

Digital Audio Degradation

Studio and Consumer Environments

10/13/2008

MasterWork Sound Studios

Written by Thomas Deyesso



Every time an audio source is recorded, edited, replicated, and listened to, it undergoes signal degradation. In fact, the moment a sound is generated, either in an acoustic or digital environment, is the only time that the sound will ever be as “pure” as it is in that original state. Every stage that the captured sound moves through is just an additional generation of degradation it must be subjected to. Signal degradation is a topic most often ignored by amateur audio enthusiasts but is also ignored by some self-proclaimed, seasoned audio engineers. In this article, digital audio degradation will be discussed through the following subtopics: capturing audio, digital signal processing (now referred as DSP), media duplication, and end-consumer tampering.

I. Capturing audio

In order for an audio source to be recorded onto any type of medium, it must be captured through the use of a device. The most common device used to capture audio is a microphone, which is simply a device that converts acoustical energy to electrical energy. The construction and mechanics of microphones is outside the scope of this article, but it suffices to mention that depending on the build quality and type, microphones have a characteristic sound assigned to them. This is because they have been mechanically constructed for certain purposes in the studio (i.e. vocal, ambience, percussion, etc.); the different constructions mainly focus on the directionality of capture and frequency tailoring. However, even if a microphone is being used for the purpose assigned to it by its manufacturer or by industry standards, only a snapshot of the audio signal is being captured by the microphone due to the microphone’s physical constraints. Hopefully, the captured audio sounds as similar as possible to the original signal – that is in fact one of the goals of an audio engineer. Nevertheless, naturally, there are compromises along the way, and in this first stage of studio work, the audio signal has undergone its first generation of signal degradation.

It should also be noted, although not discussed at length here, that when a microphone converts acoustical energy to electrical energy, the electrical energy then passes through an analog to digital (A/D) converter. The A/D converter allows the digital audio workstation (DAW), which is usually a computer, to store a digital image of the audio signal in the form of digital bits. Like microphones, the build quality of an A/D converter can greatly affect its ability to convey audio information accurately. To avoid getting too deep into this topic, feel encouraged to look up terms such as phase-locked loop or jitter in the context of analog-to-digital conversion to discover more about the quality build of an A/D converter.

II. Digital signal processing

Once audio has been captured, it now passes onto the stage of editing and processing. During this stage, the audio is edited and mixed to suite the project and depending on the preference of the engineer, effects (reverb, compression, equalization, etc.) are now added. Some engineers prefer to add effects with outboard equipment before the DAW receives the audio, but as you may guess, the engineer is not able to change any effect settings unless the audio signal is re-captured with the preferred effect settings. Either way, while DSP can help achieve the goals of the engineer/artist for any given project, there is always a destructive component involved. These DSP operations serve as multiple generations of signal degradation.

This unavoidably destructive component in DSP can be largely explained through a discussion of bit depth and word length. In the DSP environment, a computer’s processor conducts all calculations to a pre-set point of precision, which is computed every time there is a sample in the signal (discussed later). To define this pre-set point of precision, bit depth is used, which is a measure of the number of binary digits used to compute a “digital word.” An audio signal that has a bit depth of 16 then has its digital words computed to 16 binary digits (and no more beyond that); likewise, an audio signal of bit depth 24 has its digital words computed to 24 binary digits, and therefore, has greater precision. When a higher bit depth is used, a computer’s processor is able to conduct every DSP operation (such as adding gain, reverb, or compression) to a longer word length relative to using a shorter bit depth. Practically speaking, every DSP operation yields some sonic degradation, because after conducting many DSP operations, the intended word length of the audio signal will be truncated to the actual word length of the audio signal thus altering the waveform of the audio signal in an unintended way. Over time, this cumulative effect can have negative, audible results but by using a longer word length, this degradation, which is really just data truncation, can be kept to a bare minimum. So, this generation of signal degradation is unavoidable but it can be reduced to some acceptable level.

In modern-day entertainment, the CD still reigns as the main medium for audio (although this is slowly changing); however, the bit depth of material cannot exceed the CD's mechanical limit for word length, which is 16 bit. It is unfortunately the practice and belief of some recording engineers to work only at 16 bit because that is the final destination word length. However, there are ways to achieve better sonic results – by working with material that is recorded with a relatively higher bit depth than the destination bit depth. Many engineers understand this principle, but do not totally understand the essential process of dithering that must be included.

