MAKING AUDIO SOUND BETTER
ONE SQUARE WAVE AT A TIME

Leif Claesson

Omnia Audio, Cleveland, OH USA

ABSTRACT

Audio mastering engineers have felt increasing pressure over the years to master recordings at ever increasing loudness levels as compared to other contemporary recordings, by way of dynamic compression, peak limiting, and hard clipping. This pursuit of loudness adds distortion, reduces fidelity and ultimately impairs audio quality.

When radio stations play the already compromised audio from contemporary music through their typical FM processing chains, this confluence of degradation causes serious audio quality issues on air. Besides leading to listener fatigue, this can even be mistaken for poor reception by listeners, ultimately causing tune-out.

When pre-distorted contemporary recordings are encoded through typical perceptual codecs, the codecs end up wasting bits trying to faithfully encode the distortion components, in lieu of the original audio.

This paper shall examine what the music typically endures when broadcast on FM, and how that understanding led to the invention of the "undo" algorithm, which is applicable to both the broadcasting and recording industries, and automatically repairs some or most of the damage caused by these mastering techniques by adaptively de-clipping and de-compressing the mastered recordings.Insert abstract text

THE FM LOUDNESS WARS

For the past three decades, FM broadcasters have been engaged in what have become known as the "loudness wars", using ever-advancing FM audio processors to make their stations sound louder than the competition on the dial while still (presumably) operating within legal modulation limits.

During an A/B comparison between two different audio sources, listeners tend to prefer the louder of the two, even if audio quality is somewhat reduced. Even loudness differences as small as 0.1dB, which are imperceptible as a whole, can still impart differences to fullness, warmth, and clarity, which the listeners in turn perceive as a more favorable sound.

FM modulation imposes physical and electrical limitations which limit peak waveform excursion. If additional loudness is desired when the signal is already at its maximum, something has to give. In the case of an FM signal, what gives is the top of the waveform, which is quite literally sheared off in a process known as "clipping".

Depending upon the particular processor employed, the use of clipping can provide additional loudness without excessively degrading the audio, or can increase loudness but at the cost of introducing audible artifacts and distortion.

To a certain extent, then, one can see where the FM loudness wars – at least in their infancy – were somewhat understandable: Stations wanted to be perceived as clear and powerful to make an impression when scanning up and down the dial. As is so often the case with such things, this power was quickly abused as stations continued to increase the amount of processing and chased one another around in an effort to be the absolute loudest signal in the market with no regard for audio quality.

LOUDNESS COMES TO THE STUDIO

Like an FM modulated signal, other analog and digital mediums have similar limitations that dictate waveform excursion. During the process of recording, mastering, or mixing records, various techniques are used to reduce the peak-to-average ratios of signals such as dynamics compression, limiting, overdriving analog-to-digital converters, and hard digital clipping.

While these techniques do not cause significant problems when used judiciously, the loudness wars have crept into the studio over the past 10 to 15 years, particularly in the recording of popular music, which now has virtually no dynamic range. Instead of music going from quiet to loud it goes from clean to distorted with a resulting brick-shaped waveform. Such recordings have a busy, flat, and lifeless sound right out of the studio, a situation that is only made worse when the content passes through a typical FM air chain where it undergoes further dynamic range compression and clipping. The result is nearly unlistenable audio that invites listeners to tune out.

A look back at the 80’s

The easiest way to see the progression of this problem is to examine some waveforms from older and more modern recordings. Below is a kick drum transient from the original release of Dire Straits’ “Money for Nothing” from 1985. The lines are at 0dB full scale.

