Listening Rooms
Immersive vs. Stereo vs. Studio

James D. (jj) Johnston

Former Chief Scientist

Immersion Networks

Some considerations in listening rooms

- A listening room is for the final delivery of sound (including home theatre, etc)
- In stereo, you have only a front soundstage created physically.
  - Do you live with that?
  - Do you want to add something "behind" you via diffusion or the like?
  - Do you want to use some kind of stereo speaker system to create a better illusion?
  - Do you want to emphasize the details in the content?
  - Do you want to produce content in that room?
  - Does the room have to be a "critical listening space?"
  - Does it have to be, instead, a space to reproduce all recordings, flawed or not?

The usual outcome

- For pure listening, the room that provides a pleasing sense of space, including some side reflections (necessary for panpotted material) and a rear diffuse reflection is often preferred.
  - This is often called a Live End/Dead End (LEDE) setup.
  - The time domain characteristics of such rooms can hide some artifacts. This is not necessarily thought to be a flaw.
  - Some of us find the masking of time domain details annoying.

- For production or recording, opinions vary:
  - Some argue to experience the production "like the user will do"
    - Some recording or production artifacts may be hidden as a result.
  - Some prefer a critical space to listen for problems in the recording.
    - Such spaces expose flaws that an LEDE space can mask completely.
    - Those who prefer this solution may not know how this sounds in an LEDE room.

LEDE

When live end-dead end rooms were invented, much of the emphasis in both mono and stereo reproduction was directed at "people playing on stage," recordings were mostly live, and the goal was to reproduce something like a live event.

Later, the idea of a live event was less important, multitracking, pinpoint editing, and the like were developed in order to make audio the way the artists or producer wanted. The idea of a "live space" went to the wayside, at least to some extent. For this kind of pinpoint accuracy, a less diffuse space was often helpful.

This left LEDE rooms in conflict with some kinds of music, but still very desirable for other kinds of music.

Now, we move to multichannel music

- There have been a few approaches, for instance:
  - 3-channel front presentation (1933)
  - Quad (1960s, 1970s)
  - 5 channel (3 front, 2 side)
  - 5 channel (3 front, 2 back)
- And finally
  - 7 Channel,(L, R, C, LS, RS, LR, RR) that provides enough auditory cues to pan around the room.

  - I've left out theatre designs like 5 front, 2 side, 6 channel with a single rear, and a few other methods.

All of those formats were presented in a variety of fashions that did not necessarily encourage "immersive" sensation.

Production methods were often very "theatrical," with SFX in the side and back, dialog center front, and main content in Left and Right. (This was not necessary for the more advanced formats, but the production systems by and large still take this approach).

Immersive Systems

- 7.x.0 (in the plane) systems are the first to provide more capability.
- 7.x.4 (4 "height" speakers)
- 7.x.2 ("proscenium speakers")

- The industry has mostly settled on 7.x.4, or sometimes 7.x.6 formats. For audio via loudspeakers, further progress is most likely in diffuse vs. direct radiating systems that have occasionally been demonstrated a few times, but never adopted. There is a hint of diffusion in the ".6" systems, but I can't say more at this time.

- For those wondering about the "x," that's for subwoofers, OR for low frequency effects. More on that later.

So, Immersive Sound, what is it?

- It means that with proper production and implementation, you can place the listener in a space that is unrelated to the actual listening space (or in headphones, but we're talking about listening spaces today).

- That, in a nutshell, is what building an immersive room is all about.

What would the perfect immersive room be?

- It would, frankly, be a close to anechoic space, with 'n' identical speakers (for all main channels, "up" included). In this fashion, you can simply create, simulate, or whatever else you choose to do in this otherwise "does not exist" space.

- In order to make inhabiting this space comfortable, and not "break the illusion" the room would also have to have microphones and processing to overlay (with echo cancellation) the sounds made IN the room sound like they are in the desired immersive space.

Who's going to do that?

Nobody

This is not a useful approach

Bear in mind we are talking about small venues.

Large theatres are a different, more complex issue that can be stunningly good if you are richer than Croesus. But when you do that, if you're doing live, BEWARE THE LATENCY, including that from the 1115 fps speed of sound. Physics is not your friend here.

