CHAPTER 16

Analog and Digital Processing

The mastering engineer must recognize when a recording is so good that the interests of the client are best served simply by leaving it alone. And there are recordings for which so little work is needed that the gains due to processing would not warrant the losses due to the same processing! For although equipment is getting better, there is no such thing as a transparent audio processor. This chapter is about how we measure and interpret performance, as there is an interaction between objective degradation and subjective improvement. Let's take a journey into the twilight zone between the objective and the subjective.

I. The Ironies of Perception vs. Measurement

Although we'll be using test measurements, we must remember that each single measurement only provides a small part of the picture. An audio processor is like an object inside a house with no doors, only a number of small windows that you can peer into. By looking at the object through each window's unique angle we can find out more, and add up the clues, but we can never be totally sure of what we are seeing, and must always leave open the possibility that there may be some aspect we cannot see, some mystery as to why this equalizer sounds "good" and this other one sounds "bad."

For example, here are a couple of "objective" measurements that just don't add up!

What Makes it Sound Bright?

I've discovered a digital filter that measures "dull" but sounds bright! The TC Electronic System 6000 lets the user choose between different low-pass filters for the A/D and D/A converters. Some of the filters roll off significantly above 16 kHz (at 44.1 kHz sampling), so you'd think they would sound dull. But instead, to my ears, the 16 kHz filters called *Natural* and *Linear* sound more *open* and *clear* than the particular 20 kHz filter called *Vintage*. However, there are other converters whose filters extend to 20 kHz and which sound even more *open* than the TC's Linear filter. So measured bandwidth cannot tell the whole psychoacoustic story. We look into the audible effects of filtering in Chapter 18.

The Fallacy of Typical Weighting Curves

We have equipment in our studio whose noise floor measures as low as –120 dBFS to as high as –50 dBFS (after A/D conversion). However, much of this equipment is perceptually quiet: if I have to put my ear up to the loudspeaker to hear the hiss, then I consider it insignificant. Interestingly, the weighting methods[1] by which converter manufacturers commonly measure noise bear little relationship to human perception. One particular converter whose A-weighted noise floor is –108 dBFS sounds significantly quieter than another converter whose A-weighted noise floor is –115 dBFS! The reason is that the often-cited, A-weighted curve does not adequately consider the ear's greater sensitivity in critical bands. It turns out that the converter which measures better (A-Weighted) produces significantly more energy circa 3 kHz, where the ear is most sensitive, and the A-weighting filter does not take into account the significance of this critical band. To be psychoacoustically accurate, noise measurement standards should adopt a curve closer to the measured noise floor of the human ear, such as the 9th order curve used by some of the best-sounding dithers (see Chapter 4). This curve is called "F" weighting.[2]

> *"Never turn your back on digital."*
>
> —Bob Ludwig.

There are many other areas in which traditional measurements do not correlate with what our ears tell us, particularly in the evaluation of low bit rate coding systems. These systems measure quite well with standard techniques, but once the ear has been trained to hear their errors, we can easily identify artifacts we've never heard before with analog technology: described by some as *chirping*, or *space monkeys*. Let's see if we can objectively find out why some analog and digital processors sound better than others. Just remember that measurements look at an object through a few narrow windows, and there may be a different, or better, explanation for sound quality than what I've come up with.

II. Measurement Tools We Can Use While Mastering

FFT Measurements

FFT stands for *Fast Fourier Transform*. To really learn how to interpret (and not misinterpret) an FFT requires a college-level engineering course, and although I cannot claim to be such an expert, I have learned just enough to be dangerous! High-resolution FFT analysers, such as SpectraFoo™, are very reasonably priced, thanks to the exponential increase in CPU power and they provide an essential *early warning system*, a protection from the vicissitudes (bugs) of digital audio. *Never turn your back on digital*, says Bob Ludwig, or as I say, *you're only one mouse click away from disaster!* It's a whole new world based on software designed by fallible human beings.

FFT for Music

Figure C16-01 in the Color Plate section shows SpectraFoo in action during a CD mastering session.

At the middle top is a bitscope, currently showing 16 (and only 16) active bits, an indication that the dither generator is probably doing its job. This bitscope can reveal if some digital device is malfunctioning, since one of the symptoms of a disfunctional processor is to toggle unwanted bits, or hold some bits steady when there is no signal. Bitscopes can also show if there are any unwanted truncations caused by defective or misused processors. However, the bitscope is only one of the small windows we can look through; it can easily miss problems, or seem to indicate problems which require further interpretation. For example, some equalizers produce idle noise when the music goes to silence. This can be perfectly normal, but will show up on the bitscope as activity. Toggling the equalizer in and out while observing the bitscope will ascertain if that is the source of the problem or some other anomaly in the signal chain.

