

43 Latency

9 April 2021

Latency and Delay

I'm Doug Fearn and this is My Take On Music Recording

Many years ago, I was having a conversation with my friend Bill, which we did every week for about 30 years. This was via Amateur radio, using Morse code.

My station was set up so that I could hear what was happening on the frequency all the time. Functionally, this meant that my own sending was just a somewhat louder signal in my headphones than the other signals nearby.

This technique is handy because it allows the operators to interrupt each other, perhaps to ask for a repeat of a word if it was obliterated by a static crash or a deep fade.

I liked using this technology, called "break-in," because of its utility, but also because it made me feel part of the environment. I could hear if static crashes from thunderstorms were getting louder, or if someone else started sending on the same frequency, potentially disrupting my conversation with Bill.

On this night, the frequency was quiet and there were not many other signals near our frequency. But I became aware of a weak signal, right on our frequency. It was gradually growing stronger. I stopped sending to listen. The signal stopped, too, after a short pause. I resumed sending. So did the mystery signal.

The signal was not only weak, but it sounded like it was underwater. I estimated the delay to be at least a hundred milliseconds.

By this time, I was pretty sure that what I heard was my own signal, after traveling all the way around the world. Yes, this is possible.

All radio waves travel in a line-of-sight path, which means that you would expect every signal to fly off the Earth, into space.

But the Earth has a layer of charged particles, called the ionosphere, about 200 miles above us. And these charged particles, or ions, can reflect radio signal if it's in a certain frequency range. The shortwave frequencies, from 3 to 30 MHz, are the best frequencies for reflection off the ionosphere and back down to earth. Well, it's more complicated than that, but that scenario is fine for this example.

The wonderful thing about the ionosphere is that it is always changing. It varies significantly between daytime and night. It varies with the seasons. And it is profoundly affected by the number of sunspots, which peak and fall on an 11-year cycle.

My 100-Watt signal traveled all the way around the world by way of reflection from the ionosphere, down to earth, where it bounced back up to the ionosphere again. This is what makes long-distance communications possible.

My estimate of the number of “hops” my signal made that night is roughly 8. Eight times my little signal encountered the ionosphere, was reflected back to earth, reflected off the ground or sea, and back to the ionosphere.

The loss over this path was well over 100dB, so it took exceptionally quiet conditions to make it possible for me to hear my own signal after traveling this path.

Radio waves, like all electromagnetic radiation, travel at the speed of light: 186,000 miles per second. That’s 671 million miles per hour. You would think that at that speed, light and radio waves would move so fast as to be virtually instantaneous. But it’s not really instantaneous, just very fast. And at significant distance, the delay is perceptible to humans.

My best-guess estimate now of the time delay for this trip is about 140 milliseconds. That’s roughly comparable to the time delay in a tape machine running at 15ips, which I was very familiar with and could easily mentally compare it to my signal delay.

Hearing my own signal after it circled the Earth was an exciting experience, and as I thought about it later, I tried to picture where my signal landed in its multiple reflections from the earth. Truth is, it could have gone via many different paths, in any direction. And it probably did travel by multiple paths, resulting in a smearing of the arrival time of the signal I heard coming back.

Interestingly, Bill did not hear the echo of my signal nor of his. I think my recording and musical experience made my hearing much more sensitive to detecting this.

This is an example of a time delay. You experience this all the time in your daily life, possibly in some cases without being aware of it.

I want to talk about this phenomenon as both a useful artifact of physics and as a potential source of harm to our recordings.

Here’s another example, this time with signal delay in wires.

When I was working at WPEN in Philadelphia early in my career, the station often provided a remote base for NBC reporters to send their reports to the NBC Radio Network. This happened when there was a major news story happening in Philadelphia, or a big sports event.

We used one of several production studios to feed the network. Often, it was simply recorded back at NBC in New York for later broadcast. But sometimes we would feed the entire network directly, live.

Back then, radio networks relied on leased telephone lines to distribute their programming. These were ordinary copper wires, but permanently dedicated to a fixed path, and equalized and amplified to provide quality audio.

