locus would have to be on (or close to) the equator, and extend symmetrically over about $\frac{1}{4}$ to $\frac{1}{2}$ the Sphere's circumference.

Here's the main thrust of this article: because the Scheiber Sphere allows us to geometrically visualize the amplitude and phase relationships in a two-channel surround system, it gives us a useful reference point to consider whether one type of encoding is better or worse than another. We can decide "better" or "worse" on rational, mathematical grounds, rather than "I like it" or "I don't like it".

In order to make this kind of decision, we have to remember just what the Scheiber Sphere representation means: we're mapping a 2-dimensional space (every possible position in a circle around the listener) into a 3-dimensional space (the Sphere). So, what makes a good mapping?

For one thing, it should be "smooth" or "continuous". This means there is "gradual", rather than abrupt, transition from one point on the sphere to the next. After all, the circle of possible sources around the listener has no such discontinuities; why should the encoding? We'll have more to say about this later, with even better justifications.

We also want the encoding line on the Scheiber Sphere to be as long as possible (subject to certain restraints, as we'll see). Why? Well, we want to use as much of the "encoding space" as we can. Consider the worst possible case: all sources encode to the same (any) point on the sphere. That's mono! (There's no way to distinguish one encoded point from another — they're all the same point — so there's no way to specify directionality.) But by using as much of the encoding space as possible, points that are "far apart" in "real space" will also be relatively "far apart" in "encoding space". This not only makes the decoder's job easier, but also makes the system more immune to phase and amplitude errors. The more encoding space used, the greater the amplitude and/or phase differences which encode any two points, and so the *less* relative effect any phase or amplitude errors have on positional accuracy.

Combining these two criteria suggests an optimum encoding locus — a circle. A circle is the "smoothest" curve we can draw on the sphere. And if it is a great circle (a circle with the same center as the sphere), it will have the greatest possible length (Figure 3).
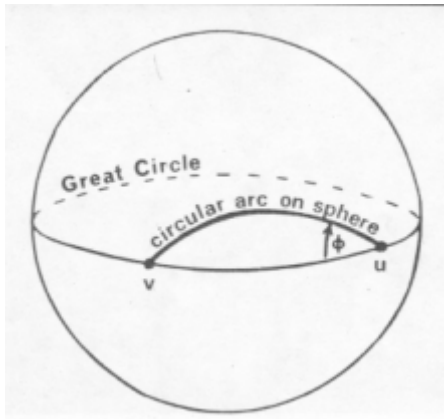
Figure 3. *An example of the locus traced on the Scheiber Sphere by a sound pan-potted between two channels.*

Someone will no doubt point out that we could increase the length of the encoding locus by twining it around the sphere. (The most obvious example would look like the stitching on a baseball.) This is true. The catch is that, although we have moved any two points further apart *along the line*, they will almost certainly be closer together *on the sphere* than they would be if they were on a circle. And it's the position on the sphere that counts. A circle is the best compromise between using as much of the encoding space as possible while keeping the encoded points "as far apart" as possible. Any other encoding scheme trades off one against the other.

Now that we've established some criteria for the "goodness" or "badness" of an encoding scheme, let's look at some of the systems and see how they stack up.

The earliest system was "Regular Matrix". It appears on the sphere as a line equivalent to the "equator". The locus does not rise above or below, showing that there is no relative phase shift between the channels. RM encodes directionality by simple amplitude and polarity variations. Center front is lateral groove motion, center back is pure vertical motion, due right is a 45-degree motion of the right wall only, and so on. Figure 4 shows how the stylus motion is half the angle of the intended direction. (This approach is essentially equivalent to an early proposal of Peter Scheiber.)
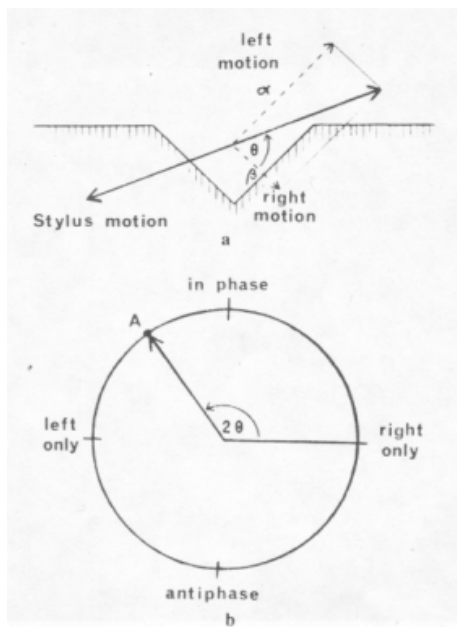