At top right is a stereo position indicator, which is frozen at a moment when the information is slightly right-heavy. At left is a meter that conforms to the K-14 standard (see Chapter 15). The meter shows the hottest moment of a rather hot R&B piece (which I would have preferred to reduce, but the client desired it this hot!). For the record, this material was monitored at -8 dB, which really makes it K-12 material. Just below the bitscope is a correlation indicator, revealing that the material is significantly monophonic. I prefer a correlation indicator to an oscilloscope; meter deflections closer to the center of the scale indicate less correlation from channel to channel and likely a larger or more spacious stereo image. However, I always use my ears to confirm the image is not too "vague" and perform a mono (folddown) test to make sure the sound is mono-compatible.

At mid-screen is the spectragram, showing spectral intensity over time. This can be useful to identify the frequencies of problem notes, or simply to entertain visitors! At bottom is the spectragraph, whose general rolloff shape gives a vague idea of the program's timbre (though most times I disregard the spectral displays, since the eye candy of the visual display distracts our aural senses).

Figure C16-02 in the Color Plates shows SpectraFoo during a pause in the music, with only the bottom four bits toggling, confirming that the dither is working correctly, since dithers which use heavy noise-shaping exercise several bits. Note that the bitscope shows four bits toggling (since dither is random, in this snapshot, bit 15 is at zero) and that the spectragraph shows the curve of the dither noise, which can be identified by its shape as POW-R type 3 or a similar 9^{th} order curve. Using this analyzer, you can often determine the type of dither used by the mastering engineer on recorded CDs.

The level meters had not decayed fully when this shot was taken. The correlation meter fluctuates very slightly near the meter's center, showing that the dither is uncorrelated between channels (random phase). I always glance at this display at the beginning and end of the program, to make sure no bugs or patching errors have crept in. I carry a SpectraFoo umbrella even if it's not raining!

II. Measurement Tools to Analyze your Equipment

Let's sort out what happens beneath the knobs. As in geometry, the shortest distance between two points is a straight line, so too in audio — both digital and analog — the cleanest signal path contains the fewest components. The converter used to be the most degrading piece in the studio, but although they have greatly improved in recent years, we should still avoid extra conversion whenever possible. For analog tapes, it's best to do all the analog processing on the way to the first and only A/D conversion. But these days mixes are often on digital tape, and as there are a lot of desirable analog processors which the mastering engineer may prefer because they sound more *organic* than their digital equivalents, the tonal benefits of analog processing might outweigh the transparency losses of an extra conversion.* The best defense is a good offense, and it is possible to reliably measure signal below the noise with an FFT analyzer. An FFT can confirm if a digital processor is not truly bypassed when it says *bypass*, which can be pretty deleterious (see Chapter 4). Jitter (see Chapter 19) is irrelevant to FFT analysers, which strictly look at data.

Even though the analyzer can only examine 24 bits (the limitation of the AES/EBU interface), it can measure distortion 40 dB below the 24-bit noise floor! This is because Spectrafoo is a 64-bit floating point system. So we can compare the distortion of processors which truncate at the 24th bit versus others which use 48 bits or so internally and then dither up to 24 bits. Whether we can hear these differences is a different question. Psychoacoustician J. Robert Stuart has demonstrated that we can hear a 24-bit truncation in an 18-bit system. The ear's dynamic range is approximately 20 bits (120 dB), but this varies with frequency. At certain frequencies we can even hear below 0 dB SPL!

How Many Bits is Enough?

In color plate ***Figure C16-03***, we compare 16, 20, and 24-bit flat-dithered noise.[3] The levels of all the "bins" add up, so at 16 bits, the curve which looks like it rides at approximately -124 dBFS (level of individual bins) totals to an RMS level of about -91.2 dBFS RMS, the theoretical limit of a properly-dithered 16-bit system. But discrete signals at some frequencies can be heard as low as -115 dBFS in a properly-dithered 16-bit system, below which they are buried in the noise. Psychoacoustically, for the vast majority of popular and classical music, 16 bits properly done are just enough to do the job right. But as soon as we post-produce, copy, process and change gain, we accumulate noise and need professional headroom, or perhaps we should call it *footroom*† since the top, at 0 dBFS, is a constant.

Psychoacousticians studying the limits of the human ear have determined that 20-bits is enough

* And losses can be minimized using upsampling (see Chapter 1).