In a vacuum, or through our atmosphere, electromagnetic signals travel at the speed of light. But in wire, they are slowed down somewhat, although it’s still extremely fast. The speed is still around 75-85% of the speed of light, or about 150,000 miles per second.

WPEN would send the news report via our own dedicated line to the network headquarters, a distance of about 85 miles. From there, the audio went to all the affiliated stations. First, it went to the West

coast, then across the South, and back north. Philadelphia was actually at the end of this long stretch of wire. It always amazed me how good the network audio sounded, traveling that far and through who knows how many amplifiers and equalizers.

All the studios at WPEN you could chose the signal in the monitor speakers from a variety of sources, including off the air so you could hear it as it sounded to the listener. The other common selection was the console output, which you needed to use for any internal production, or when the studio was operating on a tape delay of several seconds, during a call-in show.

It scared me when I would get assigned to feed the network live. Millions of people would hear anything I did wrong. It was a lot of pressure.

The first time I had to do this was for a big sports event happening in town, and it was with Joe Garagiola, former baseball great and later sports broadcaster. A nice guy, easy to work with.

We were in Studio F, one of the production studios. I was in the control room and Joe was in the small studio in front of me. I was waiting for the cue to start the feed to the network, which required monitoring our off-the-air signal. That air signal was also in Joe's headphones.

It never occurred to me that there would be a significant delay from what we were sending from the production studio to NBC, and then around the country before being received in WPEN's Master Control Room, down the hall, and from there it went to the WPEN transmitter. Our air monitor source was essentially from an AM broadcast receiver.

Joe started talking, and immediately stopped. His delayed voice was coming back in his headphones after this journey through several thousand miles of telephone lines. Fortunately, I realized the problem right away and switched to the console feed for my monitor and Joe's headphones.

The delay was probably around 50mS. It's nearly impossible to talk with that kind of delay in your headphones.

Lesson learned. If you put audio through enough wire, it will take a perceptible amount of time to get to the other end.

We experience this kind of delay all the time, if you watch TV news. Today, phone lines are rarely used for TV broadcasting. The signal is relayed over almost any distance through geosynchronous satellites, that remain in a fixed position above a point on the Earth. They do this by orbiting the Earth at a distance of about 22,000 miles. Even local TV newscasts use satellites most of the time because it is a more reliable than using a point-to-point radio link back to the TV studio.

The satellite path introduces significant delay in the signal. The path up and back might be as long as 50,000 miles, depending on how far apart the locations are on Earth. That's around 250mS. But add to that the digital latency at four points in the path. One is at the ground station, which encodes the signal in a bit-reduced format to save bandwidth, like a MP3. The next two are in the satellite, as it receives the signal and decodes it, processes it, and then sends it back down. And lastly, the decode process in the receiving equipment on the ground. I have no idea how long these encode-decode processes take, but it is significant. My guess is that the total path time is around one second.

If the reporter is interacting with the host at the TV studio, the conversation becomes filled with long pauses, since now it takes time for the question to reach the reporter, and more time for the reporter's response to get back to the studio host. That doubles the delay, resulting in a couple of seconds between question and answer. That does not seem like much, but I can tell you from editing a conversation that even 10 or 20mS off can make a conversation seem slightly out of sync.

The cell phone system is another example of latency. It can sometimes be difficult to carry on a comfortable conversation because the callers keep stepping on each other, especially if one tries to interrupt the other. If you are on a call and the person on the other end is in the same room as you, you can hear the latency. I estimate it is usually about 500mS, or half a second. That seems like a small delay, but it makes having a normal conversation difficult.

This delay is mostly from the time it takes to convert your voice from analog to digital, bit-reduce that digital stream so it takes minimal bandwidth, and then reverse the process at the other end. For local calls, the delay caused by the radio links to and from your phone is negligible, but can become significant on a long-distance call. If you think a MP3 is highly bit-reduced, cell phone audio is just a shadow of the original audio.

You'll experience the same kind of delays with on-line meeting services like Zoom or Skype. The delay will be a combination of the digital encoding, decoding, and the time it takes for the signal to travel from one point to the other.