Figure 4. *Angle θ of stereo stylus motion in a record groove corresponds to an angle 2θ on a circle representing possible stereo positions.*

Looks awfully unsophisticated, doesn't it? It's actually an excellent way of encoding directionality. It meets our criterion of optimality — it's a great circle. And it has another advantage which is not immediately apparent. It encodes directionality explicitly. If I want a sound to be encoded at some particular direction, the RM specification tells me *exactly* what kind of a signal to lay down. There is no reference at all to four original sources, or four tracks of a tape, or anything like that. "Quadraphonics" never appears.

The reason for this is that "Regular Matrix" isn't a matrix at all — it's a kernel. In a matrix, a set of *distinct* values is transformed from one dimensional space to another dimensional space (usually with different dimensions, but not necessarily). For example, in a quad matrix, four channels are transformed into two. Or for that matter, in a "discrete" system, four channels are transformed into four. In a kernel system, the same kind of dimensional transformation takes place, except that it occurs as a *continuous* function, not just for discrete values. In this way, RM is able to dispense with the idea of four channels, or four sound sources, and make a direct transformation from the space around the listener to encoding space.

The advantages of such an approach can be better seen by looking at the Scheiber Sphere representation of QS, shown in Figure 5a. QS is a quadraphonic system. It takes four unrelated sources, mixes them down to two tracks, and tries to separate them later, with as little crosstalk as possible. The locus shown for QS (sometimes called "belt-and-suspenders") actually shows far too much. You see, the QS matrix is defined, like any matrix, for *only four points*. These are the four "corner" points, roughly equivalent to the 45-, 135-, 225- and 315-degree points in RM. (Hence Sansui's claim that QS is "derived" from RM.)
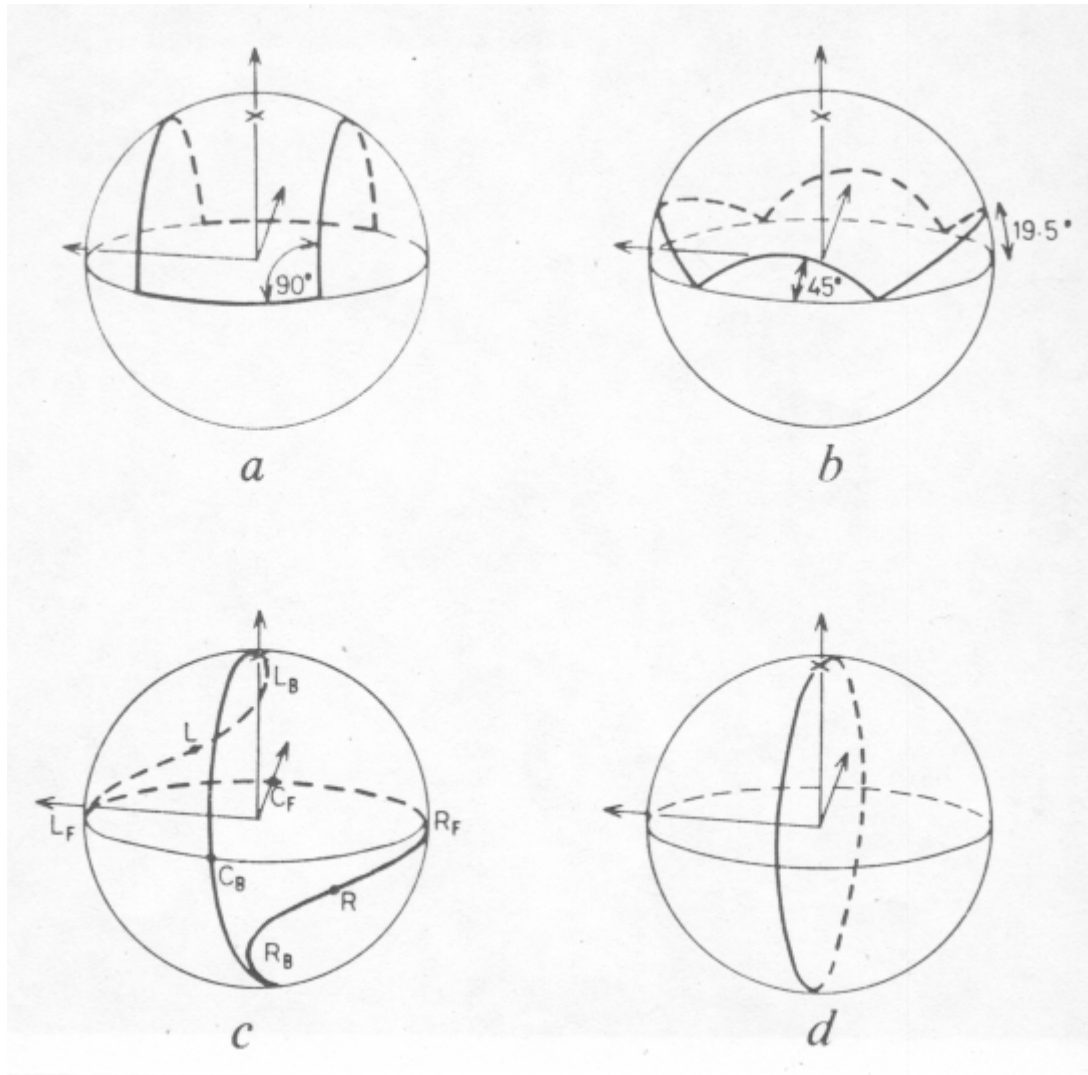