† This is a made-up word, not an official term!

for good A/D and D/A performance. Anything more is just gravy, and it's very rare to find a "24-bit" converter with better than 18-20-bit noise level. For processing, however we need the additional *footroom*, better than 24 bits, because the frequency-content of digital distortion is far more annoying to the ear than analog distortions which are much louder. **This is because distortion created during digital processing yields harmonic components which beat against the sample rate, producing dissonant inharmonic beat or intermodulation products.** For purist processing, we may need as much as 48 to 72 bits, especially for extreme gain changes, complex filtering, compression, or to avoid cumulative distortion when cascading processes. It's a myth that there's no generation loss in digital processing; **little by little, bit by precious bit, sound suffers with every DSP operation.**

Figure C16-04 in the color plates shows the noise floor of a popular dither called POW-R type 3 at 16-bit (red trace). For reference, we show the noise of flat 20-bit dither (orange), and 24-bit dither (green). POW-R's shape is designed to maximize performance by keeping the noise at or near the ear's low-level sensitivity at various frequencies. POW-R dither reaches 20-bit performance in the critical upper midrange (circa 3.5 kHz) where the ear is most sensitive. Thus, much of the low level ambience and reverberation that would have been masked is revealed, even with 16-bit reproduction. This performance can only be achieved by recording at a longer wordlength to begin with, as noise accumulates and the SNR gets slightly worse when you add final dither to the processed source.

Analog versus Digital Processing

Cheap versus Good...Is It Really Accurate?

Many people have argued that the reason we notice harshness in some digital recordings is that digital audio recording is more *accurate* than analog. Their claim is that the *accuracy* of digital recording reveals the harshness in our sources, since digital recording doesn't compress (mellow out) high frequencies as does low speed (15 IPS) analog tape. *Accuracy*, they say, is why we have regressed to tube and vintage microphones. But I say this is only a half-truth, since most of these arguments come from individuals who have not been exposed to the sound of good digital recording equipment, which is not only accurate, but can even be *warm and pretty*. Cheap digital equipment is subject to edgy sounding distortion which can be caused by sharp filters, low sample rates, poor conversion technology, low resolution (short wordlength), poor analog stages, jitter, improper dither, clock leakage in analog stages due to bad circuit board design and many others, such as placing sensitive A/D and D/A converters inside the same chassis with motors and spinning heads. It takes a superior power supply and shielding design to make an integrated digital tape recorder that sounds good; compare the sound of an inexpensive modular digital multitrack (MDM) with the Nagra Digital recorder—4 very expensive tracks versus 8 cheap ones.

When it comes to processing, numeric precision is also expensive, even though it's all software. Numeric imprecision in digital consoles

MYTH:
It's a digital processor, so there's no generation loss.

MYTH:
It's a Digital Console. It must be better than my old analog model!

produces problems somewhat like noise in noise in analog consoles, but there is an important difference: noise in analog consoles gradually and gently obscures ambience and low-level material and usually does not add distortion at low levels. However, numeric imprecision in digital consoles causes quantization errors (which increase at low levels) destroying the body and purity of an entire mix, creating edgy, colder, sound, which audiophiles call **digititis**. Since digital consoles do not make sound warmer, depending on the quality of their digital processing—and the number of passes through that circuitry—it might be better to mix through a high-quality analog console.

> *"Audio processing is the art of balancing subjective enhancement against objective degradation."*
> —Bob Olhsson.

Even though good digital equipment is getting cheaper at an exponential rate, it is still expensive to produce excellence in digital recordings. That's why analog tape and analog mixing remain very much alive at this point in the 21st century.

Two Fine Equalizers, One Analog, One Digital

In my opinion, much inexpensive tube equipment is overly warm, noisy, unclear and undefined, and the common use of "fuzzy" analog equipment to cover up the problems of inexpensive digital equipment is a band-aid, not a cure for the loss of resolution. Not many people have been exposed to recent audiophile-quality tube equipment, and only the best-designed tube equipment has quiet, clear sound, tight (defined bass), is transparent and dimensional, yet still warm. Audiophiles feel a well-designed tube circuit can be more linear and resolving[4] than a low-cost solid state circuit. I certainly feel I hear more through some amplifiers than others. Modern-day tube designers often make innovative use of low-noise regulated power supplies on filaments and cathodes, a practice which was impractical in the 50's.

Figure C16-05 in the Color Plates section shows the low distortion and noise performance of a well-designed, popular state-of-the art analog tube equalizer, the **Millennia NSEQ-2** (red trace). For reference, 20- and 24-bit noise are shown in blue and green, respectively. Notice that the tube noise of the NSEQ is about 10 dB greater than 20-bit, making it a *virtual 18-bit analog equalizer*. However, this performance is dependent on the analog gain structure used. If you drive the equalizer harder, its noise floor will be lower compared to maximum signal, and distortion may or may not be a problem. Since the Millennia's clipping level is around +37 dBu, it may be perfectly legitimate to drive it with nominal levels of +10 dBu or even higher, provided the source equipment doesn't overload! Yet even with nominal levels of 0 dBu as was used for this graph, this tube equalizer is extremely quiet. Its noise is inaudible at any reasonable monitor gain unless you put your ears up to the speaker, demonstrating that noise-floor is probably the least of our worries. 1/2" 30 IPS 2-track analog tape has even higher noise, but no one complains about it for popular music.

