The Use of Early Delays

There's more to delays than you think—you can create entire room ambiences with early delays. Here's how...

By Dave Moulton

Back when I got into this crazy business (in the late '60s), delays were a wonderful but not very subtle effect. They were generated by using the travel time between the record and playback heads of tape recorders. Varispeed was not a standard item then (in fact, we were still just trying to get the decks to run on time—forget about varying their speed!), so the available delay times were simply the times that occurred at 7.5 and 15 ips (about 130 and 65 milliseconds, respectively). Sending the delayed version of a sound to the opposite channel was awesomely, totally creative, even if it was obvious and audible to a fault.

However, when I tried sending the delayed signal to the reverb and *then* to the opposite channel, that was beyond totally awesome. That was, well, mega-awesome. As I worked with this effect, which has now become a cliché and the basis for an entire family of signal-processing tricks, I began to think about the acoustics and psychoacoustics that make these effects work. Since then, little by little, I've worked up my own model of an approach to the use of early delays, which is based on the acoustics of rooms.

The general use of early delays didn't become possible until digital delay lines became commercially available in the mid-'70s. The tape delays we used before were too long, and were blatantly obvious when we used them. We perceive a delay of 65mS (milliseconds, or thousands of a second) as an echo. The dividing line between early, inaudible delays and audible echoes seems to be around 40mS, depending on the material. When I began to play around with delays, I was struck by the fact that really short delays aren't consciously heard. This is exactly what the Haas Effect is all about, as I discussed in last month's article on panning, 'The Phantom Image.'

However, the presence of these short, inaudible delays in normal everyday acoustical listening makes use of a very powerful set of features in our auditory perception system. Because of this, the control of these delays is a basic and useful tool for the producer and recording engineer. What makes these time-based effects even more powerful is that their impact is subliminal or preconscious, so that we're not even aware of how much impact their presence is having on us.

I define short, or *early*, time delays as the reiterations of sound events that follow within the first twentieth of a second, or 50mS, of the original sound. They are usually caused by room reflections, and are closely allied to the idea of ambience, the sonic signature or character of a room. Generally, such reflections have an amplitude within 10dB of the original sound. In conventional reverberant rooms, there are between 6 and 10 of them for each sound event, all arriving at our ears from different directions.

Odd as it seems, a sound is not usually single event, but an amalgam of at least seven events. A normal rectangular room of medium size (12×15×8 feet, for instance) has a floor, a ceiling and four walls (unless, of course, it's Spinal Tap's motel room just before checkout). Therefore, any sound made in that room will travel to the listener via a direct path and six reflected paths. The original sound will take about 10mS to travel to the listener, and the longest reflection path that represents only one reflection will probably take about 30mS. Therefore, all of those first reflections will get to the listener over a period of about 20mS, from different directions.

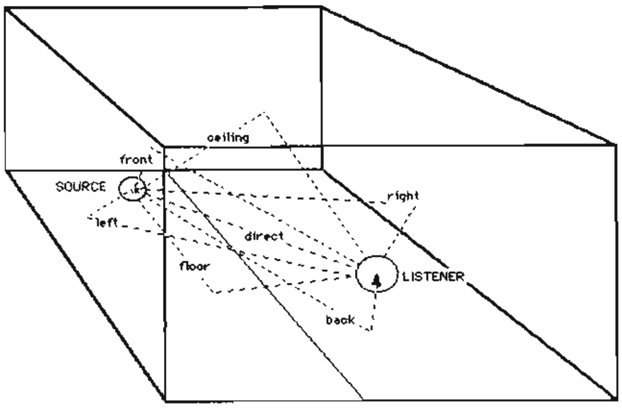


Figure 1. Seven-path model of early reflections in a room. All of these sound-paths are perceived by the listener as part of a single sound, until some of the paths get so long that they begin to be heard as reverb or an echo.

Now then, we don't *perceive* these reflections as such. Our auditory system integrates this family of early delays with the original sound so that what we consciously hear appears to be a single sound that's actually an amalgam of direct sound and early reflections. This compound sound has a much richer timbre and a more interesting character than the single direct sound alone. Musicians have known this for years, which is why playing out of doors (where there are no early reflections) is not very popular.