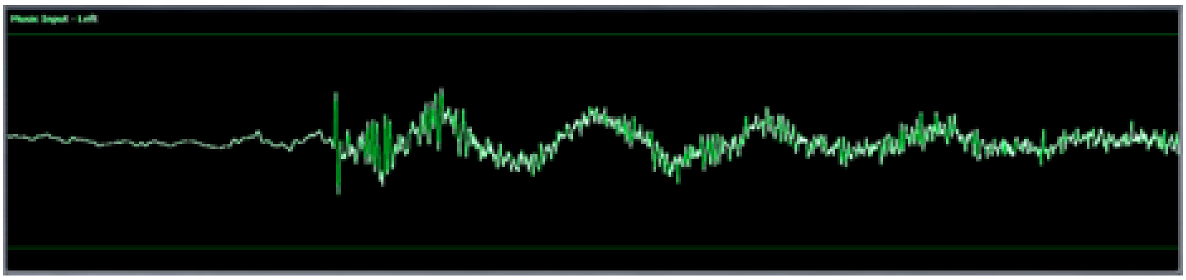

Figure 1 – “Money for Nothing” – Dire Straits - 1985

Notice that no one seemed concerned about maximizing volume; after all, the industry had just gained a 30dB improvement in signal-to-noise ratio afforded by the recent transition to the compact disc, offering headroom they could use - or not. Provided you turn up the volume control, this recording still sounds fantastic.

Ten years later

Next, we see Alanis Morissette's "Ironic" from 1995. Here, the waveform is maximized without being modified. This is the point at which we ran out of headroom on CDs and around the last time we heard a mainstream CD that was sonically unaffected by the mastering loudness wars.

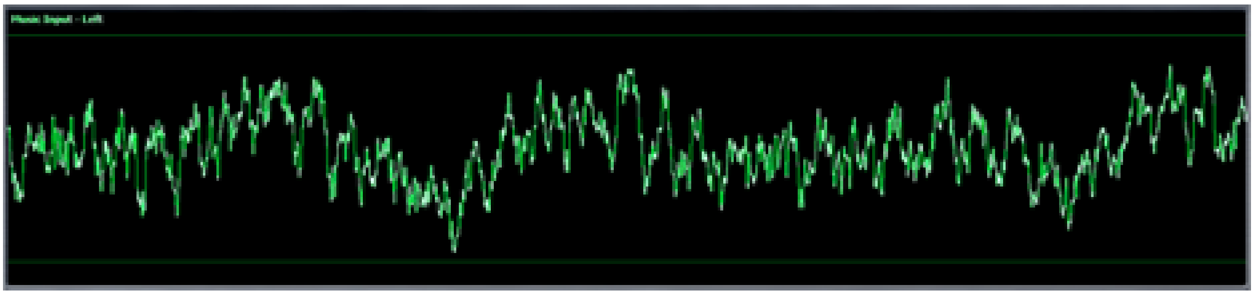

Figure 2 – "Ironic" – Alanis Morissette - 1995

Trouble ahead

1999 brought forth the Red Hot Chili Pepper's *Californication* album. It doesn't look very clipped by today's standards, but suffered sonically and was recorded without any limiting or distortion-cancelled clipping, or oversampled clipping to prevent aliasing. It was simply turned up digitally and overloaded. The Red Hot Chili Peppers may have been pioneers in this type of mastering, but by 2003, most of the music industry followed.

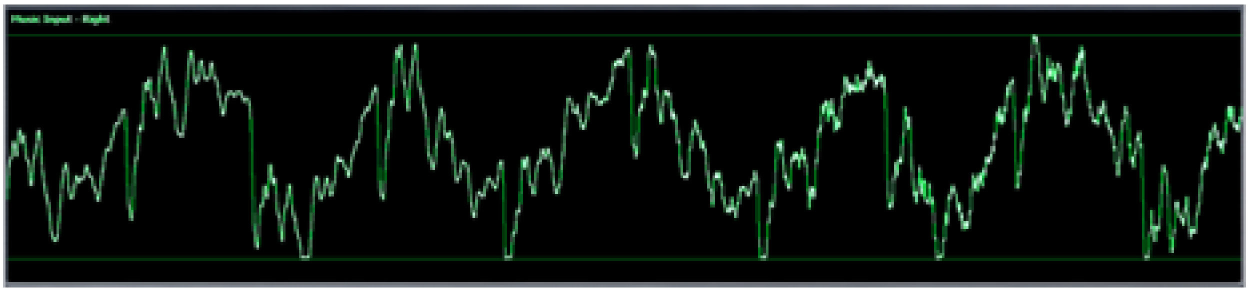

Figure 3 – "Californication" – Red Hot Chili Peppers - 1999

New millennium, new problems

Next, we see Green Day's "Wake Me Up When September Ends", in which every kick drumbeat is a square wave. This changes the sound of the drum from "kick" to "crunch". At first it distorts the sound, but the damage doesn't stop there. Sound waves are additive and if slammed into the rails, something has to give; in this case, what gives is the rest of the audio, which is forced to mute for the duration of the kick drum.