For this FFT, we set up a D/A converter, feeding the NSEQ and then an A/D and the FFT. A digitally-generated 1 kHz -6 dBFS 24-bit dithered sine wave feeds the D/A. We adjust converter gain so 0 dBFS is +18 dBu, and boost the equalizer about 6 dB, till just below A/D clipping. The equalizer is coasting at this level, since it's around 19 dB below its clip level! If you are looking for extreme "tubey" effects, you can drive the equalizer even harder, and also realize a greater SNR, provided the converters can handle the hotter level, certainly the equalizer can.

Notice that the equalizer's distortion is dominated by second, third, and fourth harmonics, which tend to *sweeten* sound. For comparison, in yellow is the performance of the superb **Z-Systems** digital equalizer, dithered to 24 bits, boosting 1 kHz 5.8 dB with a Q of 0.7. Its harmonic distortion performance is textbook-perfect (no visible harmonics on the FFT). Some engineers use the word "dry" to describe the sound of a component that has little or no distortion. Looking through other "windows" we find that harmonics are far from the only sonic differences between these pieces of gear. Tubes, power supplies and transformers can *loosen* the bass, which can sometimes be desirable; the digital equalizer retains the tightness of the bass;* the digital and analog equalizer's curves are also different, though the ZQ-2 does a nice job of simulating the shapes of gentle analog filters. Equalizer curve shape and phase shift probably make up other areas of delicate sonic difference between models of equalizers.

The premium price of both the ZQ-2 and the NSEQ reinforce my point that high-quality analog or digital recording is expensive. At the time of this writing, it will be a number of years before there's enough power in a typical computer plug-in to come up to the quality of the best outboard processors.

"Nasty" Digital Processors

Truncation distortion can be fairly "nasty." For example, in ***Figure C16-06*** of the Color Plates section, we compare the analog Millennia NSEQ (orange trace) versus the digital Z Systems set to truncate at 20 bits, no dither (black trace).

Don't try this at home! I think there are better ways to add *grunge* than turning off the dither. Much of the ambience, space, and warmth of the original source have been truncated, lost forever, converted to low level grunge (severe inharmonic distortion and noise). Even a small amount of non-harmonic distortion can be bothersome. Which sounds better, an analog processor with a smooth but higher noise floor, plus second and third harmonic distortion, or an undithered digital processor with a lower average noise floor plus inharmonic distortion?

Poorly-implemented digital compressors produce severe inharmonic distortion, which is without integer relationship to the fundamental. ***Figure C16-07*** in the Color Plates compares two digital compressors, both into 5 dB of compression with a 10 kHz signal.

* Since digital equalizers don't soften the bass like some tube units, you may wish to "loosen" the bass with compression or some other tool.

In orange is a single-precision, non-over-sampling compressor, and in black a double-sampling compressor implemented in 40-bit floating point. Note the single-precision compressor produces many non-harmonic aliases of the 10 kHz signal, especially in the critical midband. Nasty-sounding first-generation compressors are still common in low-cost digital consoles and DAW plugins. It takes a lot of processing power to double-sample. I'm convinced that the proliferation and misuse of cheap digital processing has degraded the sound quality of much recently-recorded music.

The Magic of Analog?

Static distortion measurements don't explain every reason why some compressors sound *excellent* and others hurt your ears. There are analog processors which are so *magical* that though they are not transparent, they add an interesting and exciting sonic character to music, or to put it another way, *their subjective cure is better than their objective disease.* Analog tape recording is a perfect example of this type of process; measured objectively it's noisy and distorted, but subjectively it can kick ass! If psychoacoustic research had been a bit more advanced on the audible effects of masking distortion and noise, then perhaps we may not have pursued this expensive search for 144 dB extremes. For example, the noise floor of the Sony-Philips DSD system is not particularly special (about 120 dB in the audible band), but it sounds excellent, indicating that low-noise must not be our only goal. We may even conclude that part of the good sound is due to masking; maybe -120 dB is just enough to cover the ugly parts of the distortion of even some of our best analog and digital gear. In addition, noise-free recording media can be very *sterile-sounding* because all the nits and cracks and distortions caused by the musicians and their amplifiers are completely revealed by the quiet media. So, sometimes, adding extra noise can be more beneficial to the music than working noise-free. Perhaps one of the many reasons why analog tape sounds more *musical* to many people...noise can be very euphonic. We should certainly experiment with noise-masking and make our decisions on what is best for the music. *[Please see sidebar, **Clarity or Fuzz**.]*

I think that many classic analog compressors' warm, fat yet clear sound signatures come from a unique combination of attack and release characteristics, which may be emulated in a digital processor. There are some plug-ins which emulate classical analog compressors but to my ears they do not come up to the job; I think they will get better over time when the cost of DSP goes down. Currently, plug-in designers are forced to minimize the DSP load of their processors or users complain they can't fit a plug-in on every channel strip (as if this is desirable). Certainly the Weiss digital compressor does not sound *digital*, so we know that it can be done with programming skill and expensive DSP.