Figure 5. *Some representative horizontal pan-loci. a, Sansui QS; b, closer approximation to great-circle locus consistent with four-channel master tapes (not in commercial use); c, CBS SQ (note cusps and left-right asymmetry due to choice of front-sector mapping and limitations of encoding from four pair-wise blended channels); d, a great-circle locus consistent with ambisonic encoding.*

The rest of the Scheiber representation for QS applies to the locus for sounds which are intended to appear between the four corner points. As in all "quadraphonic" systems, QS assumes that the four loudspeakers carrying our four signals are to be treated in pairs, with sounds "between" the speakers being localized by panning the sound between the pair. (The "belt-and-suspenders" is derived from this assumption.) This is called Pair-Wise Mixing (PWM) and it *doesn't work!* This is not a matter of opinion; it's fact. PWM cannot position sound sources to the side of the listener. You can prove this for yourself.

Set your system to mono. Face the front speakers, standing about six feet away. If you are properly centered, you should hear a sharp, precisely-focused image, as if there were a single source of sound in front of you. Now slowly turn to the left or right. Stop when you're at a right angle to the speakers. Where is the apparent sound source now? It's hard to tell. It *should* appear to be coming precisely from the right or left. Instead, it is blurred and indistinct. Or it may appear that the speakers have become two independent sound sources. In any case, you didn't get what you expected! Something more complex than simply putting the same signal at equal strength on adjacent speakers is needed to get convincing side images. Pair-Wise Mixing *doesn't work*. Unfortunately, all "discrete" systems and all matrix systems are built around this incorrect method of encoding source position. There's an expression in the computer industry: GIGO. "Garbage in, garbage out." You can't get useful results if you start with the wrong assumptions!

