

Digital Dynamics Processing: It's All In The "Samples!"

*By Frank Foti
November, 2000*

Introduction

One of the most emotional topics for discussion in broadcasting is audio processing. Why? Well, it's dynamic in so many ways! Not just in the actual level adjusting action of the device itself, but the plethora of viewpoints, opinions and philosophies about this combination of technology and "black magic". I must admit that as a former chief engineer and now designer of this stuff, I too can easily become embroiled in a good chat about processing. From the Telos/Omnia R&D offices via the Internet, I've shared many a thought about this exciting "personality enhancement" to many radio stations. So it stands to reason that the topic we are about to visit will certainly draw a few comments from the many passionate folks who study, design, adjust, and work with audio processing.

This paper is an outcome from our recent development work on our newest processor, the **Omnia-6fm**. No, this is not a sales related piece about all of the new "doo-dads" and "whatcha-ma-call-it" functions. It's a short essay on the key element of dynamics processing...sampling. We have found that this sole function, in digital audio, can be the basis of difference between a merely acceptable design and an excellent one. If I may modestly mention here, this is probably one of the elements that has led to the world-leading success of our **Omnia.fm** product. I'll say nothing more on that, as I'm sure you know the story by now. But this issue of sampling came to light for us again in this latest development phase, and I'm sharing some of that new information with all of you here in this forum.

Sampling Rate: A Fast Review...

If you're experienced with digital audio, I suggest you skip this section. If not, please read on. In any digital audio system, the basis for audio data acquisition and transportation is related to something known as *sampling rate*. Simply defined, this is a function by which the audio signal is sampled and measured very quickly at a specified rate. This is usually referred to in the form of a frequency, or the rate of the samples as they are acquired. Hence the term *sampling rate*.

Now, for audio signals, it has been determined that to reliably capture the entire audio spectrum so as to reproduce it with low distortion, the minimum sampling rate should be at least twice the rate of the highest known frequency. By example, our hearing range is known to be 20Hz – 20kHz. So the lowest sampling rate for full range audio should be at least 40kHz. This (2x) of the highest audio signal is known as the *Nyquist frequency*. The compact disc audio player employs a sampling rate of 44.1kHz. In the professional audio world, this is the lowest rate used to obtain full range audio. Most recording studios employ 48kHz or even 96kHz for recording and mastering purposes.

This essay is not written to debate the merits of which rate is optimum for pure aural purposes. There are many excellent papers on the topic of higher-sampling rates for improved sonic performance that may interest the reader, and we have listed several on our website (www.omniaaudio.com). This paper rather, will focus on how sampling rate affects the operation of a dynamics processor. This issue of sampling rate is, in fact, very influential to other functions besides the audio signal itself, which will become more apparent soon!

Aliasing Distortion: The Enemy Of Digital Audio

Within the digital audio system the frequency spectrum *must* be contained to precisely one-half of the Nyquist frequency. Why is this important? Since the sampling rate defines the limited range of audio spectrum, any signal that tries to exceed this range will run out of spectrum! When this occurs, the excess frequency will *reflect*, and travel back down into the audio spectrum as a distortion component. This is known as *aliasing distortion*. For instance, in a 48kHz sampled system, any signal that tries to exceed 24kHz will reflect itself back down into the audio domain.

Example: Let's say that in our 48kHz system, some form of audio harmonic of, say, 15kHz is present in the source (such as the second harmonic of 30kHz). Since 30kHz is 6kHz beyond the system's reproducible maximum of 24kHz, aliasing product at 18kHz is generated from the 6kHz of *reflection* that occurs due to lack of spectrum. This is proven by subtracting the difference between the desired 30kHz signal and the 24kHz Nyquist frequency, which results in a "leftover" 6kHz. Further subtracting 6kHz from 24kHz yields the aliasing component at 18kHz.