An Analog Simulator-Pick your flavor of grunge

Figure C16-08 in the Color Plates compares the NSEQ to the Cranesong HEDD-192, a digital *analog simulator* of excellent sound quality.

The Cranesong (blue trace) has been adjusted to produce a remarkably similar harmonic structure to the NSEQ. For this graph, its levels have been purposely set to produce more distortion than the Millennia was producing. Amazingly, the ear thinks it's hearing an excellent analog processor without any imaging or resolution loss. But the low-level grunge at the bottom of the picture looks mighty suspicious; looking through this "window" you might think the Cranesong was truncating important information. But two important factors ameliorate: First, the Cranesong's grunge is about 12 dB lower than that of a truncated device and thus is likely masked by the noise and the euphonic harmonics. Secondly, the HEDD has a unique summing internal architecture that does not alter, truncate or recalculate the original source signal. The Cranesong clones the original source and sends that to its output, while mixing in the calculated distortion, thereby largely preserving the ambience and space of the original. The low level distortion in the figure is part of the additive distortion signal and not a result of recalculations to the source. In other words, only the distortion is distorted! We took this measurement first at 44.1 kHz; at 88.2 and 96 k. As you can see in the two figures on the next page, at 96 k the low level grunge is virtually gone, and the Cranesong's distortion is even cleaner, if that's not a contradiction in terms!

Cooking Better Sound—Naturally

There are certain analog consoles whose character is highly prized because they add spice, dimension and even punch to a mix. One name that comes to mind is API, which to my ears has an excellent combination of desirable linearities (like headroom and bandwidth) and nonlinearities. I think the subtle "grit" in their discrete opamps could even be slight intermodulation distortion, which does just the right thing for rock and roll yet is subtle enough for jazz and classical depending on how you drive the stages (a matter of taste). I think the transformers add some punch or fattening via saturation and 2nd and 3rd harmonic distortion as well as some upper harmonics and a touch of phase shift (which could add some *dimensionality*).

Our role as mastering engineer is like that of the master chef who knows just how much and what kind of spice is useful to add *pizzazz* without overcooking or spoiling the flavor. By the middle of our careers we have collected a sizable analog and digital spice rack! The Cranesong can mimic three types of naturally-occurring analog distortion, called **Triode, Pentode** and **Tape.** The **triode** control adds a pinch of **salt**, pure second harmonic, which, being the octave, is quite subtle, almost inaudible with some music. It can *clear up* the low end by adding some definition to a bass, but it can also thin out the sound too much. The **pentode** is extremely versatile; it provides both

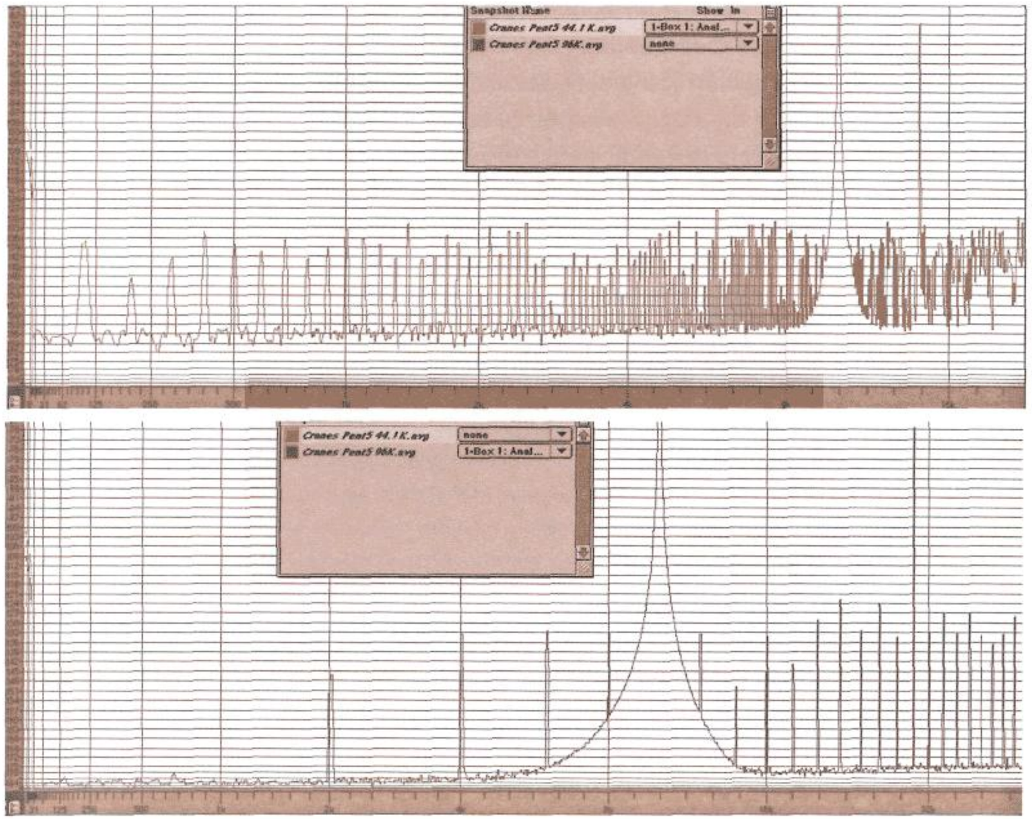
Clarity or Fuzz, which is best?