Multitrack recordings are generally made without the early reflections from the room. This is because we want isolation between instruments so we can control the quality and character of each individual instrument during mixdown. To do this, we mic the instruments closely, and have come to accept the resulting in-your-face, close-up quality of the basic tracks as our preferred sound quality. However, if the truth be known, these sounds are flat and one-dimensional, compared to their true acoustical counterparts. When we hear them with delays added to compensate for the lack of early delays on basic tracks, we know we prefer them this way, without necessarily knowing why.

However, there are two problems. First, when a sound and its delay are recombined, the time of delay becomes the basis for a pitch that is generated as a function of the mix of the delay and original sound. This effect is called "comb-filtering," because the spiky frequency response curve itself looks like the teeth of a comb. Without going into detail, a sound delayed by 1mS (for instance) mixed with its original version will generate a frequency response curve with peaks at 1kHz and all whole number increments of 1kHz (2kHz, 3kHz, 4kHz, etc.) and deep notches at 500Hz, 1.5kHz, 2.5kHz, etc. The response curve will look like this:

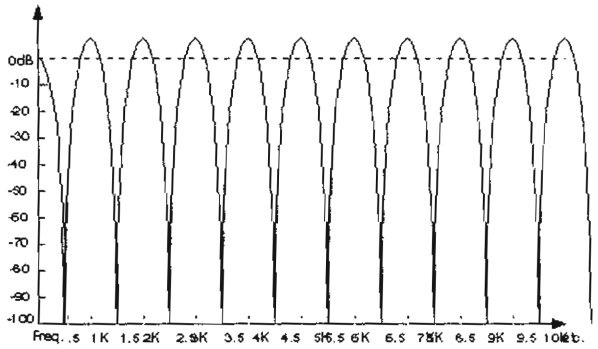
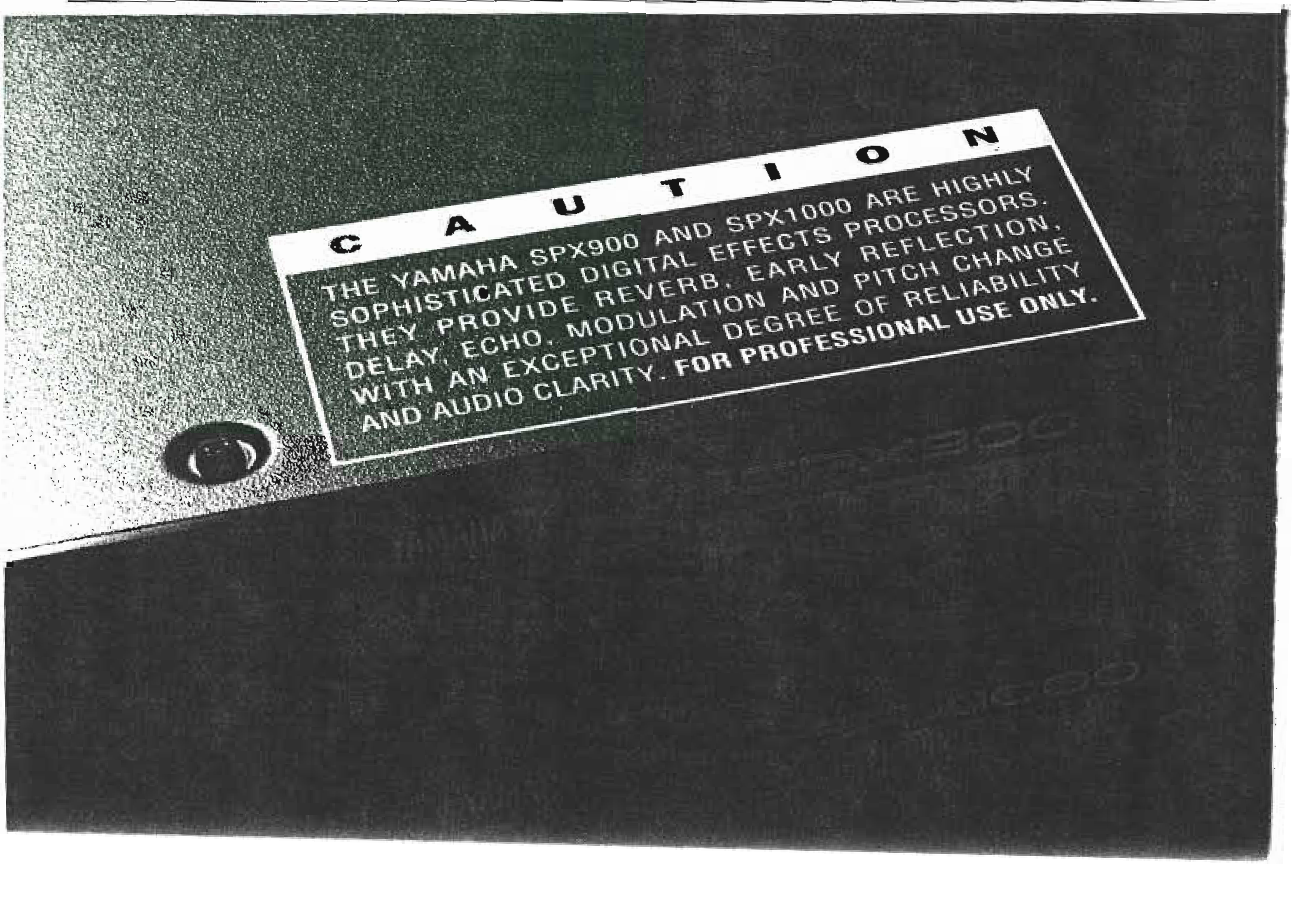


