Dither

Keeping Your Digital Audio Pure from First Recording to Final Master

Part I

Dither is not the most important technical detail to learn about, but if you want to get your digital audio done just right, then you should learn about dither. Especially if you want to learn why your digital reverb has been leaching the ambience out of your music, when it's supposed to be *adding* ambience. Or why your CDs don't sound as spacious as your 24 bit sources and want to avoid that veiled, dry, and lifeless feeling!

Follow that Sample

Let's start with a little lesson in DSP (Digital Signal Processors). Many workstation and processor manufacturers ignore the critical issue of **wordlength**. Let's examine what happens to digital audio when you change gain (or mix, equalize, compress, sample rate convert, or perform any type of calculation) in a digital audio workstation. It's all arithmetic, isn't it? Yes, but the accuracy of that arithmetic, and how you (or the workstation) deal with the arithmetic product, can make the difference between pure-sounding digital audio or digital sand paper.

All DSPs deal with digital audio on a sample by sample basis. At 44.1kHz, there are 44,100 samples in a second (88,200 stereo samples). When changing gain, the DSP looks at the first sample, performs a multiplication, spits out a new number, and then moves on to the next sample. It's that simple.

Instead of losing you with esoteric concepts like two's complement notation, fixed vs. floating point, and other digital details, I'm going to talk about *digital dollars.* Suppose that the value of your first digital audio sample was expressed in dollars instead of volts, for example, one dollar and fifty one cents--$1.51. And suppose you wanted to take it down (attenuate it) by 6 dB. If you do this wrong, you'll lose more than money, by the way. 6 dB is half the original value (it has to do with logarithms; don't worry about it). So, to attenuate our $1.51 sample, we divide it by 2.

Oops! $1.51 divided by 2 equals 75-1/2 cents, or .755. So, we've just gained an extra decimal place. What should we do with it, anyway? It turns out that dealing with extra places is what good digital audio is all about. If we just drop the extra five, we've theoretically only lost half a penny--but you have to realize that half a penny contains a great deal of the natural ambience, reverberation, decay, warmth, and stereo separation that was present in the original $1.51 sample! Lose the half penny, and there goes your sound. The dilemma of digital audio is that most calculations result in a longer wordlength than you started with. Getting more decimal places in our digital dollars is analogous to having more bits in our digital words. When a gain calculation is performed, the wordlength can increase infinitely, depending on the precision we use in the calculation. A 1 dB gain boost involves multiplying by 1.122018454 (to 9 place accuracy). Multiply $1.51 by 1.122018454, and you get $1.694247866 (try it on your calculator). Every extra decimal place may seem insignificant to you, until you realize that DSPs require repeated calculations to perform filtering, equalization, and compression. 1 dB up here, 1 dB down here, up and down a few times, and the end number may not resemble the right product at all, *unless adequate precision is maintained.* Remember, the more precision, the cleaner your digital audio will sound in the end (up to a reasonable limit).

The First Secret of Digital Audio

Now you know the first critical secret of digital audio: *wordlengths expand*. If this concept is so simple, why is it disregarded by some manufacturers? The answer is in your wallet. While DSPs are capable of performing double and triple precision arithmetic (all you have to do is store intermediate products in temporary storage registers), it slows them down, and complicates the whole process. It's a hard choice, entirely up to the DSP programmer/processor designer, who's been put under the gun by management to fit more program features into less space, for less money. Questions of sound quality and quantization distortion can become moot compared to the selling price.

Inside a digital mixing console (or workstation), the mix buss must be much longer than 16 bits, because adding two (or more) 16-bit samples together and multiplying by a coefficient (the level of the master fader is one such coefficient) can result in a 32-bit (or larger) sample, with every little bit significant. Since the AES/EBU standard can carry up to 24-bits, it is practical to take the 32-bit word, round it down to 24 bits, then send the result to the outside world, which could be a 24-bit storage device (oranother processor). The next processor in line may have an internal wordlength of 32 or more bits, but before output it must round the precision back to 24 bits. The result is a slowly cumulating error in the least significant bit(s) from process to process. Fortunately, the least significant bit of a 24-bit word is 144 dB down, and most sane people recognize that degree of error to be inaudible.

Something For Nothing?

But suppose you want to record the digital console's output to a 16 bit medium, like the CD. Frankly, it's a serious compromise to take your console's 24-bit output word and truncate it to 16 bits. After processing, the mastering engineer uses a technique called *dithering* to take long wordlengths, and cleanly turn them to 16-bit for the CD. First, must ensure that our DAW is high resolution (has very low distortion at low levels) and can be bit-transparent when called upon. **Bit-transparent** means that the output is identical to the source, from the most significant to the least significant bit, that the DAW does not increase or decrease the source wordlength.