If RM is the origin of QS, QS is certainly a bastard child. Somebody at Sansui wasn't thinking. They took a system which specifies *all* directions and stripped it to the point where it specifies only four!

QS has other problems. Notice the abrupt shift where the "belt" and "suspenders" meet. This is called a cusp. It occurs because the matrix is only defined at four points, and we have filled in the missing sections by assuming Pair-Wise Mixing. Even if QS were otherwise correct (which it isn't), the cusps are undesirable. First, they represent an abrupt change in the encoding locus. This is inconsistent with the smooth way things vary in the "real world". Second, it makes the points on either side of the cusp closer together than they should be. Therefore, *sources* on either side of the cusp will be "closer together" in the decoder's output than they should be. Logic circuits can't help; in fact, the cusp tends to blur the distinctions the logic needs in order to act properly.

If QS is a "naughty" matrix, SQ is, by comparison, absolutely sinful. Figure 6a shows the SQ locus, for clarity, for just the front and rear quadrants. (Notice that rear positions are encoded simply by phase changes. Since all rear positions are on a line of longitude, they all have the same relative amplitudes in the encoding. This is an excellent example of the longitude of the encoding *not* corresponding to the horizontal position of the sound source.) In Figure 6b, the pair-wise locus for side sources has been added. As you can see, SQ really "breaks the rules" we've established. Notice especially the severe cusps at the transition points where left-front and right-back signals "connect" to the rest of the locus.
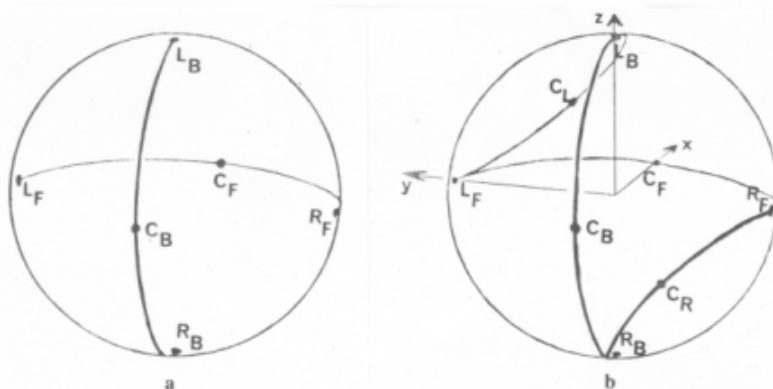

Figure 6. Pair-wise pan-locus of SQ matrix system. Angles at Lf and Rb are 0°; those at Rf and Lb are 180°

Someone will probably point out that all we want to do is get back our original four channels, and that the shape of the matrix's locus really doesn't matter. But this misses the point. Surround-sound systems were intended either to accurately reproduce a "live" sound field, or to allow the producer complete freedom in placing sound sources around the listener. Any system which cannot do these things is flawed from the beginning, since it doesn't meet the most basic requirements of surround sound.[4]

We now have the answer to the question asked at the beginning. The single most important characteristic of a surround-sound system is its ability to accurately reproduce directionality. (If you said "good channel separation", you can now see why that is not applicable. Since Pair-Wise Mixing among four channels doesn't work, "channel separation" has no meaning.) Our point of reference is either live sound or the performer's imagination. A four-channel master tape is not a legitimate reference since it does not encode directionality correctly.
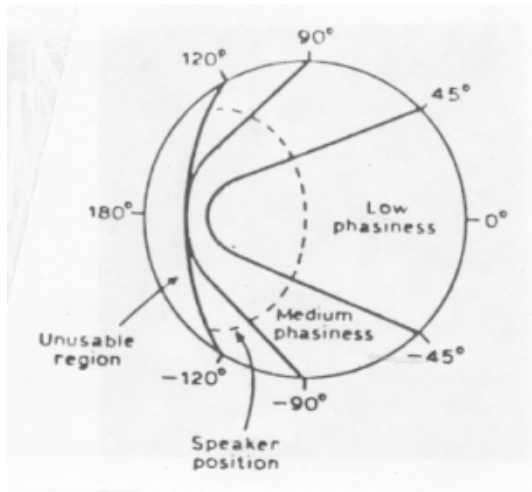
We have already looked at RM and seen that it is "good". The one remaining system is also "good". It is the UHJ kernel for Ambisonic encoding. It is shown in Figures 7, 8, 9 and 10. (In each case it is shown from the side or top, rather than in a "perspective" view.)