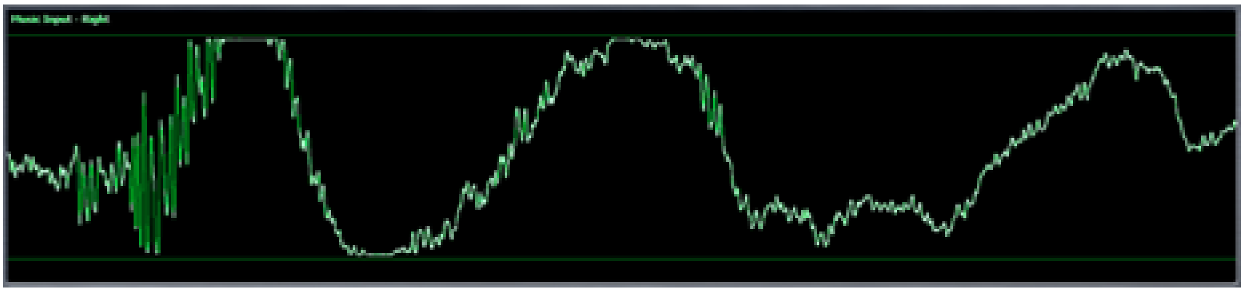

Figure 4 – “Wake Me Up When September Ends” – Green Day – 2004

What happens on the radio?

What happens when this recording is played through a typical FM air chain? Due to phase rotation and non-phase linear filters, the result is clipped edges that don’t stay put at the edges any longer and can end up in any portion of the waveform, including through the zero line as seen below. The clipping from the original waveform is still clearly visible, but nowhere near the edges. A flat-top at the edge would have at least helped contribute to on-air loudness while preventing further distortion downstream, as there would have been no peak for the final clipper to clip.

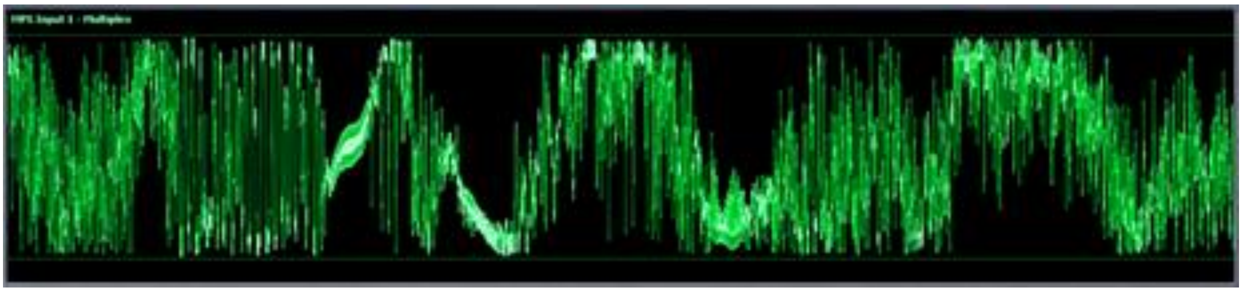

Figure 5 – “Wake Me Up When September Ends” – Green Day – 2004 – after FM processing

This subject was examined in some depth in a 2001 AES paper entitled “What Happens to My Recording When it's Played on the Radio?” co-authored by Frank Foti and Robert Orban. At the time, these two heavyweights in the field of broadcast audio processing were bitter rivals and agreed on very little, but managed to un-don their boxing gloves long enough to write a paper together to attempt to curb the misguided notion that increased loudness in mastering would translate to increased loudness on the air.