Generally speaking, normal digital audio systems like CD players and DAT machines have nothing to worry about. But any digital system that has the potential to create harmonic energy within the system must be careful of creating aliasing distortion. This is possible within a dynamics processor, as there are numerous functions (hard limiters, for instance) that generate signal spectrum which will try to exceed the Nyquist frequency. Just as a clipper creates harmonic products in an analog design, the same holds true in the digital domain. Even if a low pass filter is applied, it will not suppress harmonics generated *within* the digital system which try to exceed the Nyquist limitation. Our previous design efforts have proved that we are the only audio processor manufacturer to observe this, and to take extra pains in designing our digital processing system so that added processing distortions were not generated in the digital domain.

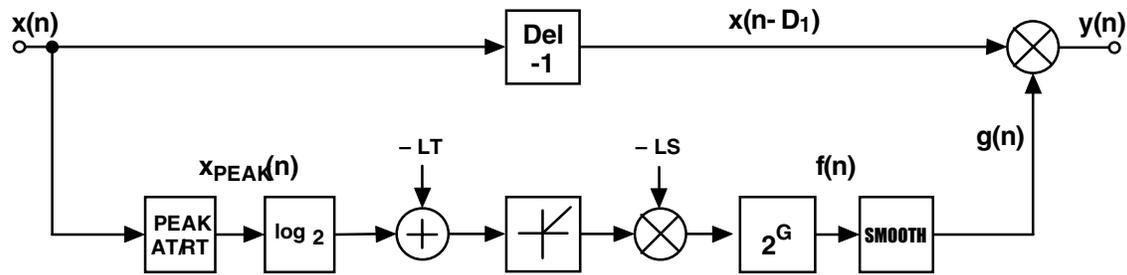
This is the basis of this discussion.

Digital Processors: More Than Just Clean Clipping!

The record speaks for itself about how our products perform with respect to the final limiting ("clipping") function. Through musical program, as well as dynamic technical testing, we have proven our theories correct with respect to the design of a non-aliasing digital hard limiter/clipper. For us this is old news. What is newsworthy is our detection of what occurs in dynamics control when processing generated distortion. Our analysis of such is at the root of this discussion (and our position on which has been supported by another digital audio researcher as well¹).

To better understand this phase of the topic, we must look at how a dynamics controller such as a compressor or limiter operates. Simply stated, the audio signal is passed through a gain function block that will change or manipulate the gain by means of a *control signal*. It is the action of this control signal that we are most concerned about. Following is a simplified block diagram of a feed-forward peak limiter:

¹ Mapes-Riordan, D.: *A Worst-Case Analysis for Analog-Quality (Alias-Free) Digital Dynamics Processing*, 105th Convention of the Audio Engineering Society (AES), San Francisco 1998, Preprint 4766



The audio signal appears at the input node $x(n)$. It is passed through to a delay block ($DEL-1$) and then onto one side of a multiply function where it appears as a result of $x(n-D1)$. Additionally a *control signal* is created in the feed-forward section that appears as $g(n)$ and is attached to the other portion of the multiply function, thereby creating a dynamic gain control where the audio signal $x(n-D1)$ is multiplied by $g(n)$ which yields the resultant level as represented as $y(n)$. For our discussion here, the critical signal is the control node which is illustrated as $g(n)$.

It is this signal that determines not only the output level of the dynamic function, but also the aural texture of how the function block will sound. To relate this concept to an analog design, think of it as the operation of a four-quadrant multiplier component, or VCA (voltage controlled amplifier). As anyone who has designed a gain controller will tell you, this signal must be generated with care or additional processing related distortions and artifacts will occur. Naturally there will be some sonic tradeoffs, as this function of the multiplier is to affect the dynamic gain function. But if the control signal reacts too abruptly, or not in a *smooth* fashion, IMD (Intermodulation Distortion) will be easily heard in the signal. This is what is commonly described as a *busy or squashed* sound.

Within the digital design we must be careful of all these parameters, plus one additional component: our old nemesis *aliasing distortion*. How can this occur? The answer lies in the characteristics of control signal $g(n)$. To understand the aspects of this, we must first consider what can transpire in this signal, especially when employed as a dynamic peak limiter.