When an engineer records at 24 bit and finishes his/her work and decides to truncate the material to 16 bit because the artist wishes to press CD's, all the additional information that was kept in the extra 8 bits of data, which is usually the ambience of the song, in the 24 bit recording are now lost. So, the hard work is just thrown out with the extra bits, and the signal's data integrity suffers for it. This is because when the signal is quantized, the bit numbers are rounded off without due regard to the original waveform and the engineer ends up with a truncated version of the signal with additional, often undesirable distortion. Consider this example:

A word length of 2 bits with the following values:

1.0	1.3	1.6	1.9	2.2	2.5	2.8	3.1
-----	-----	-----	-----	-----	-----	-----	-----

If these values were truncated to the lowest tenth-decimal value to be saved as one bit values, you would observe the following values:

1.0	1.0	1.0	1.0	2.0	2.0	2.0	3.0
-----	-----	-----	-----	-----	-----	-----	-----

Relative to the original signal, the truncated signal would now have additional frequency content different than the original. So, in order to resolve the word length close to the original word length values without actually have the extra bits to encode the original word length values, dither is used. The process of dithering uses a complex algorithm that actually incorporates random noise into the signal. Even though the word length is reduced when dither is used, the random noise algorithm, over the course of time, will average out the signal word length values to the original word length values. The key to understanding dither is that the new word lengths do not actually contain the original word length values – **they only average out to the original word length values over time**. So, not only will dither actually resolve a waveform to its higher bit value counterpart, it also lowers the noise floor by reducing distortion effects from adding incidental harmonics (ironically, by adding noise). This part of signal degradation is critical and an engineer will not be able to avoid it without a proper understanding of dither and when to use it.

Another component of audio that should be considered here is sampling rate. By definition, sampling rate is the rate at which an A/D converter samples the electrical signal it's receiving every second. So, an A/D converter of sample rate 44,100 Hz samples the signal 44,100 times every second. Currently, it is still being determined for the reasons why operating at higher sample rates tend to sound better than lower sample rates. Some believe the reason to be the increased bandwidth with higher sample rates even though any sample rate above the commonly used 44.1kHz only has additional frequency components that are inaudible to humans (humans have an upper frequency response limit around 20kHz). Others believe the reason to be the filter slope used in sample rate conversion. A converter filters out frequency content above its sampling rate capacity with a predesigned filter slope. Smoother slopes tend to sound better than sharp ones, but as you may guess, tend to cost much more in design and implementation. This aside, it has been determined that remastering material at higher sample rates yields better audible results. This is because the distortion of DSP operations during post-production, such as filtering and compression, will be spread out over a wider frequency bandwidth so much of the unwanted noise will be thrown out during downsampling. This wider bandwidth includes additional frequencies that are inaudible to humans so the negative effects of DSP operations are relatively small.

III. Media duplication

After an audio project has been recorded, mixed, and mastered and requires no further processing, it is usually duplicated onto a hard medium in order to be delivered to its audience. There are a great number of different formats that audio can be placed upon, and the processes inherent to each different format placement vary. Usually, media duplication is handled by a duplication/replication plant that produces the audio project in mass quantities, but in recent

years, it is not uncommon for the musical artist to duplicate their project themselves. Due to contemporary relevance, the scope of this section will only cover the CD and DVD formats, while formats such as vinyl and tape will be ignored.

There are two ways in which data can be written to both CD's and DVD's. Most commonly, audio data is usually "written" to the disc medium with an optical laser that transcribes the corresponding data to the surface from either a master disc or hard drive. The other method used in large replication plants is the stamping method. Without going into long detail, the stamping method operates as follows: a "father" disc is created from the master CD or DVD. This father disc is the inverse of the master disc and thus, can be used as the primary stamper for small duplication runs, but usually is used to create yet another "mother disc," which is the inverse of the father disc. "Children" stampers are then made from the mother disc and it is these stampers that will do the majority of the replication. The reason these master discs are called "stampers" is because that is exactly what they do: they are mechanically pressed up against a large number of blank discs in order to impart a number of "pits" and "lands" into the blank discs (see figure 1). The pits and lands are the physical property of a disc that an optical disc reader uses to translate into an electrical signal, which is used for the translation of the audio.