Figure 2. Frequency response curve resulting from the mix of a signal at 0dB and a 1mS delayed iteration of that same signal at 0dB. Peaks will be at +6dB and nulls will be less than -100dB. The graph is drawn with a linear horizontal axis for frequency, as opposed to the more conventional logarithmic or octave-based scale.

This "comb-filter" frequency response is dramatically different from the frequency response of the original signal, which is represented by the dotted line at 0dB. Comb-filtering is clearly audible (!) and, as I said before, has a very clear pitch, which in this case is a pitch whose fundamental is 1kHz. This pitch will impose itself on whatever sound is being delayed and recombined with itself.

Second, we also have to consider our auditory mechanism. When we are working with music, we tend to take this incredible system for granted. Making beautiful and powerful music is very difficult and requires so much concentration that we really don't have the time or mental resource left over to think about what is really going on in our neurons. However, this is a situation where I've found that a little extra knowledge really made a big difference in the quality of my recordings, so listen up!

The auditory system can be thought of as an extremely complex and highly developed local warning system. It is *extremely* sensitive, and it has an elaborate, three-dimensional localizing feature. This system can locate an object in space, and it does this both by: 1) direct



observation of the high-frequency components of the sound at each ear; and by 2) analysis of the assorted room reflections of that sound, including their various delays *and* the above-mentioned angles of arrival. Furthermore, this amazing system also locates and establishes the limits of the room we are in by observation of the phase differences of low-frequency reflections. Such a highly complex and multistage echo-location system works mainly in the pre-conscious realm, and our *conscious* perception of sound is, in fact, a highly developed construct that has comparatively little relationship to the original physical sound pressure waves. The system gives us useful information that we can trust and act on quickly, *without thinking*! The merging of early reflections with the direct sound is a clear example of this: in just 50mS, the brain receives and processes enough data to clearly present to the conscious mind an explicit and uncluttered image of a specific sound coming from a specific point in a given space!