There's nothing wrong with using fuzz if it produces the right esthetic result. With high-resolution digital recording, tube equipment can add a nice flavor.

Or, it can be used as a useful cover-up, a *fuzzy band-aid*. A client once told me, "Bob, your mastering is so much clearer than the mix, I'm starting to hear all the mistakes!" Yes, high-resolution processing revealed more and more of the source, but this came at a price, all the warts were revealed. I solved the problem by fuzzing up the sound slightly with some delicate *tape style* harmonics.

For if the performance is not the absolute best, or the mix is not wonderful, or the sound is just better when it's not perfectly clear—then *fatness, masking fuzz, or analog distortion magic* may be just the right approach for the music. In mastering I usually prefer to accomplish this by first passing the signal through the highest resolution electronics, which add little or no distortion, and then add a touch of the fuzzy sauce with a selectively *fuzzy* component or a noisy dither. This approach is methodical, controllable, and reversible.

Clearly, artful use of noise can mask and therefore ameliorate some low-level distortions. Ironically, digital recording's super low-noise may be its greatest enemy.



Comparing Cranesong HEDD 192 in Pentode mode at two different sample rates with a 10 kHz −15 dBFS test tone.

At top, 44.1 kHz SR, at bottom, 96 kHz. Note the different frequency scales since the higher sample rate displays harmonic frequencies of the audio signal up to 48 kHz.

salt and *pepper*. At lower levels it adds third and fifth harmonics, which are dangerously seductive, producing a unique presence boost and brightness with little grunge or digititis, especially at 96 kHz SR (pictured). At higher levels, additional odd harmonics add grit and some fatness, like an overdriven pentode tube—a Marshal amplifier in a 1 U rack-mount box! Past the fifth, subtle amounts of seventh and ninth harmonics add a sometimes desirable "edge."

The Cranesong's **tape** control is the **sugar**, which when mixed in, can sweeten the pentode pepper, yielding flavors from red to yellow, green or Jalapeño! The celebrated third harmonic (an octave plus a fifth) sweetens and fattens the sound, much like analog tape. **Tape** also produces the fat sound of analog tape, which helps to "glue" a mix together. **Tape** can help digitally-mixed sources that may be well-recorded but miss some of that "rock and roll fatness." The control produces largely second and third harmonic distortion, but as it's advanced, some additional higher harmonics, emulating analog tape performance. Too much sugar gives slow, muddy molasses, a rarely desirable quantity, but available if you need it. But just a light amount can act as a sweet-sounding bandaid to ameliorate truncated or edgy recordings. Regardless, space and depth have been permanently lost if there was truncation prior to the use of the Cranesong.* No one is sure why, but critical listeners have observed that adding delicate amounts of harmonic distortion in just the right proportion appear to enhance the depth and clarity in a recording. The trick is to know the exact amount.[5]

Single Precision, Double Precision, or Floating Point?

First-generation digital processors gave digital processing a bad name. But single precision 24-bit processors are going the way of the Dodo, at least in respectable audio equipment. All things being equal

* Though Digital Domain's K-Stereo process does a pretty good job of restoring that lost ambience.

(and they never are) 32-bit floating point processors are generally regarded as inferior-sounding to 48-bit (double-precision fixed), and 40-bit float. Some newer floating-point devices, such as the software program **ChannelStrip** by Metric Halo, work in 64-bit and have impressively low measured distortion. However, one designer, Z-Systems, has produced a 32-bit floating point digital equalizer using proprietary distortion-reducing techniques that sounds very good and measures as well as some other equalizers using longer wordlengths. Ultimately the skill of the designer determines how nice the device sounds. The mathematics involved are not trivial, and the designer's choice of filter coefficients can make as much difference as his choice of wordlength.

Figure C16-09 in the Color Plates shows that with a single precision processor, even a simple gain boost can ruin your digital day. A dithered 24-bit 1 kHz tone at -11 dBFS is passed through two types of processors, each boosting gain by 10 dB. The distortion of the single precision processor (red trace) is the result of truncation of products below the 24th bit. Nevertheless, the highest distortion product, at -142 dBFS, is extremely low. I believe the *sound* of a single 24-bit truncation may not be audible, but cumulative truncation adds enough inharmonic distortion to become annoying to the sensitive ear. In blue we compare the perfectly clean output of a 40-bit floating point processor which dithers its output to 24 bits. I measured similar performance with a 48-bit (double precision) processor and 32-bit floating point processor, which both dither to 24 bits.

