Realities of the Stereo Illusion

A musician's guide to what makes stereo work, and how to make it work for you

When we listen to sound recordings, we rarely notice the playback loudspeaker. Instead we hear pianos, singers, voices, electric guitar amps, and orchestras. We have little sense of the loudspeaker itself as a musical instrument, or of its essential musical character. However, by physical definition, the loudspeaker is, in fact, a musical instrument. Without formal recognition, it has become the predominant musical instrument of our time.

For producers, engineers, composers, performers, and other music professionals, the quality and intensity of the stereophonic illusion determines the essence of our craft and the basis of our business. Despite this fact, many music professionals do not understand the basic principles of how stereo perception actually works, or why we like one recording over another. A basic understanding of the principles of stereo and a few simple methods for listening critically to a recording can be powerful tools for more successful recordings and for a heightened awareness of what makes great recordings work.

Beyond the aesthetic benefits, there are other advantages to the effective use of stereo. Studies by Floyd Toole, an active researcher on the perception of loudspeaker playback, show that stereo can significantly enhance the perceived quality of a recording.

Music Production & Engineering Department Chairman David Moulton is an acoustical consultant and a developer of speaker technology.
Furthermore, it tends to mask perceived defects in frequency response, distortion, and other sonic aspects that are clearly heard in monaural playback. In other words, effective stereo makes whatever you do sound better.

**The Phantom Image**

In audio reproduction, as in all acoustics, there are vast differences between what we hear and what we think we hear. As noted author and recording engineer John Worley once stated: “Mono is a hallucination where the sufferer believes that he/she hears music coming from a small wooden box. Stereo is a more severe ailment, where the sufferer believes that he/she hears music coming from an imaginary point in space between two small wooden boxes.” A key to this “hallucination” is the “phantom image.”

The phantom image is an illusion created when two loudspeakers emit exactly the same sound (meaning the same complex wave, phase-locked together). This situation is paradoxical in two ways. First, it is a condition that does not occur in pre-technological nature. Second, it is a condition that our perception system is not equipped to identify.

---

**THE MONITORING SYSTEM**

The monitoring system is your musical instrument, your point of departure for creating an exciting and successful musical experience for your listeners. It is absolutely critical that you be comfortable and familiar with your monitoring system, and that you have reasonable control over its operation.

An interesting part of the production problem is that the music you create on your monitoring system may be played back over a wide variety of other systems by your listeners. Therefore, a large part of your work will involve making sure that your production will sound good over a wide range of playback systems. For this reason, it is not absolutely necessary that your system have “perfect” performance characteristics. What is important is that you know how your system sounds in relation to other end-user systems, which may include home hi-fi systems, table radios and televisions, automobile playback systems, and headphone-based Walkman-type units. There are a few basic criteria your monitoring system should fulfill to help you meet these challenges effectively.

Your system should be known to you. You should be intimately familiar with the system and its behavior with a wide variety of program material, including your favorite music and some conventional commercially successful material from all genres.

Your system should be known to the industry. It should be similar to the types of monitoring systems normally used in audio production work so that you can reasonably assume that what you are hearing bears some relationship to what the original creators/producers of the material heard. For this reason, stick with tested systems, playback enhancements, and other sonic treatments are to be avoided.

Your monitors should be installed in a stable, robust, and reasonably standard configuration. If you can’t exercise considerable control over the “acoustics” of your playback space, near field monitoring (i.e., having the loudspeakers very close to you) is probably desirable.

Another variable in your monitoring system is the playback level. Because playback level has such a dramatic effect on how we hear and perceive musical qualities, you should take several steps to maintain a reasonable standard.

First, your monitoring level should be carefully selected and documented. The reference level should be the loudest level that you are reasonably comfortable with when you first start listening for the day (I use 90 dB SPL, when “pink noise” is coming from both speakers and left and right meters are both showing 0VU). Using a sound level meter (Radio Shack sells one for about $50), you can determine what this level is experimentally, and then routinely replicate it in every work situation. (Another toy for your Halliburton case.)

Once determined, your monitoring level should be used for all decisions and adjustments. When you are actually working in production, all decisions should be made with the system set at the reference level. You should mark the level of the monitor level control and avoid touching it except for special cases. Do not routinely fiddle with the level and do not turn it down when you need to talk to someone (shut off the tape deck or mute the console instead).