Figure 7. *Quality of mono and stereo reproduction shown on Scheiber Sphere viewed from right side. "Speaker position" curved indicates appearing to come from one speaker only in stereo.*
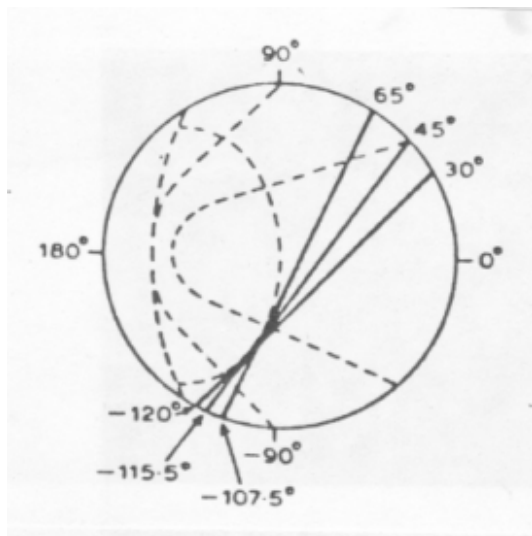

Figure 8. *Three possible choices of two-channel encoding system having optimized mono and stereo reproduction.*
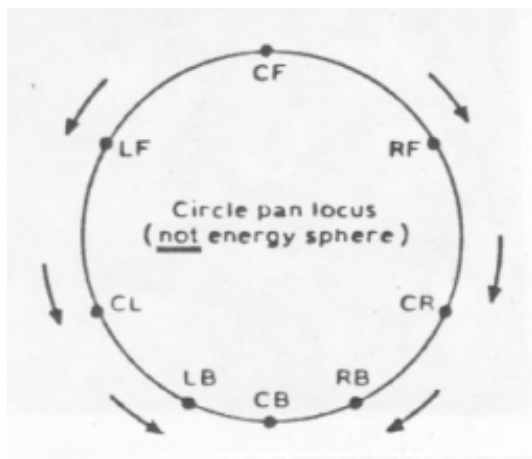

Figure 9. *Optimized non-symmetric distribution of different encoding directions within the circle "pan locus" of the systems shown edge-on in Figure 8.*
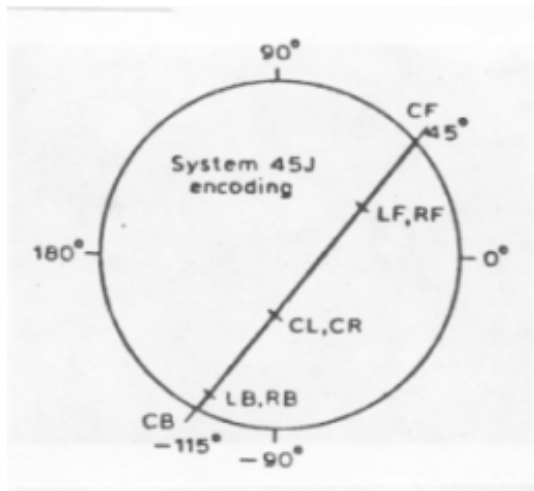
Figure 10. *Scheiber Sphere (side view) picture of two-channel version of System 45J encoding.*

Like RM, it is a circle locus, with all the advantages attendant thereto. But it has two features which distinguish it from RM. The first is that the source encoding is not symmetrical from front to rear. Almost 240 degrees is used to encode the front hemisphere (from CL around to CR), leaving the rest for the rear. The reason for this is that the ear is much more critical of positioning accuracy and stability in the front than it is the rear. By giving plenty of encoding space to the front sources, the ear has more information to work with, and (as explained earlier) the apparent positions of the front sources are more immune to slight changes in amplitude or phase.