Double Sampling?

The most advanced digital equalizers and dynamics processors use double sampling technology, which means that the internal sampling rate is doubled to reduce aliasing distortion. High-quality linear phase filters are used in the internal sample rate converters. I'm not certain this has audible meaning for equalizers,[6] but dynamics processors benefit because non-linear processing generates severe aliases of the sampling rate, and the higher the sample rate, the less aliasing.

Figure C16-10 in the Color Plates compares two excellent-sounding digital dynamics processors, the oversampling Weiss DS1-MK2, which uses 40-bit floating point calculations, and the standard-sampling Waves L2, which uses 48-bit fixed point.

To compare apples to apples, both processors are limiting by 3 dB, with the Waves in red, and the Weiss in green, set to 1000:1 ratio. Note the oversampling processor exhibits considerably lower quantization distortion. However, the switchable safety limiter of the Weiss, which is not oversampled, produces considerable alias distortion even at 1 dB limiting (orange trace). At 88.2 kHz and above (not shown), the Weiss safety limiter and the Waves perform measurably better, and double sampling may not be needed. Thus there is considerable advantage of doing all our processing at higher rates, which moves the distortion products into the inaudible spectrum above 20 kHz. Then, sample rate convert to 44.1 kHz during the last step, which filters out most of the high-frequency by-products.

Despite the measured differences, the "window" we've chosen, (steady-state sinewave performance) probably has little to do with the perceived performance of these two excellent-sounding limiters. Because steady state measurements have little or no relationship to audible performance of limiters. I believe the key to the ear's reaction is the *duration* of the limiting action. In typical use, limiters go into gain reduction for a very short time. At limiting ratios of 1000:1, with instantaneous attack, and fast release, these processors produce only momentary distortion, shorter than the human ear's sensitivity to distortion (about 6 ms according to some authorities). But if a user overpushes a limiter so that it is working on the RMS levels of the material as well as the peaks, then its sinewave-measured distortion becomes audibly significant.

Compressors, however, are different animals, and double sampling is critical for them, because a compressor may be into gain reduction for a good percentage of the time. I feel that double-sampling contributes to the Weiss's *robust* and *warm* sound when used as a compressor. While Heavy Metal recordings employ considerable distortion for effect, classically they employ analog processors for this purpose to avoid the inharmonic aliases of typical digital processors.

Better Measurement Methods?

It should be clear by now that we can easily measure simple phenomena that are probably too subtle to hear (such as single tone harmonic distortion near the 24 bit level). But we can hear (perceive) very complex phenomena that are difficult to describe with measurements (such as the sound quality of one equalizer versus another). What we will need to better describe such complex audible phenomena are *psychoacoustically-based* measurement instruments that have not yet been invented. Current research and development of coded audio such as MP3 (that benefits from the ear's masking) could lead to better noise and distortion analysers that can discriminate between distortion we can and cannot hear.

The Bonger—A Listening Test

Since current steady-state sine-wave measurements are misleading when measuring nonlinear processors like compressors, a more effective measurement method is by listening: using the **gonger** aka **bonger**, originally developed by the BBC's Chris Travis and available on a test CD from Checkpoint Audio (see Appendix 10). This test is a pure sine wave that modulates through various amplitudes, in order to exercise and reveal any amplitude non-linearities in the signal path. Just play the bonger through the device under test and listen to the output for noise modulation, buzz or distortion.

Identity Testing—Bit Transparency

Any workstation that cannot make a perfect clone should be junked. The simplest test is the identity test, or bit-transparency test. Set a digital equalizer to flat and unity gain, then test to see if it passes signal identical to its input. Some people scoff at this test, since analog equipment almost never produces identical output. But the test is important, since digital equipment can produce egregious distortion as we have seen. The bit scope can aid in null testing: it is quite likely that a device is bit-transparent if you selectively put in 16 bits, then 20, then 24, and get out the same as you put in. You can also watch a 16 or 20-bit source expand to 24-bits when the gain changes, during crossfades, or if any equalizer is changed from the 0 dB position. A neutral console path is a good indication of data integrity in a DAW. After the bitscope, your next defense is to perform some basic tests, for linearity, for distortion with the FFT, and finally, test for perfect clones (perfect digital copies). The **null test** confirms bit-for bit identity: Play two files at the same time, inverting the polarity of one and mixing the two together. There must be zero output or the two files are not identical. Since designers are fallible human beings, you should carry out basic tests on your DAW for each software revision.

Choose Your Weapon

So, which to use, analog or digital processing? A few years ago, I didn't like the sound of cumulative digital processing. I could tolerate a couple of the best-designed single-precision units in series. After that, it was back to analog.