As an aside, people on their phones, or at a Zoom meeting, tend to talk way too loud because there is no sidetone in their ear. Sidetone was a system developed by the Bell Telephone Company over a century ago. It fed a portion of your outgoing audio into your earpiece, which made a conversation feel much more natural. I don't know for sure, but I suspect this was done to prevent people from shouting into the telephone and causing crosstalk, which creates interference with other phone calls.

This probably isn't practical on a cell system, due to power demand, and probably latency would also enter into it. A delayed signal into your ear would not be tolerated. It's a shame, because it would cut down on the noise of people talking loudly into their phones in a public space.

Humans are very good at perceiving very small delays, and those delays cause us discomfort, whether we are consciously aware of it or not.

So far, I've been using delay and latency somewhat interchangeably, but in my mind, I divide this effect into delay, which could be useful or annoying, and latency, which is time delay I categorize as a delay in a digital path. Latency is almost always bad.

Another common example of latency is when the audio and video are out of sync.

The video and audio are usually using a bit-reduction compression technology. The time to encode the audio may be significantly different than the video encode time. It throws the two off. Usually this is corrected down the line, but not always. In some compression schemes, the delay varies by the content of the visual portion, so it is not a fixed offset. That's annoying. At least to me. Fortunately, the technology has improved and this problem is far less common than it was a few years ago.

The annoyance factor of out-of-sync audio depends on whether the audio leads or lags the video. If the audio is slightly behind the video, our brains seem to accept this as credible, because we are accustomed to that kind of delay in the real world. Light travels at 671,000,000 miles per hour, while sound goes about 761 MPH. The speed of sound varies somewhat with altitude and other factors, but you get the idea.

Another way of looking at the speed of sound is in feet per second, or, more usefully, in feet per millisecond. Sound travels at about 1100 feet per second, and that's close enough to 1mS per foot for a rule of thumb.

So, a reflection off a wall 50 feet away is delayed roughly 100mS.

A common experience (in most of the world) is thunder and lightning. Unless the lightning hits right next to you, you will always see the flash before hearing the thunder. You can use that delay to determine how far away the lightning is. This can be misleading, however, because most lightning is in the clouds and does not hit the ground. I think of the delay as how far above the lightning is. And you may notice that the sound of lightning hitting the ground is different from lightning within the clouds. Think about the geometry of how the sound reaches you under different lightning scenarios.

This kind of delay is noticeable to some of us even when someone is talking to you from a distance. The delay is short, but noticeable if you are paying attention. The sound always comes slightly after we see the person's lips moving, and our brains accept this as normal

But what if the audio comes before the video? Our brains do not like that one bit. And this happens a lot because the video encoding often takes a lot longer than the audio encoding.

This is another example of subconscious discomfort in most people.

OK, so if electronic signals travel at the speed of light, or slightly slower in wires, why do computers have so much latency?

Back in 1979, with my new Apple II computer, I was amazed that there could be any perceivable latency. What was going on inside this little box that could take any amount of time? When you typed something, how could the letters take a perceptible amount of time to appear on the screen? It was inconceivable to me that this would take so long as to be noticeable.

Our computers are now millions of times faster than my old Apple II, but the latency is still there. As I type this script, I can detect the delay from when I press the key and when the letter appears. And this is with a fast computer running Word, a bloated but useful word processor. How can it take any time at all?

Well, a computer scientist could explain all the steps your keystroke goes through, starting with a polling signal that constantly sweeps through the keys looking to see if one is pressed. It does not take too much time to store that byte of data in memory, but it isn't instantaneous. Retrieving it takes time, too. And so does the processing time to check my spelling, and figure out how to format the text. Even the computer screen takes a finite amount of time to convert the incoming data into the proper pixels to turn on or off. It all adds up.

Add to that your computer's other tasks, like background checks with software providers, and you can understand why there is latency in a general purpose computer.

I can type pretty fast, and I watch the screen while I am typing. I've gotten used to the latency, but when I have to backspace to correct something, I almost always go back too far, because of the latency on the screen.

So, what does this have to do with recording?

Well, latency can ruin your record.