A few definitions:

- Direct sound: Sound representing a source, that arrives before anything else, i.e. the sound directly from a source. (Occlusion makes this interesting, until you realize you're just confusing which is the direct source as far as the listener is concerned.)

- Early reflections: Sound that arrives soon after the direct sound from a given source. This can provide important information to the perception of the source, in terms of width, etc.

- Indirect sound: Sound coming after the "early reflections," most often, but not always, reverberant. As well as reverberation, things like delayed echoes (longer than 40ms give or take) can fit in this category, but are, of course, better treated as filtered, delayed direct sources, depending on the question of "is the echo from a point, or is it diffuse?" That's a question of intent. The producer gets to have the intent.

What do we need to do?

- First we must be able to overlay the playback room with the intended environment of the immersive content.

- Second, we must avoid adding too much "coloration" to the playback.

- Third, the playback room can't be inordinately live, so it doesn't overhang the intended environment.

Overlaying the playback room

- The signal from the intended environment must arrive before any playback room reflections.
  - Yes, this means both the "direct sound" that provides direction cues and some distance cues, and the "indirect sound" added to create the desired environment.
  - No, it does not mean "bounce the indirect sound off the room wall." While this might work sometimes, it can just as well go bad depending on the intended environment.
  - For once, the "precedence effect" is your friend.

- The signals from all speakers must arrive AT THE SAME TIME to the central listening position.
- The GAIN to the central listening position must match from all channels.

More on the overlay issue

- Any "early reflections" in the intended environment must be preserved.
  - That means that they should be treated like "direct sound" as far as the playback system.
  - You DO NOT WANT to design in early reflections from the playback room. This will confuse things in unpredictable ways.
    - No, this does not disrupt direction perception IF THE IMMERSIVE MATERIAL IS PROPERLY PRODUCED, using both time and gain rules to set the intended direction. Notice my failure to use the word "PAN" here, thank you kindly.
    - This quite surprisingly means for a 7-channel horizontal mix, you never EVER have signal arriving from only one speaker. Never. Ever.

Delay and gain compensation for speakers and speaker locations.

- This is not a good idea. It is **the LAW!**

- This is not ONLY for direct sound, a good simulation of an intended space may need to depend on this, especially the gain balance.

- Remember, this is important, because:
  - You are not just creating "images" between two speakers.
  - You are creating images between all 7 speakers, as well as diffuse sensation that doesn't "come from any speaker."
  - Lack of balance will cause unpleasant results.

- Balance your system. ALL OF IT!!!!!!

Coloration issues

- This is listed as a second issue. Most any room will have some playback coloration, however:
  - If it's the same in all directions, you're probably going to be ok.
  - If it's not too extreme, you're probably going to be ok.
  - The issue is timbre from different directions, and less on overall neutrality.
  - Giant bass modes, however, will mess things up emphatically. More on bass, subwoofers, and LFE later. There's a whole set of issues and debates on that, wherein the issues from physical acoustics conflict with the psychoacoustic issues, and both conflict with many 'subwoofer' systems.

- All in all, moderate coloration, if the same in all directions, is "ok." Of course it could be better, but this is the same thing as "balance" between speakers, the balance is much more important.

Room too live

- It is very easy, using even 1950's equipment, to create a sensation of greater distance in the intended environment, EVEN IN MONO (in fact in some cases easier in mono than stereo).
- It is *(&(& hard, bordering on impossible, to move a source CLOSER into a playback room if it's inordinately live.
  - Yes, you can do that. I've made array speakers that do that. For that, not only do you need semantic (location) information, you have to adapt that to the individual hardware.
  - You can create a decent sense of that by ensuring that the arrival from the playback room is:
    1. Much later than the direct from the speakers
    2. Rather lower in intensity than the direct from the speakers (r^2 is your friend here).
- A really live playback room is not going to do it.

# There is, however, an interesting confluence here:

- While you'd like a production room to maybe be quieter than an end-point listening room, both have the same goals.

1. Balanced system (gain, delay, timbre)
2. Clear first arrival of the speaker signal at the listening position
3. Of course, properly located loudspeakers. (With a decent semantic mix, you can tell the playback process where your speakers are within a fairly wide range, but GET IT RIGHT!)
4. No blatant modes or long-duration reverberation in the room.

What did we do for a production room?