The other difference is that UHJ is tilted. This has several interesting consequences. Think back to RM for a moment. A center-back encoding source is encoded anti-phase. The closer a sound is to center back, the more nearly anti-phase it is. Now, you know what your system sounds like when the channels are anti-phase: the image is blurred and indistinct, and there can be a "pressure in the ears" sensation. Well, a stereo listener would hear rear sources in RM reproduced with the same kind of effect. (This also applies to SQ and QS.) This isn't desirable, unless the producer specifically wants that effect.

Anti-phase is the worst possible condition for a signal. If we add or subtract phase shift, the ear will hear an improvement, since we are moving the phase closer to either 0 or 360 degrees. Therefore, by tilting the UHJ locus, we have added phase shift to rear-hemisphere sources, *reducing* their phasiness for stereo listeners.

There is a further benefit. As we have seen, QS, RM and SQ all encode center-back sources anti-phase. When the two tracks are summed for mono broadcasts, center-back sources cancel out, and sources near center back are attenuated in proportion to their closeness. Traditionally, therefore, matrix recordings have not included center-back sources, simply to accommodate the guy listening on his car radio! Since UHJ's tilt removes the anti-phase condition, sounds may now be placed at any position.

Since you can't tilt the rear down without bringing up the front, some phase shift is introduced there, too. This has the advantage of reducing the 6-dB "build-up" of a center-front signal to about 3 dB, a further enhancement to mono compatibility. It also introduces some phasiness into the sound for stereo listeners. The exact tilt of the UHJ locus was chosen by weighting the conflicting factors of how much reduction in rear phasiness was desirable, how much increase in front phasiness was acceptable, and how closely to "constant level" sources should be in mono. The result is an encoding that is not quite "great circle", setting a bit below the sphere's center. It was felt that improvements (or the best compromises) in the areas discussed, outweighed the slight loss in available encoding space.

So far, we've only been discussing encoding. What about decoding? (We are using "decode" in the sense of deriving appropriate signals to feed to our loudspeakers, not trying to get back to the "original" channels. In the cases of RM and UHJ, we have seen that this consideration has no meaning.) You may be surprised to learn that, given a great-circle, or near great-circle, encoding

locus, it is possible to derive speaker-feed signals (for any number of speakers) which correctly satisfy a large number of psychoacoustic criteria for accurate sound localization. And no kind of "logic" circuitry is required for this: a "simple" decoder suffices.

In other words, the two-channel UHJ system is capable of a performance level that "discrete" four-channel systems cannot even *approach*. What does the listener hear? Well, all awareness of the speakers as sound sources vanishes. Particularly in the front, sounds are spread evenly, with no "bunching up" at the sides, or a "hole in the middle". The room is "full" of sound, with no "emptiness" to the sides. The acceptable listening area is enormous — far larger than for stereo or quad. You can even walk near the speakers, and you have to be almost on top of one before its output dominates the sound.

Think about it. This is the way it *should* be! Surround-sound reproduction ought to be much superior to stereo, and, with UHJ, it is. But "discrete" and matrix systems are actually worse. You have a narrower listening window than with stereo, side sounds aren't localized properly, and the front imaging is no better than it was before.

Remember that the goal of surround reproduction is accurate localization at any point around the listener, nothing more or less. "Discrete" and matrix systems are inherently limited to gimmicky "quadrifontal" (four-source) effects, and are incapable of correctly reproducing hall ambience. (Remember that an accurate system can be degraded to produce "artificial" sounds, but not the other way around!)