The *control signal* is comprised of functions that must respond to the audio envelope. For a peak limiter, there are qualities of the algorithm where the initial transient response of a short duration peak signal are either recognized or ignored. In the event that the limiter algorithm reacts to the transient information, it must do so in a manner that does not try to respond faster than the sampling rate of the limiter. This is very critical to the upper third portion of the audio spectrum; basically, above 5kHz for FM Stereo transmission. Transient audio signals in this range can respond very quickly. Should they trigger action in a peak control algorithm that is faster than the Nyquist frequency of the sampling rate, aliasing components will result in the control signal $g(n)$.

While most limiters are designed to operate with a normalized set of time constants that are well within the sampling rate of the system, most broadcast processors employ some added multiple timing functions that are not user controlled. They are designed to operate on the short duration peaks, and reduce them in a sub-audible manner. But should the action of these added time constants result in response that approaches the Nyquist frequency of the system, aliasing distortion will result.

Due to the complex non-linearity of these signals, they are not easily measured. To do so would take careful analysis with predictable tone bursts of high frequency signals, and a very fast storage scope and audio analyzer. They are actually much easier to model in a computerized synthesis of the algorithm. Suffice to say that any additional high-speed dynamics components which try to become part of the control signal will yield aliasing, if the sampling rate is not large enough. Our research has shown conclusively that *some* aliasing will result even in a 48kHz sampled system; therefore *any* system which employs sampling at a lower rate – say 32kHz sampling for dynamics control – is bound to cause aliasing, even in the peak limiters. Any designer who contends that sampling rate for high frequency control does not matter has failed to

grasp this uniquely important aspect of processing. In listening tests comparing low-sampling rate controlled processes with higher rates, listeners unanimously choose the higher rates. The difference is very obvious, even in casual listening.

Sampling Rate Matters A Great Deal!!

Our research has shown that to properly generate a dynamics limiter control signal void of *any* digital artifacts, the sampling rate should be a minimum of 96kHz. This is three times the rate of the 32kHz sampled systems known to cause aliasing distortion.

Realize also that the section of audio processing we are discussing is the area *before* the final limiter/clipper. Should aliasing distortion be generated and added to the audio signal, it is impossible to remove. Therefore, it does not matter if the final limiter employs a very large sampling rate, for if the signal prior to final limiting has been contaminated with aliasing artifacts, damage in the final stage is unavoidable and irreversible.

The new **Omnia-6fm** avoids any possible contamination by employing 96kHz sampling in the upper bands of the limiting section. That, coupled with our proprietary non-aliasing clipper method, allows clean, smooth and crisp reproduction of presence and high frequencies. This is quite critical for managing the steep boost of the preemphasis curve (where 15kHz is amplified by 17dB for 75 μ s in North America, and 50 μ s in the rest of the world).

Conclusion

With a new generation of digital processors coming to market, there will no doubt be much discussion about sampling rate once again. Some will tell you that using a lower sampling rate conserves machine cycles, then turn around and waste these same saved cycles needlessly in an over-sampled clipper. In meetings with a number of the Telos/Omnia DSP engineers, we examined this approach most carefully to ensure that we were not missing anything. Our unanimous conclusion is that the proponents of limited sampling rate control processors – not us! – have indeed missed something. Something very important to the sound of audio processing.

As a designer, it doesn't really matter how we employ the available DSP cycles. As long as we provide you with a great sounding product, and at a decent cost, who cares how many cycles or DSP chips we use? Did it matter how many transistors were in your first portable radio all those years ago? No, it didn't! All that mattered was the way it *sounded*. That axiom held true during the early days of transistor radios, and it's still true today with DSP technology.

Anyone who claims that 32kHz sampling is sufficient for FM Stereo dynamics audio processing is way behind the times. Sampling rate matters. It's the basis of our *digital audio spectrum*, and the more samples we have, the better the resulting performance. This issue rings true again and again as research into digital audio strives to raise the bar. And it's no different in the dynamics domain.

Happy listening...

Frank Foti
Omnia