If processing digitally, be aware of the weaknesses of the equipment. Until manufacturers adopt more powerful processors, and processing power catches up, limit the number of passes through any digital system. Each pass will sound a little bit colder even using 24 bit storage. **A mix made through a current-day digital console may or may not sound better than one made through a high-quality analog console**, depending on several factors: the number of passes or bounces that have been made, the number of tracks which are mixed, the quality of the converters which were used, the outboard equipment, and the internal mixing and equalization algorithms in the digital console. While no console equalizer currently has the power of a $6000 Weiss, economically it's a lot simpler to replicate a good equalization algorithm for 144 channels than performing the equivalent in analog hardware, so there is hope for the digital console's future, when silicon will be cheaper.

"The Source Quality Rule: Always start out with the highest resolution source and maintain that resolution for as long as possible into the processing."

And there's no turning back; 24-bit recording and high sample rates are taking over, and they sound better, so for mastering we can choose from the best of several worlds, and we make our choices by balancing the benefits and the losses:

- (some) very transparent, low-noise, pure-sounding digital gear
- (some) good-sounding, reasonably-transparent, low-noise analog gear that we can use to add a little sugar, salt, pepper, or spice, or simply to prevent the sound from getting colder
- a digital processor that simulates analog distortion or warmth.

Why Is Good DSP So Expensive?

Intellectual property is the most nebulous thing to a consumer. It's easy to see why a two-ton

Mercedes Benz costs so much, but the amount of intellectual work that has gone into a one-gram IC is not so obvious. It can take five man-years to produce good audio software, created by individuals with ten or more years of schooling or experience. Similarly, when the doctor takes ten minutes to examine you, prescribes a 10-cent pill and then presents you with a $100 invoice, remember you're paying for all that knowledge and experience. This doesn't mean I'm against socialized medicine, I just want to re-emphasize the reasons why intellectual property and good DSP are so expensive.

The Source-Quality Rule

An important corollary of this discussion is the **source-quality rule:** *Source recordings and masters should have higher resolution than the eventual release medium.* **Always start out with the highest resolution source and maintain that resolution for as long as possible into the processing.** When mastering, one consequence of this rule is to reduce the number of generations and copies, and if possible, go back one or more generations when a new process must be added or applied.

This rule even applies when you're making an MP3 or other data-reduced final result. Consider a lossy medium like the (rapidly obsolescing) analog cassette. Dub to cassette from a high quality source, like a CD, and it sounds much better than a copy from an inferior source, like the FM radio, by avoiding cumulative bandwidth losses, as wider bandwidth sounds better. In other words, the higher the audio quality you begin with, the better the final product, whether it's an audiophile CD, a multi-media CD-ROM, MP3, or a talking Barbie doll. It may seem funny, but you'll never go wrong starting at 96 kHz/24 bit if the product is to end up on 44.1 k/16 bit CD. Sample rate conversion should be the penultimate process, followed by dithering.

In Summary

Mastering engineers do not have to think about the meaning of life every time they perform their magic; many engineers simply plug in their processors, listen, and make music sound better. But I also like to consider just why things sound better, because it helps me avoid problems that are not obvious at first listen, and also dream up innovative solutions. I hope that this chapter has inspired you to dream up some innovations of your own!

---

1 See the Appendix for references on noise filters. Ironically, all the standard noise-weighting filters should be revised, because they have no relationship with human perception of very quiet devices such as A/D and D/A converters.

2 And even then, the F-curve is an approximation, since the ear's perception of noise is much more than just a frequency response curve, as Jim Johnston explains: Noise should be measured separately in each critical band and compared to the ear's threshold for that critical band.

3 Most of the SpectraFoo™ screenshots were taken at an FFT resolution of 32K points (32000 "bins") with about 4 second average time and Hanning weighting. The actual amplitude of details on an FFT depends on its resolution, so FFTs are only directly comparable if the same methods are used.

4 The term *resolving*, when applied to the sound of tube circuits, is itself an unquantifiable audiophile subjective term. It's fair to say that audiophile negative reactions to some ugly-sounding solid-state circuits use inexact terms such as *resolution* and *transparency*, which may be proved to be simply distribution of harmonics or differences in frequency response. And maybe not!

5 For the curious, K-Stereo and K-Surround do **not** use harmonic distortion to enhance depth. They use other psychoacoustic principles.

6 Although the makers of the double-sampling Weiss Equalizer, GML plugin, and the Audiocube feel that double sampling is important for equalizers. Some engineers like the sound of high frequency curves that extend beyond 20 kHz, even if that is later cut off when the sample rate is halved at the output of the equalizer. And Jim Johnston (in correspondence) states that when a digital filter has response extending to half the sampling rate, it can produce some really odd and unexpected frequency responses, indicating that double sampling is important for such type of equalizers.