- It's quiet. This is always nice.
- It's dead, dead, dead. Not anechoic, but other than reflections from the speaker fronts, it's close. Even with 11 speakers in the room, the dominant reflection from the central position is the foam-covered entrance door.
- Of course everything is time aligned to a single sample.
- Likewise, even though we use 2 kinds of loudspeakers, they are both within 1dB from 40 to 18000, and the mains (7 horizontal) are 1dB from 20 to <stupid, way above 20K>
- No, you're not ever going to do that at home. (Well I might, some day, if I become inordinately rich and build a new house, but don't hold your breath!)
- Note: This room has discovered some profoundly annoying problems with almost all immersive production we've found on the market in terms of signal processing errors, noise, and other odd artifacts. No, we don't even try to use "encoded" channels, yeah, you can hear that alright. But that's why we built it that way!

For a listening room:

- Fix any disastrous bass modes or "ringing" in the room.
- Equalize delay, gain and overall timbre.
- Move your speakers at least 2-3 milliseconds in from any flat surface.
  - That's for "direct radiators."
  - Things like planar and other dipole, bipole, or omni speakers are not well suited for immersive reproduction. What they do very well is making your listening room sound good. But here, remember, we do not WANT your listening room's sound, rather the intended production's sound.
  - Make sure any channel sounds like any other.
    - A good test for this is to pairwise channels, and present a stereo signal you know well to two channels (any two). You should get a good stereo image. If not, something is wrong
    - Once you've done that, try a good immersive mix. You may be impressed.
- Note: This does not create a critical space. In my experience, this often works to your advantage, I'm sorry to say.

Pointing into ear canal

Behind Pinna Shadow

L   C   R

LS   RS

LR   RR
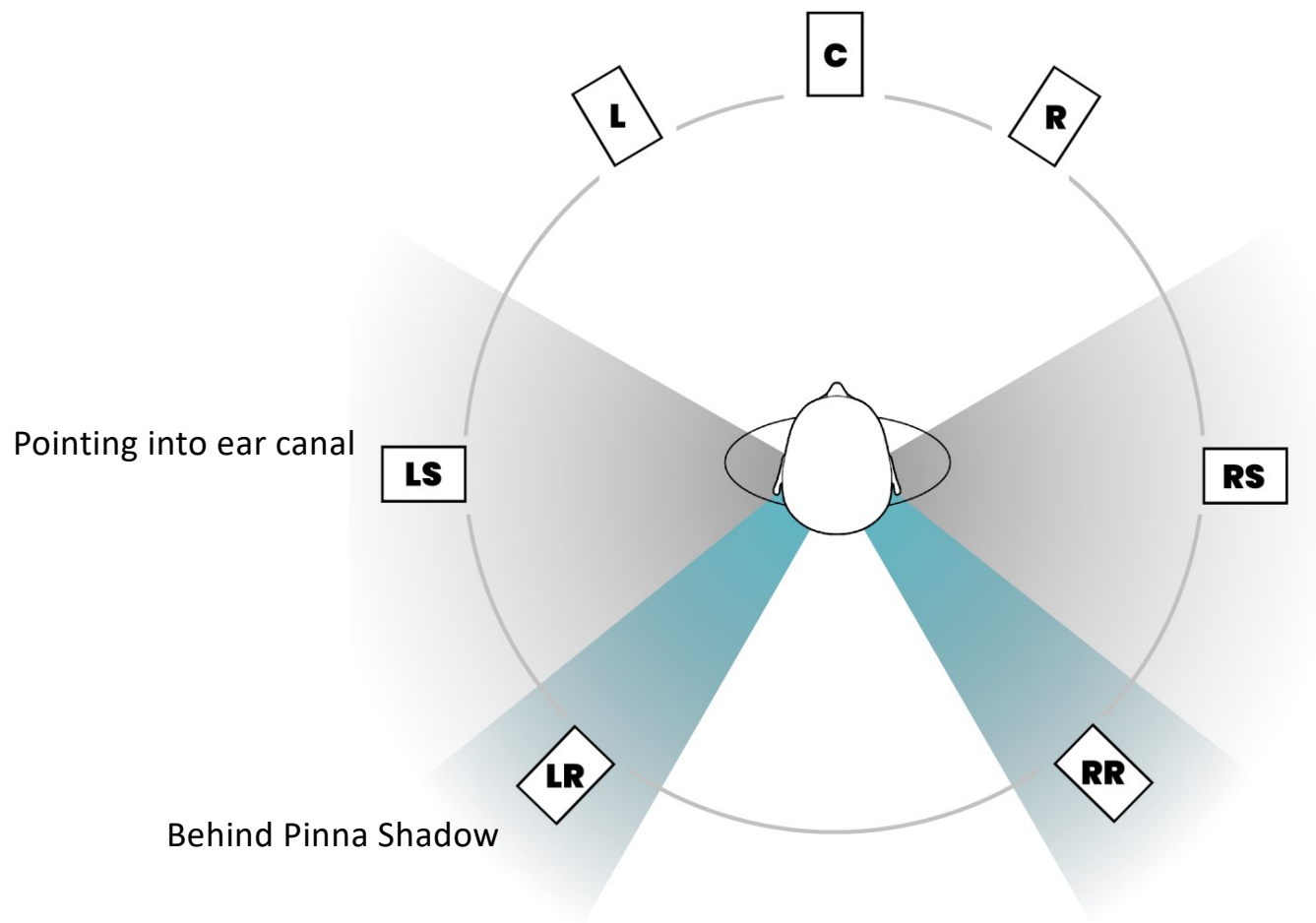
Tricks to fix issues:

- Fire-rated absorbers for annoying walls, ceilings, etc.
- No vertical space? Maybe better to virtualize the top speakers. Yes, it's workable. Unfortunately not in current consumer equipment, but I do expect that may pass with time.
- Again, distance from walls. You may find it advantageous to move symmetric pairs (referring to dead-front as zero point) in and out from a wall, remembering that you separately control the delay to each speaker. There are room systems that will do that for you.

- Which brings us to: (DRUM ROLL PLEASE)

Subwoofers

The role of the "subwoofer"

- Based on correct, proper research it is shown that there is little if any directional sensation below 90Hz.
  - Yeah, I'm not aware that's disproven, and I suspect it may never be.
  - It's probably right.
- So, many "minispeakers" have a transition in the 90 to 130Hz range. (Yeah, that 130 is a mistake, you'd never expect that, right?) This is done because it "doesn't matter." (Yes, those are quotes of contempt, boys and women.)
  - **WRONG. JUST WRONG!**

A simple experiment:

- Take a 50Hz sine wave in headphones. Of course, headphones that can do that.
- Set it to comfortable, but non-distorting level.
  - Note, this will be more like a sensation than any tonal feeling
  - If you create harmonics you'll mess this up!
- Slowly shift the phase of the tone in one ear.
  - Oh look, it changes the spatial perception.
  - No idea where it's from, but the "width" and "sense of space" change quite surprisingly.

Yeah, so what?

- Simply put, in-phase bass at all frequencies sounds like a small, pressure-driven room.

- If you want "wide" and "spatial" you need to get those mains down to a minimum of 40Hz. Now with some of the "mono the bass" recordings, which is often still done these days quite deliberately, you won't hear any difference. You want either a natural spatial recording (meaning with either near-coincident <I suggest a .9 millisecond diameter there) or widely spaced miking, and then you'll see what it should sound like.

- The outcome:
    If you can't do that, better to cut the bass, and you'll mostly preserve the spatial effects. At least, you won't have two horribly conflicting sensations that really mess things up once you learn to hear this.

The LFE channel(s)

- Ok, that's fine. "E" stands for effects.

- For use as bomb blasts, tornadoes, spaceship explosions, and the like, yeah, that's good.

- Go ahead. But stick to effects.

  NOTE: Even the effects would be better in at least 4 channel with the woofers in a square arrangement, but ok. Better "the planet blew up" than "the planet rattled a bit and there was a big white flash."

Some Materials for Absorbers (not diffusors!)

- Owens-Corning "soundsoak"
- GIK Acoustic
- Acoustimac

- ALWAYS CHECK THE FIRE RATING AND DO NOT EVEN THINK ABOUT "egg crate". FIRE RATED. ALWAYS! CHECK YOUR LOCAL CODES OR HIRE SOMEONE WHO KNOWS THEM!!!!!