Good Advice

Once you've verified your workstation is bit-transparent, then proceed with editing, with the goal of maintaining the integrity of your original source. **Do not change gain** unless you need to align the gains of two pieces you are editing together. **Do not** normalize (normalization is just changing gain). **Do not equalize. Do not fade in or fade out.** Just edit. This is to avoid additional DSP or degradation when the mix gets to the mastering studio .Leave the segues, fadeouts and gain changes for the mastering house, where they can properly handle the long wordlengths necessary for smooth fades (so that's why your last fadeout sounded like it dropped off a cliff!). Follow these simple guidelines and your digital audio will immediately start sounding better.

Part II

Dither

How to Dither Let's look at that long sample word. Whether it's 24 bits or 32 bits, we have to find some way to move the important information contained in the lower (least significant) bits into the upper 16 bits for recording to the CD standard. Truncation is very bad. What about rounding? In our digital dollar example, we ended up with an extra 1/2 cent. In grammar school, they taught us to round the numbers up or down according to a rule (we learned "even numbers...roundup, odd...round down"). But when we're dealing with more numerical precision and small numbers that are significant, it gets a little more complicated.

It turns out the best solution for maintaining the resolution of digital audio is to calculate random numbers and add a different random number to every sample. Then, cut it off at 16 bits. The random numbers must also be different for left and right samples, or else stereo separation will be compromised.

For example:

Starting with a 24-bit word (each bit is either a 1 or a 0 in binary notation):

```
                      Upper 16 bits        Lower 8
Original 24-bit Word MXXX XXXX XXXX XXXW  YYYY YYYY
Add random number                         ZZZZ ZZZZ
```

The result of the addition of the Z's with the Y's gets carried over into the new least significant bit of the 16-bit word (LSB, letter W above), and possibly higher bits if you have to carry. In essence, the random

number sequence combines with the original lower bit information, *modulating* the LSB. Therefore, the LSB, from moment to moment, turns on and off at the rate of the original low level musical information. The random number is called *dither;* the process is called *redithering*, to distinguish from the original *dithering* process used to during the original recording.

Random numbers such as these translate to random noise (hiss) when converted to analog. The amplitude of this noise is around 1 LSB, which for 16 bit lies at about 96 dB below full scale. By using dither, ambience and decay in a musical recording can be heard down to about -115 dB, even with a 16-bit wordlength. Thus, although the quantization steps of a 16-bit word can only theoretically encode 96 dB of range, with dither, there is an audible dynamic range of up to 115 dB! The maximum *signal-to-noise* ratio of a dithered 16-bit recording is about 96 dB. But the *dynamic range* is far greater, as much as 115 dB, because we can hear music *below the noise.* Usually, manufacturer's spec sheets don't reflect these important specifications, often mixing up dynamic range and signal-to-noise ratio. *Signal-to-noise ratio* (of a linear PCM system) is the RMS level of the noise with no signal applied expressed in dB below maximum level (without getting into fancy details such as noise modulation). It should be, ideally, the level of the dither noise. *Dynamic range* is a subjective judgment more than a measurement--you can compare the dynamic range of two systems empirically with identical listening tests. Apply a 1 kHz tone, and see low you can make it before it is undetectable. You can actually measure the dynamic range of an A/D converter without an FFT analyzer. All you need is an accurate test tone generator and your ears, and a low-noise headphone amplifier with sufficient gain. Listen to the analog output and see when it disappears (use a real good 16 bit D/A for this test). Another important test is to attenuate music in your workstation (about 40 dB) and listen to the output of the system with headphones. Listen for ambience and reverberation; a good system will still reveal ambience, even at that low level. Also listen to the character of the noise--it's a very educating experience.

Some Tests for Linearity
You can verify whether your digital audio workstation truncates digital words or does other nasty things, without any measurement instruments except your ears. Obtain the disc *Best of Chesky Classics and Jazz and Audiophile Test Disc, Vol. III*, Chesky JD111.* Track 42 is a fade to noise without dither, demonstrating quantization distortion and loss of resolution. Track 43 is a fade to noise with white noise dither, and track 44 uses noise-shaped dither (to be explained). Use Track 43 as your test source; you should be able to hear smooth and distortion-free signal down to about -115 dB. Then listen to track 44 to see how much better it can sound. Try processing track 43 with digital equalization or level changes (both gain and attenuation, with and without dither, if it's available in your workstation) to see what they do to the sound. If your workstation is not up to par, you'll be shocked. Use a quiet, high-gain headphone amplifier to help reveal the low level problems.

*available at major record chains or through Chesky Records, Box 1268, Radio City Station, New York, NY 10101; 212-586-7799. The hard-to-find CBS CD-1, track 20, also contains a fade to noise test.