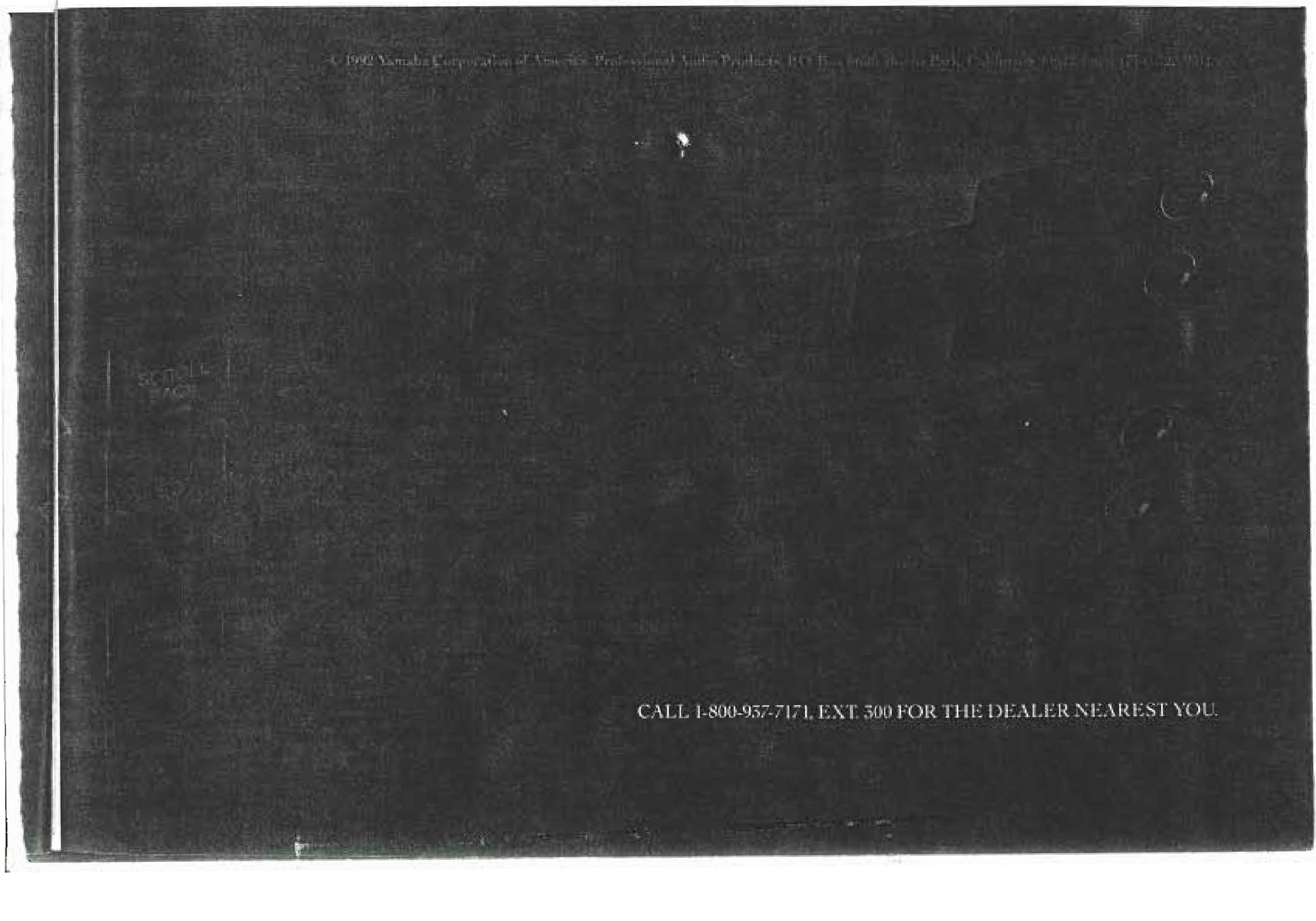
You can set up many rooms or ambiences for different instruments in your mix. Our ears are very tolerant of diverse information.

If we think about the actual information in visual terms (an exercise I always find useful and entertaining when I'm trying to think about these things), the sound and its early reflections are the equivalent of a light bulb in the mirror room in an amusement park fun house. We see dozens of reflected lights that are pretty much indistinguishable from the real one, and are faced with an illusion that is chaotic and confusing. To help us survive, the ear evolved to sort out the audible equivalent of this fun-house chaos and tell us, almost instantly, which sound is the real one.

Now here comes the really weird part. We have a paradox, a conflict in what we know. On the one hand, if we recombine a sound and its delay, we get comb filtering, which really changes the frequency response (and timbre) of the sound, in major-league ways. On the other hand, I've just been telling you (and you've probably noticed this on your own as well) that we don't hear the early reflections as such; they combine with the direct sound to create a richer version of the sound, but without obvious or major damage to the quality of the sound.

How can this be? Is there comb filtering? Yes. Do we hear it? Well, yes and no (I told you this is weird). Here's the scoop: if the original sound *and* its reflection have the *same angle of arrival*, we hear comb filtering and a dramatic change in timbre. If the original sound and its reflection arrive from *different* directions, we barely hear the comb-filtering. There is no appropriate explanation yet as to why this is so. The comb filtering is still there and can be easily observed or measured with a real time analyzer. You can also demonstrate this for yourself if you have a delay line. Take a noise signal and run it to one speaker. Split the signal and run it through the delay and mix it with original signal going to that speaker. Voila! Comb filtering! Big time! Now send the delay to the other speaker instead. Voila again! No comb filtering! The signal sounds like itself. Awesome!