There are still many SQ supporters, which is unfortunate, since SQ, as well as QS and "discrete" systems, has so many deficiencies. It needs very sophisticated hardware (i.e., directional enhancement decoders) to do poorly what UHJ does simply and superbly. Larry disagrees with me (and he's probably right), but I feel that the adoption of SQ as the surround-sound standard would kill surround sound. Once tens of thousands of SQ decoders had been purchased, and people then had a chance to hear UHJ, they would be so disillusioned that they might give up on surround sound altogether.

Some of these SQ supporters are claiming to have "improved" or "enhanced" versions of SQ. These versions sufficiently alter SQ encoding/decoding parameters that they are not fully compatible with the $500 to $1000 DES decoder you just bought. (A pleasant thought!) And to the extent that such systems really *are* improved, they must approach the kind of "great circle" encoding locus, and psychoacoustically-correct decoding, that are already embodied (and patented) in UHJ.

If someone comes along claiming an "improved" SQ system, ask him the following questions: What psychoacoustic criteria for correct localization does your system meet? Is the listening area wider, and frontal imaging better, than with regular stereo? Is the system part of a comprehensive mathematical model of the imaging process, with predictable and controllable effects, or is it just a hodge-podge of ideas you "thought would work"? Is a simple, "unaided" decoder capable of better sound quality than the use of four "discrete" channels? Caveat auditor! UHJ is *all* these things and more. I only wish there were some way I could give a demonstration for all our readers.

_____

[1]Describing a "space" as "three-dimensional" isn't redundant. A mathematical "space" may have any number of dimensions, from 1 up.

[2]A number of readers have complained that I'm too quick to jump into trig. The problem is that there's no way to explain certain things without delving into "higher" math. (Of course, since it had no bearing on the fundamental principles involved, I skipped over the matrix algebra in the Tate DES article in the Autumn issue.) And there's no way to explain everything so that everyone can understand it. I may be a degreed engineer, but there's plenty of technical material I can't follow. If I think it's important (or just interesting) to understand, I make the effort to learn more about it, either by asking experts or studying books. If you want to learn about trig (or most other technical subjects) in a semi-painless way, try Schaum's Outline Series.

[3]There's another subtle point of confusion which may be bothering readers. You'll remember that relative phase was encoded as elevation or depression with respect to the equator. Yet I said that signal polarity was considered along with amplitude differences. Traditionally, we think of polarity as being related to phase. For example, if I invert the polarity of the amplifier connection to one of my speakers, I say that it is "out-of-phase" with respect to the other channels. So where so we get off putting the polarity in with the amplitude, rather than with the phase?

Technically, inverting the polarity of the signal *does* introduce a phase shift of 180 degrees, for all signal components. But it isn't the kind of shift we associate with ordinary phase-shifting circuits. It can be obtained simply by reversing the signal leads (it's that practical — it isn't always) or by introducing another gain stage (most amplification inverts signal polarity). Such a phase *inverting* circuit is *never* called a phase *shifting* circuit.

The British recognize this difference in their terminology. Instead of saying that a speaker pair wired the wrong way is "out-of-phase", they say it is "anti-phase". This better expresses the idea of polarity inversion.

I agree that this is not the best of explanations and that there is same room for argument. But the simple inversion has never been considered true phase shift (at least in any text book I've seen). And if we didn't consider polarity, there would only be half as many possible angles which could be used on the Scheiber Sphere. It doesn't matter whether the convention of associating polarity with amplitude is correct, so long as it's useful (which it is).

[4]If you own any "discrete" tapes, or CD-4 records, you'll have noticed that you never hear sounds coming from the sides. This is no accident. Either consciously or unconsciously, producers realized that sounds panned to the sides never appeared there. So they limited themselves to placing sounds across the front, with distinct sources at the left- and right-rear, and occasional center-rear sources. If there were ever any more profound condemnation of "discrete" systems, it is these recording practices.

Last updated: August 1, 2000