So Little Noise, So Much Effect
-96 dB seems like so little noise. But strangely, engineers have been able to hear the effect of the dither noise, even at normal listening levels. Dither noise helps us recover ambience, but conversely it also obscures the same ambience we've been trying to recover! Dither noise adds a slight veil to the sound. That's why I say, *dither, you can't live with it, and you can't live without it.*

Improved Dithering Techniques
Where there's a will, there's a way. Although the required amplitude of the dither is about -96 dB, it's possible to shape (equalize) the dither to minimize its audibility. Noise-shaping techniques re-equalize the spectrum of the dither while retaining its average power, moving the noise away from the areas where the ear is most sensitive (circa 3 KHz), and into the high frequency region (10-22 KHz).

Here is a picture of one of the most successful noise-shaping curves (courtesy of Meridian Audio, Ltd).
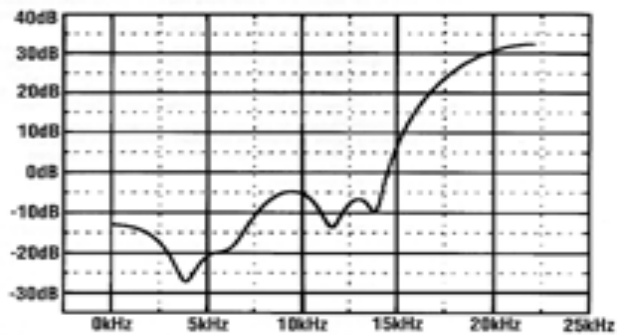
As you can see, it is a very high-order filter, requiring considerable calculation, with several dips where human hearing is most sensitive. The sonic result is an incredibly silent background, even on a 16-bit CD. The 0 dB line is around -96 dBFS in this diagram.

There are numerous noise-shaping redithering devices on the market. Very high precision (56 to 72 bit) arithmetic is required to calculate these random numbers. One box uses the resources of an entire DSP chip just to calculate dither. The sonic results of these noise-shaping techniques range from very good to marvelous. The best techniques are virtually inaudible to the ear. With 72-bit arithmetic, all the dither noise has been pushed into the high frequency region, which at -60 or -70 dB is still inaudible. Critical listeners were complaining that the high frequency rise of the early noise-shaping curves changed the tonality of the sound, adding a bit of brightness. But it turns out that it is the shape of the curve in the midband that affects the tonality, due to masking. Two or three of the latest and best of these noise-shaping dithers are tonally neutral, to my ears. It took a long time to get there (about 10 years of development), but now we can say that the best of these processors yield 19-20 bit performance on a 16-bit CD, with virtually no tonal alteration or loss of ambience from the 24-bit source.

Noise-shapers on the market include: db Technologies model 3000 Digital Optimizer, Meridian Model 618, Sony Super Bit Mapping, Waves L1 and L2 Ultramaximizers, Prism, POW-R, and several others. When using dithering plugins, be sure to use them with the right version of workstation software to retain a 24-bit wordlength until the final mastering step.

Apogee Electronics produced the UV-22 system, in response to complaints about the sound of earlier noise-shaping systems, declaring that 16-bit performance is just fine. They do not use the word "dither" (because their noise is periodic, they prefer to call it a "signal"), but it smells like dither to me. Instead of noise-shaping, UV-22 adds a carefully calculated noise at around 22 KHz, without altering the noise in the midband.

To effectively compare the sound and resolution of these redithering techniques, perform the low level test described above. Feed low level 24-bit music (around -40 dB) into the processor, and listen to the output at high gain in a pair of headphones with a good quality D/A converter. You will be shocked to hear the sonic differences between the systems. Some will be grainy, some noisy, and some distorted, indicating improper dithering or poor calculation. The winner of this test should be your choice of dithering processor, although at high gains you are exaggerating the effect of the extra high frequencies in the noise, which would not be noticed at normal gains..

Damage, Destruction, or just Deterioration?
Before digital recording and editing, every edit was destructive. Every equalization or gain change involved an analog copy, with attendant noise, or remixing the multitrack, which "destroys" or replaces the previous mixdown. After DAWs were invented, people started talking about "non-destructive"-editing, and keeping your sound in the digital domain until the end. But as we have seen, even "non-destructive" may be damaging if word lengths aren't maintained.

The Best Approach
To maintain the quality of your digital audio, always store the full output wordlength of your digital

processors. Also, be sure to *question authority.* Never take a digital processor for granted. Don't even trust BYPASS mode, unless you're sure the processor produces true clones in bypass. The following illustration (courtesy of Jim Johnston, AT&T research), shows a series of FFT plots of a sine wave. The top row is an undithered 16 bit sine wave. Note the distortion products (vertical spikes at regular intervals). The second row is that sine wave with uniform dither. Note how the distortion products are now gone. The bottom row is the dithered sine wave, going through a popular model of digital processor set for BYPASS and truncated to 16 bits. This is what would happen if you took your source, fed it through this processor in BYPASS mode, and recorded it again!

Sine Wave
16 bits
No dither

Sine Wave
16 bits
Uniform dither

Dithered sine multiplied
by 1-2^-24

(Bypass mode of a
popular effects box)

240 Hz

1 kHz

17 kHz