So, the general rule is this: when sound



arrives from a *single direction*, any delays inherent in that sound are considered to be *information about timbre*. When the same sound arrives from *multiple directions*, any delays are considered to be *information about space*. In a real room, where there never is a single sound or even just one reflection, the multiplicity of reflections from different directions reduces the audibility of comb filtering even further. The fact that we have two ears and are therefore observing multiple points in space probably helps as well. So mostly, in the real world, we don't hear the effect of comb-filtering, even though it's there in spades.

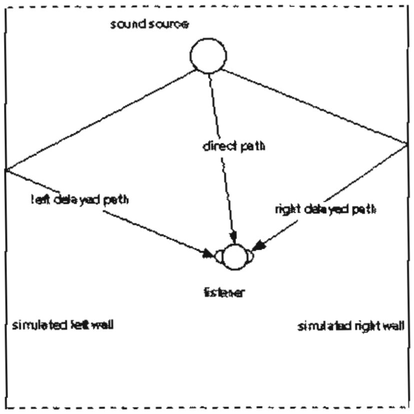


Figure 3. The simulated room for the first patch. The two delayed paths are not quite identical.

Nevertheless, comb filtering is a big problem in a crucial area of audio: the mono compatibility of stereo recordings. When we sum stereo sounds to mono, comb filtering is clearly heard. Multi-directional delays become single-source delays. What was effective and powerful spatial information in the stereophonic mix becomes unintended timbral information in the monaural mix. So, good engineering practice demands a *careful* consideration of the impact of the use of time delays on the monaural mix, particularly if the mix may be broadcast, as monaural summing is a normal feature in the reception of radio and TV. [*There's an article on mono-compatible mixes scheduled for next month.*—NB]

What we've said about early delays can also illustrate the use of effects such as flanging, phasing and chorusing. For all of these effects, three things must happen: first, there must be a delay; second, it must be recombined with the original signal; and third, the delay time must vary. With flanging and phasing, we have a special effect where the frequency of the comb filter is swept up and down by varying the delay time. Delays are very short: a 10mS delay might be a suitable maximum, because it gets the fundamental frequency of the comb filter down to 100Hz (the delay time will equal the period of the comb-filter frequency). The minimum delay should be .05mS or so, which will bring the comb filter's frequency to 20,000Hz (20kHz—the upper limit of audibility). The reason this

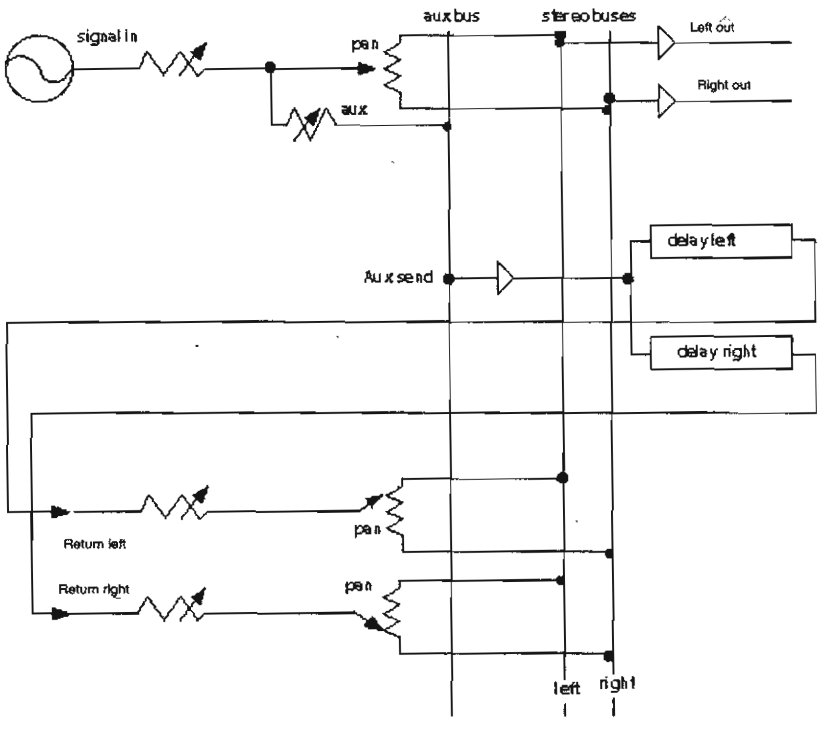
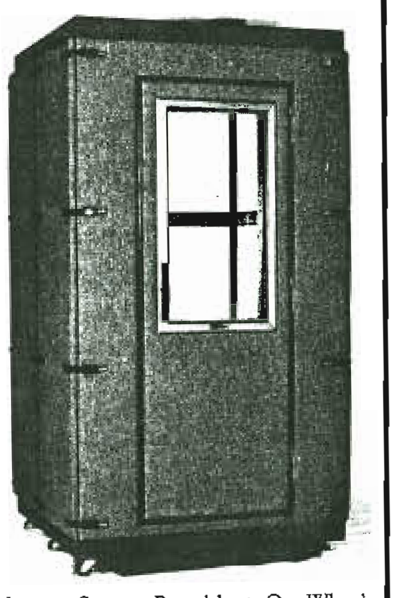