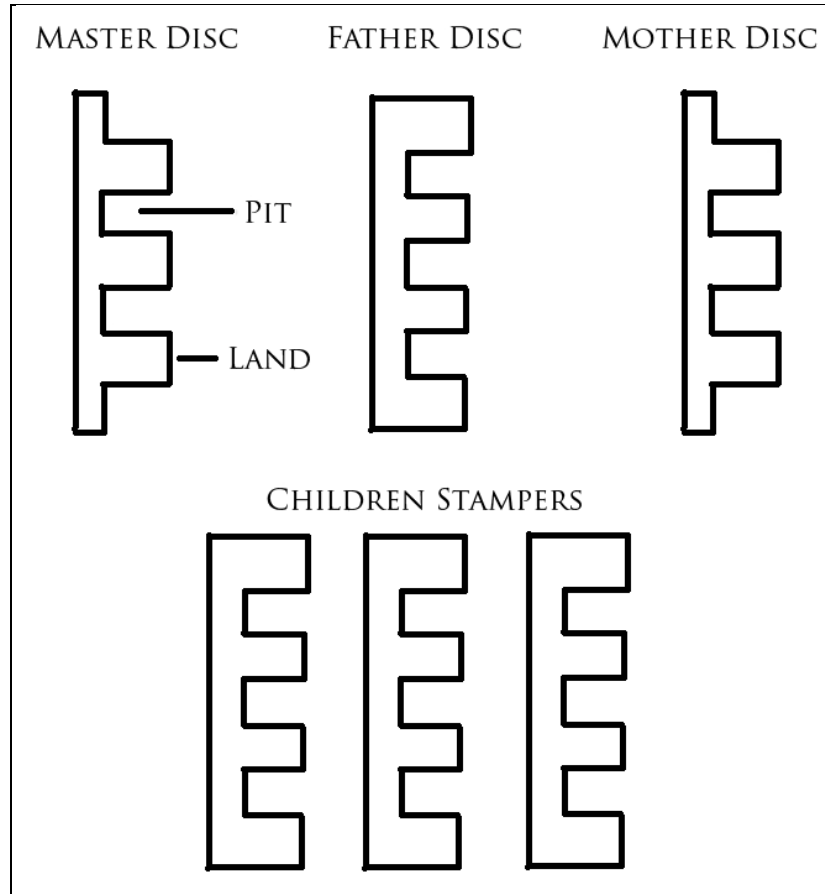


Figure 1: Pits and lands in the various discs used in the stamping replication method. Notice that the father disc has the inverse lands and pits as the master disc, the mother disc has the same pits and lands as the master disc, and that the children stampers has the same pits and lands as the father disc. When a children stamper comes in contact with a blank disc, the blank disc is imparted with the same pits and lands as the master disc, which is the entire point of this replication method!

With that tedious explanation behind us now, it must be stated that in both data-writing methods, there is a chance of signal degradation. With laser beam recorders, there will absolutely be errors in the writing process if the disc is even slightly bent, moist, covered with fingerprints or hair, etc. prior or during the writing process. Furthermore, writing errors can occur especially at higher writing speeds (very common in the consumer market) and the ill-fated E32 error, which is an uncorrectable error during the writing process, can occur. Fortunately, master CD's and DVD's are almost always written in the mastering house where superb optical drives are used to write data, error detection software is used, and there is an overall high level of quality control. Also, needless to say, if there is any type of malfunction with the optical writing drive involving the writing laser, off-center disc spinning, or any burst/short-out of

the data transmission, there will be signal degradation. In the stamping method for replication, there are strict guidelines to follow to avoid any signal degradation during the replication process. Needless to say, if there is an error in the father disc, there will be a corresponding error in the mother disc and subsequent children stampers. So, to prevent the addition of errors by the replication facility, the stamping process is done in an extremely clean room where the personnel operating the equipment even wear full body suits – this way, there can be assurance that there will be no human addition of error.

IV. End-consumer tampering

Arguably the greatest source of signal degradation of digital audio involves the final destination of any audio project – the hands of the consumer. Much of the actions of the consumer on audio are performed haphazardly and without knowledge about the consequences of their actions. However, the education of the general public about this topic simply isn't present and your average audio consumer wouldn't know any better unless they either do some research on digital audio mechanics.