While maintaining a constant playback level is important, also listen to your project at a lower level at regular intervals (some consoles have a “dim” switch, which reduces the level by 20 db). Due to anomalies in the way we perceive frequencies, the effective equalization of the recording will be dramatically different at the lower level. Be sure that the recording works at both the loud and soft levels.
The phantom image is a reaction to this paradox. To a listener on the median plane (any point equidistant from both speakers), such sounds are perceived as emitting from neither speaker, but instead from a point in space midway between them. This illusion is what makes stereo work.

The phantom image can be enhanced and modified to tremendous effect by the use of time delays and reverbation. In acoustic recording, the phantom image is the result of appropriate stereo microphone configurations and placements.

**Loudness Differences**

It is commonly believed that localization of sound (and therefore, stereo) results from small differences in loudness and time of arrival at our two ears. This is true, in part. But the reality is a good bit more complicated than that. While we are sensitive to both loudness and time differences, their relative importances are quite different.

In acoustic events, each sound arrives at each ear with a slightly different amplitude. In fact, this difference is comparatively unimportant for the localization of sound. Only gross differences in loudness (a doubling of power, or 3dB) have a significant effect on where we perceive the sound to be coming from. This means that the pan pot on mixing consoles (or the balance control on your receiver) doesn’t really do much if you sit down and listen carefully, you will notice that the pan pot only allows you to place the sound at the left or right speaker or in the phantom image in the middle. Subtle placements between these areas are unstable, and are often more a function of the tilt of your head than the fineness of your pan-pot turning.

**Time Differences**

More importantly, you should be aware that in a normally reverberant room, each sound you hear consists of multiple artifacts—the result of the multiple paths (direct and reflected) that the sound takes through the room to reach your ears. These artifacts are detected individually by their angles of arrival at each ear (each angle of arrival has its own unique sound). However, the conscious perception is not of an ensemble of artifacts, but of a single sound—a timbral summation of all the artifacts.

Regardless of that fact, the auditory system is extremely sensitive to differences in time of arrival of such artifacts. We localize a sound based on the angle of arrival of the first artifact, the direct sound arriving from the source. This is called the “precedence effect” (or the “Haas effect,” after Helmut Haas, who first published data about it in 1949).

The precedence effect takes place with even the slightest variance in time. When listening to the same (monaural) signal coming from two loudspeakers, for example, we hear the sound as coming only from the near speaker, even when the difference in distance from the two speakers is only six inches. Tilt your head and hear the sound move.

In the studio, we can use this effect by inserting slight delays into the audio path of one channel or the other, which yields far more realistic and stable shocks in localization than does the use of the venerable pan-pot.

Keep in mind that this sort of spatial information changes when the sound and its artifact arrive from the same direction (as in monaural playback). In this case, the two waveforms are “summed” by the listener, yielding significant destructive interference due to the cancellation of some frequencies of the sound. A significant change in timbre is heard. Therefore, what is spatial information in stereo becomes timbral information in mono. Producers and engineers, beware!

When applied to two loudspeakers in the same space as the listener, the precedence effect dictates that the same sound arriving from both speakers will be perceived as emitting only from the nearer one. Even though the farther speaker may have the same amplitude, its sound will not be perceived. This is why we must be on the median plane.

In stereo listening, as mentioned above, a six-inch offset is enough to conceal the presence of one loudspeaker when common information is being played over both.

When we place ourselves correctly between the speakers and the recording presents an effective stereo space, the result can be extremely powerful. Perceptually, the information we receive is equivalent to information from a room (and speakers) that are different than the room that we see we are in. We hear sonic artifacts of a new, unseen space surrounding us.

The presence of this new space is the basis of the power of stereo, that gives it its impact and entertainment value. Our brain is fascinated with the illusion: How can we be in two places at once? Which space are we really in? And the experience keeps us coming back for more.

**Problems**

While the illusion can be fascinating and its application lucrative, it can also hold many traps for the music producer. Because of the vast array of end-user systems that may listen to your work, it is best to keep in mind the following when doing a final mix or critically listening to your work.