In the old days of analog, a headphone feed for the musicians was virtually instantaneous. They heard the note they just struck just like they would hear it live.

But when digital came along, latency became immediately apparent. For someone trying to do an overdub, the delay from when they hit the note and when they heard it could ruin the performance.

I don't know what the maximum allowable delay that most musicians can tolerate would be, but I know players who can get thrown off by just a few mS of latency.

The entire feel of a rhythm track can change with just a few mS change in timing.

Blow up the waveform on your screen and compare say a drum track and a bass or guitar track. Note how little they have to be off to sound like they are not tight. It is incredible how good our hearing is at detecting these timing errors. In some music, a delay of just a few of mS could make the difference between a song that makes you want to dance and one you want to turn off.

Even with modern software and the fastest hardware, you can't get a delay shorter than a few mS.

I think some musicians, especially those that have had no experience with analog recording, have accepted the latency and it does not consciously bother them. But I sincerely believe that digital latency adversely affects a player's performance, whether they know it or not.

And yes, you can quantize every note of the performance, but I find this unsatisfactory. I prefer to hear real musicians playing in the time they feel the music demands, which needs to breathe on many songs.

The solution to latency is to give them a headphone feed that is derived before the signal hits any digital device. That's simple if you have an analog console, but it gets complex otherwise.

Some converters have direct outputs, with little or no latency, and this can be used to feed the headphones. It adds complexity, but it will undoubtedly improve the performance.

And if you think the inevitable latency in digital circuits is a problem, try adding a software plug-in. Now the latency is obvious to anyone.

This compromises the way we work, since if you need an eq or compressor operating on the track while recording, you can't do it with a software plug-in most of the time.

Personally, I use analog hardware outboard gear 99% of the time. It's not a problem for me. But it does mean cutting a track with the processing already done, which makes some people nervous. You can't easily un-do that processing if you later decide you don't like it or need it. But that was a necessity in the old days and I rarely regret cutting a track with the processing I know it will need.

Don't get me wrong, I don't want to go back to analog recording. Digital offers way too many advantages.

Another workaround is to set up your session so that headphones are not needed. I talk about this approach in the episode on Minimalist Mic'ing. It is specialized technique and not possible with most projects.

Is this a problem we will always have to live with? Computers today are much faster, but the operating system and the digital audio software are much more complex, so we have not made any significant improvement in the latency. I suspect much of the latency could be eliminated with a special, stripped-down operating system designed just for audio recording. There doesn't seem to be much interest in doing that, however.

I think IZ Corp, the makers of the Radar recorder, had the right idea when they decided to use a special version of Unix to operate their proprietary system. The operating system was very lean, and even with a relatively slow processor, latency was never a problem, even with dozens of tracks. That system never crashed on me, either.

Maybe a breakthrough in computer technology will dramatically speed up our machines and latency will become just an archaic memory. This problem does not affect 99% of computer users, so it is not likely to become an R&D priority. Video and audio producers would surely rejoice, but no computer or chip manufacturer is going to be a pioneer for that miniscule segment of the computer market.

Now let's take a look at what might be called "good" delay.

A recording made in an anechoic chamber, a room with no reflections whatsoever, does not sound good at all. We record in real rooms, and that means that we usually want some sense of the space. If we have a big studio with a great sound, the studio becomes part of the recording. It can make a huge positive difference, not only in how the record sounds, but also in how the musicians play.

These days, most of us do not have regular access to great rooms, so we use the space we have. Your room might sound pretty terrible, but a lot can be done to improve the space. Listen to some of my previous episodes on this for some guidance.

In a real room, with a mic close to the performer, the first thing we hear is the direct sound. That is followed by reflections from the floor, walls, and ceiling, plus from anything else in the room. In a great room, these reflections are well-balanced both in time and frequency response, and we love the way it sounds.

Our brain makes assumptions about the room based on the time delays involved, and that allows us to picture the room where the performance took place.