This stage of signal degradation can come in many different forms but for the most part, it involves a sort of downsampling of the audio content. For many, the internet is the greatest access source for multimedia, but oddly enough, many hosting servers and internet service providers are not adequately equipped currently to serve the general public with any semblance of high definition content within a reasonable streaming/downloading time frame. So, the general public responded to this in a number of ways. Websites who wish to host audio content will usually downsample audio from a high resolution format (WAV) to a low resolution format (MP3) in order to be able to stream content to users with as small time delay as possible. While performing this action is certainly one of necessity and utility, the audio is totally compromised and even sounds to an untrained ear as inferior because it is being truncated to a lesser refined version of its higher resolution counterpart. Examples of these websites include user-based networking communities like myspace.com and youtube.com, but there are even websites that have artist-based networked such as purevolume.com. Furthermore, audio is downsampled to lower resolution formats to yield fast transfer times between users. While I won't make any comment on the ethics of music piracy, users who illegally download/upload audio aren't even taking their illegal practice seriously. If a consumer is going to pirate audio from artists, they should at least try to preserve the audio in its original form without any resolution downsampling.

Outside the internet-centric multimedia domain, there is still a great amount of signal degradation occurring in the hands of consumers. When a consumer buys a CD, there are only a few places at most that they will be listening to the audio: their car's audio system, their personal computer's audio system, and perhaps a home stereo system. When someone listens to audio in a car, there are a number of contributing factors to signal degradation. The first of which is the poor reproduction of the audio using low quality electronics in many consumer level CD players and poorly constructed speakers. However, this isn't the worst factor contributing to signal degradation. There is an exorbitant amount of outside noise added to a car's listening environment while driving so that much of the audio content is masked in this noise. Also, the stereo image is totally lop-sided no matter where you sit in a car. So, as you can see, there is quite a lot of signal degradation while auditioning audio in a car. If someone bought a hard copy of CD/DVD project, then they might wish to put this audio content on their computer as a backup or even to serve as their primary listening environment. Usually, when a consumer transports the audio to their computer, they have already played the medium in an optical drive so there is the possibility that the optical drive has damaged the disc before they have gotten the chance to transport the audio. Also, the playback system usually incorporated into a home computer is an extremely poor environment to reproduce any audio because much of the low frequency content is usually rolled off by the small diameter of commonly used, small speakers and because consumers are rarely aware of how to place their speakers to yield an effective stereo image. When a CD/DVD is placed in a home stereo system, the medium is of course subjected to the same electronic degradation it would face in a car's media player. The speaker system is quite different in this environment however. Usually, there is a sort of excitation equalization applied in any home stereo system just to

make the audio sound “extra special.” You’re lucky if the manufacturer even tells you about it much less give you the option to disable it. Also, the speakers provided with the system are tailored much in the same way enhancing certain frequencies based on the physical mechanics and frequency response of the speakers. So, signal degradation occurs in this environment mainly because the playback of audio is “frequency flavored” and does not even come close to reproducing audio the way an artist intended.

V. Conclusion

So, what do audio professionals and general consumers do with this knowledge? Well, there are many things that audio professionals should be responsible for. Firstly, any recording engineer with experience and a working knowledge of his or her trade will be aware of what type of microphone to use for a specific recording application in order to fully capture the sound for the intent of the application. Furthermore, they should understand how important analog-to-digital conversion is and purchase interfaces accordingly. Also, even though audio professionals often misunderstand the topic, the practice of using high sample rates during recording, mixing, and pre-mastering is a good way of minimizing signal degradation due to DSP operations; they should also avoid any unnecessary signal processing as they will be adding unintended harmonic content to the audio program. The general public who downsample audio for internet use should be aware of their actions’ consequences. Hopefully, sometime in the near future, there will be no more need to downsample audio at all for internet use. Over time, server storage space and internet bandwidth will continue to expand thus making low resolution audio formats such as the MP3 obsolete. However, in the mean time, audio professionals should be aware of how to responsibly use the audio production tools set before them, and the general public should seek to educate themselves more on audio formats and how to preserve the highest file quality over time.

All topics presented here are put into practice at MasterWork Sound Studios. Not only do we upsample all audio worked on in order to minimize noise from DSP operations, we avoid any unnecessary DSP operation that would otherwise degrade audio. We also use a variety of tools and processes to ensure that our operations are fully transparent and non-destructive. You can be assured that whenever you have your audio project mastered by MasterWork Sound Studios that you are receiving the highest-fidelity product possible with the fullest attention to what is considered project-appropriate processing to make your audio sound the best it possibly can.

Property of Masterwork Sound Studios

Masterwork Sound Studios
56 Waterman Street
East Bridgewater, MA, 02333

main@masterworksound.com
www.masterworksound.com