If there is any reason to believe that a production will be broadcast in any format, it may be very important to consider how the production sounds in mono. More than 50 percent of all FM listening in the United States is in mono (according to a 1987 National Public Radio study) as well as 90 percent of television listening. For international broadcast, the mono is even more prevalent.
Also, although analog discs are clearly on the wane, it is wise to keep in mind their limitations. Specifically, stereophonic signals with strong-difference components or out-of-polarity signals cause the cutter head (during mastering) and the stylus (during playback) to move vertically. This can result in damage to the cutter head (if it strikes the aluminum substrate) and can cause the stylus to skip during playback, resulting in warranty returns of the retail product. Low-frequency signals are the biggest problem, particularly if they are loud and close to 180 degrees out of phase and/or out of polarity.

**Study Tools**

In the continuing quest for better stereo reproduction, there are several tools that should be in every engineer's Halliburton case.

The oscilloscope is an excellent tool for observing the time/phase relationships between the left and right channels of a stereo recording (using the XY or "Lissajous" display), and for observing waveforms, particularly in regard to signal overload.

The Real Time Analyzer (RTA) gives a display of the relative loudness of various parts of the audio spectrum on an ongoing basis (hence in "real time"). With repeated use, these devices yield extremely valuable insights about the nature of the program material and possible problems that will be encountered by end users.

Then, of course, there are the two essential devices that everyone carries with them and that many do not use to full potential: cars. Your ears are far more sensitive and discriminatory than almost any test equipment yet devised. Learn to use and trust them.

When you perceive a problem, you can reliably count on the probability that there is, in fact, a problem—even if you have difficulty expressing what it is or verifying it objectively. When you perceive a problem, keep working on it until you are really sure (with your ears and your equipment) that the problem has been resolved.

Stereo can be a powerful illusion. With a little understanding and a lot of care, you can harness that power for the benefit of your own music. Good luck and happy hit-making.

---

**A NEW WAY TO LISTEN**

There are three basic aspects that make up a stereo image. The first includes those sound elements common to both the left and right channels (A+B). The second includes everything that is not part of the common sound elements (the entire stereo signal minus the central image, known as A-B). The third includes all signals present on either the left or right channel alone (A, B). For analysis purposes, this third aspect is considered part of the second.

Disc-mastering engineer Bob Ludwig first suggested to me means of analyzing a recording on these terms which has proven invaluable. He noticed that the A-B version of a recording usually carries special characteristics that are unique to each producer. Listening to the A-B aspect of a recording offers unique critical insights into the way a recording is made and the way the producer works.

To compare the A+B and A-B aspects of a recording in the studio, route the stereo signal you wish to study through two inputs on the console, panned to left and right. This yields conventional stereo. Select "mono"-at-the-monitor control section to listen in mono (i.e. A+B). Then also select phase reverse (雍) on one of the two inputs to listen to A-B. Adjust the levels of the two channels to get complete cancellation of the A+B signal. Then, using the phase reverse control, switch back and forth between A+B and A-B as desired to come to understand the recording and its elements.

To achieve this process at home, in a simple and somewhat clumsy method, reconnect your speakers to your power amplifier or receiver so that each speaker (or only one, if you prefer to leave the other not hooked up) is connected only to the two red, positive, "hot" terminals. This will yield an accurate a-B signal.

In popular or multi-track recordings, the differences between A+B and A-B can be vast and intriguing. You may notice an approximately equal balance in interest between the two elements, although most central musical elements are usually in A+B (lead vocals, kick drum, bass). You may also notice a high-frequency rhythmic "framing" that surrounds and supports the A+B music. This is usually equally balanced left and right and often includes doubled rhythm guitars, keyboards, or high-frequency percussion parts.

This listening method will also reveal numerous production effects and practices including the use of reverberation and delays. In addition, various tricks with occasionally doubled words, rhythmic ping-ponging between channels, special out-of-polarity effects, and other elements of multi-track stereophony will be clearly revealed.

Most importantly, direct observation of A-B leads to a much clearer sense of how pop/rock recordings are assembled—as a kind of spatial polyphony evolving from the musical and spatial tension between A+B and A-B elements. Recordings where this polyphony is strong are effective over the broadest range of stereo systems and they usually do well in mono, as well (except that the spatial effect is lost).

While A-B listening is not for regular listening, it can provide fascinating insights. I routinely listen to all recordings at least once in A-B for critical study.