Figure 4. Console flow for left and right reflection simulations. Auxiliary bus sends signal to delay lines that return to Left and Right stereo buses. Left and right delays should be slightly different. Delay lengths should be between 10 and 50mS and usually only a few mS different. Level of delayed signals should be to taste, but typically 6dB below direct signal. Longer delay returns should have higher levels to offset the Haas effect pulling the image toward shorter delay.

MAXELL'S MUSIC MASTERING MACHINES
New Professional MS Master Studio Cassettes


DEMO TO THE MAX!!
•Maxell High Bias XLII Formulation.
•Available in 20,30,60, & 90 Minute Lengths.
•Call or Write for Our A/V Catalog.
Distributed Nationally by:
SONOCRAFT™
THE LEADING DISTRIBUTOR OF MULTI-MEDIA PRESENTATION PRODUCTS
575 Eighth Ave. New York, NY 10018
TEL. 800-274-7666 FAX 212-564-9488

WhisperRoom INC.
MUSICIAN'S ISOLATION ROOMS

Various Sizes ★ Portable ★ On Wheels
Assemble/Disassemble in Minutes
Quiet Ventilation System
Endorsed By: BMG Studios, NY
Tel. (410) 997-1440 ★ Fax (410) 997-1812
9139-F Red Branch Rd. ★ Columbia, MD 21045

SINGERS!
REMOVE VOCALS FROM RECORDS AND CDs!

SING WITH THE WORLD'S BEST BANDS!

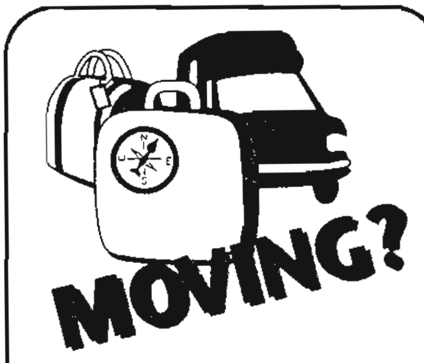
A virtually Unlimited supply of Backgrounds! The **Thompson Vocal Eliminator** can remove most or virtually all of a lead vocal from a standard stereo record or CD and leave most of the background untouched! Record with your voice or perform live with the backgrounds. Used in Professional Performance yet connects easily to a home component stereo. Not an equalizer! We can prove it works over the phone.

This unique product is manufactured and sold exclusively by LT Sound and is not available through dealers. Call or write for Free Brochure and Demo Record.

LT Sound, Dept.H R-9, 7980 LT Parkway Lithonia, GA 30058 (404) 482-4724

For 24 hour Phone Demo Call (404) 482-2485

Manufactured and Sold Exclusively by LT Sound.



MOVING?

If you're moving, remember to take your favorite magazine with you! Notify us of your new address by sending a copy of your current mailing label, and indicating the address changes required, to:

Music Maker Publications
Subscriber Services
P.O. Box 547
Mt. Morris
IL 61054

Please allow 6-8 weeks for the change of address to take effect.

Delay time	Implied room width	Implied room length
3 msec	3 ft.	6 ft.
5 msec	4 ft.	7 ft.
10 msec	7 ft.	12 ft.
15 msec	11 ft.	20 ft.
20 msec	17 ft.	45 ft.
30 msec	23 ft.	60 ft.
50 msec	36 ft.	80 ft.