In the simplest case, the sound bounces off a wall and is picked up by the mic with a delay of about 1mS per foot. In a room that has a wall 20 feet away, that is about 40mS of delay. Anyone will hear that if the reflection is strong enough. Now multiply that by many paths from many surfaces and objects, and, if you are lucky, the result is a nice distribution and balance of echoes. This adds fullness to the sound and is the key to the sound of many classic records.

Of course, if the reflections conflict or build up in a small range of frequencies, the sound is adversely impacted. It might sound “hollow” or “tubby.” Echoes that are too strong might be too reverberant for the music. That can muddy up the sound.

If the room does not have good reverberant characteristics, then your best bet might be to deaden the room as much as possible and add artificial reverberation or delays.

Often, short delays can add a lot of life to the recording.

Early in my career, I only had tape delay to use as “reverb.” This is possible because professional tape machines have separate heads for recording and playback, which have a certain minimum separation in spacing, and thus in time between recording and playback. Most of the time, you simply monitored the audio going into the machine. But you could take the playback and bring it up on a fader on your console. Pushing that up introduced the slightly delayed playback into the mix.

A variation of this would have some of the delayed audio fed back into the recording, which results in an ongoing repetition of the sound. Taken to an extreme, this will set up an oscillation that grows stronger with each repetition, like feedback in a PA system. Not useful. But small amounts can create a distinctive sound you hear on many records from the 1950s.

At WPEN, we had a tape delay machine made by a Hollywood company called Surround Sound. It had a continuous loop of magnetic tape around a large, rotating wheel, with a record head and multiple playback heads. The record head was on an arm that you could rotate around the wheel, adjusting the delay time. The radio station used this on most of the programming to give the station a distinctive sound.

I didn't like what it did to the audio, especially with music, but the station used it for many years. Eventually, they decided to eliminate it, and I bought the machine from them.

Ideally, it would have been possible to extract the audio from each playback head independently and use those to create a nice stereo spread of the echoes. That Surround Sound machine had many controls to adjust the level of each delay, the amount of feedback, and eq. But the machine was mono. I got around this on some projects where there was a surplus of tracks and I could record multiple delays on separate tracks. I could then pan them wherever I wanted and adjust the level.

For a vocal, my usual pattern was to have a first delay panned hard left, and a slightly longer delay panned hard right, and a third delay back near center, behind the main vocal. Usually I would take the third delay and send it to the EMT reverb, too.

The delay tracks did not need to be very loud in the mix. I liked them almost subliminal. This made the vocal pop out of the mix.

I found the effect less useful on instruments, except maybe on horns.

My plan was to modify the Surround Sound machine so I could have separate outputs from the multiple playback heads, but that would have been a lot of work. And it was the late 1970s, when digital delay units started to appear.

I bought a Lexicon Prime Time delay, which was extremely useful and soon became my main delay device. There were other manufacturers of digital delays, and I had a couple of those, plus an Eventide Harmonizer.

With all those delays available, I was able to create some really nice effects.

Today, we have plug-ins to do this, but I have yet to find one that does what I want. They seem overly complex for their simple function. What I do these days is to duplicate the vocal track two or three times and then simply displace them slightly in time, as needed, to achieve my old favorite effect.

Another use of delay, either tape or digital, back in the Mesozoic era of recording was before the EMT plate reverb. This displaced the onset of the reverb, which kept it out of the way of the music, and gave the impression of a larger room.

There are many other applications of delay that can be used creatively in recording. It's easy to do, and with the multiple track time offset method, there is zero cost because you don't need any plugins or hardware.

Occasionally it may be useful to have a "pre-echo." That is, the echo comes before the main sound. This was used on records in the 1960s and 70s, and although it is rarely appropriate, it can be useful on some songs. In the old days, this was done by reversing the tape and recording echo or reverb on an empty track, keeping in mind that all your tracks are now on different faders. It is a unique sound.

Delay can be your friend or your enemy, depending on the circumstances. Knowing when to utilize these effects to your advantage, and when to minimize them when they interfere with the creative process can be used to improve your recordings.

Thanks to all of you who have subscribed to this podcast on any of the 30 or more podcast providers that carry it. You can always reach me at dwfearn@dwfearn.com

This is My Take On Music Recording. I'm Doug Fearn. See you next time.