Table 1. Approximate virtual room sizes implied for various delays in patch shown in Figures 4 and 5. A 3mS delay implies a tiny room, while a 50mS delay suggests a fairly large club.

effect is such a psychic bombshell is that it sonically mimics a situation where *the wall is moving*! Psychologists call this *cognitive dissonance*—our various senses give us conflicting information. It's a really useful tool for creating profoundly entertaining effects. In this case, while our eyes tell us that everything is stationary, our ears inform us that the environment is moving! Yeeha!!

Chorusing mimics something else: the sound of multiple voices. When you hear a sound and its room reflection, the amount of delay (and phase shift) remains constant unless: a) you move; b) the source moves; c) the wall moves; or d) you've ingested too much single malt scotch (that's a joke—however, I've tested it and it is true!). This is because there's a fixed relationship between you, the source and the wall. The fixed quality of that relationship results in a pattern of delays that also have fixed phase relationships. When two or more voices sound in unison, however,

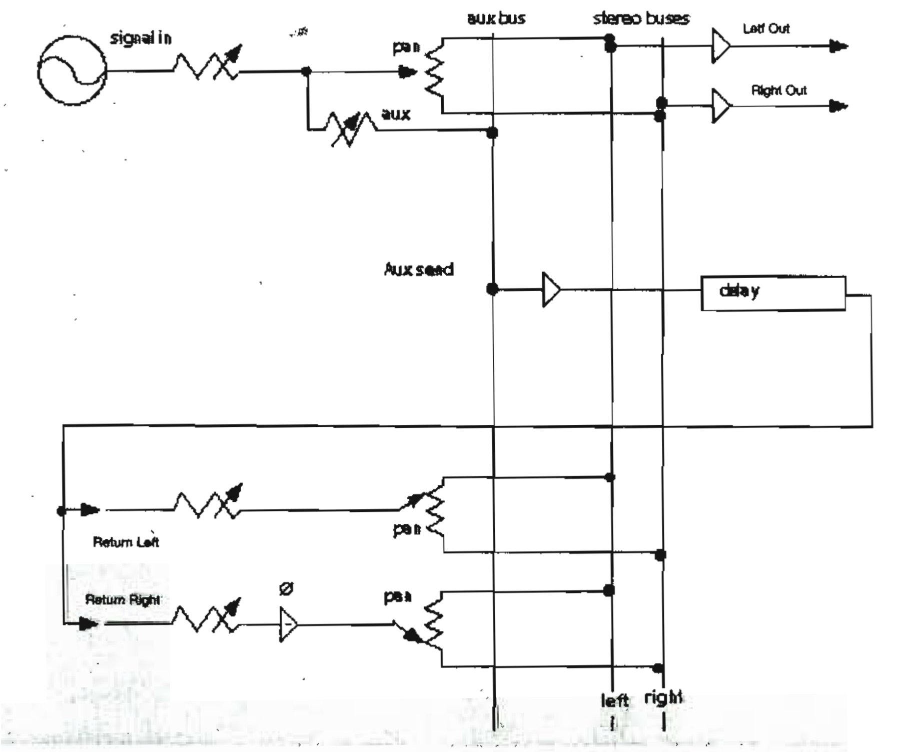


Figure 5. Alternate console flow for left and right reflection simulations. Multitrack equivalent of Middle-Side stereo microphone setup. Auxiliary bus send signal goes to single delay line that returns to both Left and Right stereo buses. Right channel return has signal inverted (polarity reversed). Left and Right return levels should be identical. Delay lengths should be between 10 and 50mS, and usually only a few mS different. Level of delayed signals should be to taste, but typically 6dB below direct signal.

While our eyes tell us that everything is stationary, our ears inform us that the environment is moving! Yeeha!!.

they cannot vibrate in a "phase-locked" way. If the voices are oscillators, each with its own stable frequency, they will "beat" with each other with a constant "beat" rate. If they are musicians, they are constantly correcting and trying to hold exactly the same frequency, so they tend to drift around each other, always holding very close to the unison but never at *exactly* the same frequency. When a whole group of voices does this, it results in a very rich, vibrant and dynamically "happening" sound. The chorusing device simply alters the delay time randomly, and within a few mS. Additional delays will mimic the sound of additional voices, and yield a richer chorusing effect.

It is possible to use a couple of delays, or even a single one, to effectively create the early delays to simulate a room around your multitrack sounds. The trick for using these effectively is to "think like a room." Figure 3 is an illustration of an easily simulated room with side reflections only.

In this case, the room has only two walls, left and right, and they are almost equidistant on either side of the direct track. Below are two basic patches for simulating this room. Take a look at the first patch.

In this case, the direct sound (there can be many direct sounds, mixed, of course) is sent in mono to both stereo busses, and is also sent via the auxiliary bus to two separate delay lines. I typically start with delays of 30 and 35mS, but this is really an issue of taste—your personal fantasy about the size of the room in which you imagine the music is playing, and the style and mood of the material you are recording.

You can also use equalization to enhance the character of the delays and implied ambience: high-frequency boost will add a hard, bright edge to the ambience, suggesting hard, reflective surfaces with no diffusion, and low-frequency boost will add a sense of spaciousness and envelopment.

In any case, the delays should be almost the same, but not quite, and the level of the longer delay should be greater, to offset the Haas effect, which will tend to pull the image toward the channel with the shorter delay. The overall level of the delayed sounds should be 6 to 10dB below the direct sound, for three reasons. The first is simply that in reality (if we "think like a room"), reflections, which have traveled further and attenuated more, will typically be 6 to 10dB lower in level than direct sounds. Second, the effect can be a little overwhelming if it is over- done. Finally, the comb filtering problems will be reduced as the *difference* between levels increases, which means that you are going to run into less trouble when you sum the mix to mono. I find that this patch works really well on the lead vocal, maybe all the vocals, the lead instrument, and the electric bass. It gives a richness and presence to these lead parts that is quite compelling. I also love to "wake up" synth sounds with this kind of treatment. They usually get an awful lot better with a couple of judicious delays like this.

Figure 5 shows a variation on this patch that only requires one delay line.

This patch mimics the Middle-Side stereo microphone placement, where the delay line serves the same function as the side microphone. It has a couple of advantages, and one big disadvantage that causes me to not use it very much. The advantages are that it requires only one delay line, and that it is absolutely mono-compatible, which is also its big disadvantage. When the stereo signal is summed to mono, the left and right returns (simple polarity reversals) will cancel each other out, so the effect of the delays is completely removed in mono, which usually really changes the musical character and impact of the mix. On the other hand, the patch is very effective, and, because there are no time differences between left and right buses, balancing it is a cinch.

Naturally, if you have more delay lines, you can set up many rooms or ambiences for different instruments in your mix. Our ears are very tolerant of diverse information, and will cheerfully accept a whole bunch of overlaid ambiences and reverbs without much trouble or confusion. You can also mix dry and wet sounds simultaneously without much trouble, and often with great effectiveness. This is an area where there is tremendous potential, and a great deal of experimentation is called for.

Have fun, but don't delay too long (that's another joke)!

Dave Moulton chairs the Music Production and Engineering Department at Berklee College of Music, and is typically a little late with deadlines. Peter Alhadeff procrastinates on his assignments in the department, and Alex Hodge usually echoes his ideas.

WANT TO CONTROL DUST IN YOUR STUDIO/LAB?

DUST

We do! We can custom make nylon/vinyl covers for all your equipment. Keep it clean and looking new. All covers are designed for cable cut-outs and available in black, blue, burgundy or gray.

BUT WAIT...WHAT ABOUT SMOKE AND DUST WHILE I WORK?

True, air-borne particules are constantly settling as you work. Now you can control your indoor environment with a STUDIO AIR FILTRATION SYSTEM. Our sytems are:
- Portable
- 99% effective towards all irritants
- Operates at pennies a day
- Hepa and Electronic Cell Filtration

Give us a call and ask us about all the advantages of an Air Filtration System.
Individual Brochures available upon request

1(800)228-DUST
Mastercard & Visa Accepted

Le Cover™
The Dust Control Company

1 N. 353 Bloomingdale Rd. • Carol Stream, Il. • 60188

Serving San Fernando Valley Since 1986

ABA

CASSETTE DUPLICATING

Real Time and High Speed Cassette Duplication

- Chrome Tape
- Printed Labels
- Speedy Turnaround
- Quantity Discounts

WE TEST EVERY CASSETTE

FOR USER-FRIENDLY SERVICE

Call:

818-222-1071

fax 818-222-1072