And the winner is…

The undisputed winner of the mastering loudness wars, though, is Metallica’s Death Magnetic album from 2008 which featured the song “The Day That Never Comes”, seen below. Notice that not only are the peaks clipped, but the flat tops aren’t exactly flat. This suggests that the clipping was done during mixing, before the mastering engineer received it. The only tool left at his disposal was a (non-phase linear) equalizer.

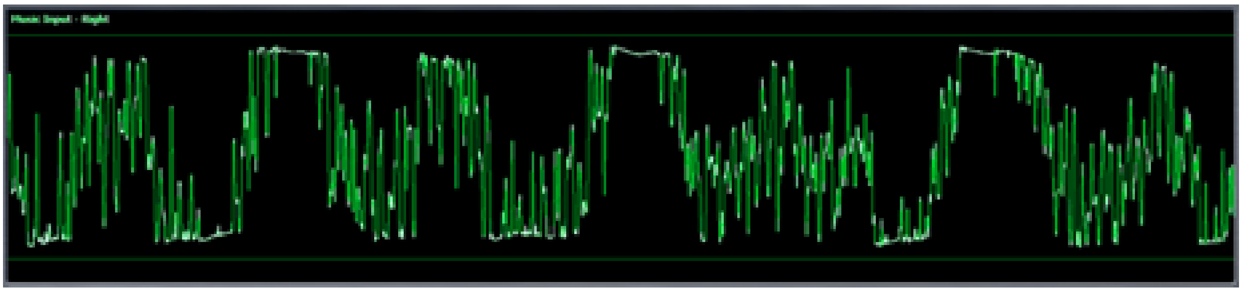

Figure 6 – “The Day That Never Comes” – Metallica - 2008

HOW “UNDO” REPAIRS AUDIO

Although the ultimate solution lies in education and improved recording practices, the “Undo” algorithm introduced in the Omnia.9 processor is an important stop-gap measure as it increases listenability of existing damaged recordings by removing clipping distortion, reducing density and increasing fidelity.

The two highest priority design criteria for the Undo algorithm were that it must be a fully automatic “set and forget” implementation, and it must not over-enhance well-made recordings.

Overview of Undo

Undo is actually a two-step process: A de-clipper and a multiband expander. De-clipping must be performed prior to any other processing, and subsequent processing inevitably changes the shape of the waveform making effective de-clipping difficult if not impossible.

Step 1 – the de-clipper

The de-clipper must first identify which parts of the waveform are intact and which are damaged or missing. This is accomplished by statistical analysis, comparing every part of the waveform against the current peak level, where the parts closest to the current peak level are more likely to be damaged than areas closer to the center of the waveform.

Since clipping is a permanent, non-reversible process, de-clipping is really a process of peak restoration that essentially involves identifying the gaps in the waveform and recreating the peak to fill in the gap. Simply drawing a line or curve to bridge the gap is insufficient, and can actually increase distortion, as it would not take into account things like a high frequency, low amplitude instrument (like a flute) being driven into the clipper by a low frequency, high amplitude instrument (like a bass guitar).

Instead, it is necessary to analyze which frequencies were present before and after the gap and bridge it by recreating the missing part of the waveform accordingly. The recreated segment is compared to the original and is used in lieu of the original only if its amplitude exceeds that of the original signal as well as yields reduced overall spectral complexity. Thus, distortion is automatically removed while ensuring that intact portions of the waveform are unaffected and distortion is never added.

Step 2- the multiband expanders

After the distortion has been removed and the clipped peaks have been re-created, the audio is sub-divided into five frequency bands and examined by short-term dynamics detectors on each band. This analysis is mostly achieved by deriving the peak-to-RMS average ratio, but this metric alone is insufficient as it can be fooled by static, spectrally sparse background sounds that our ears “hear through” even though their RMS energy content is high.

For example, using peak-to-RMS measurements of the high frequencies of the synth lead a couple of minutes into the Donald Fagen song “I.G.Y.”, one of the most dynamic pop recordings ever made, measures as devoid of dynamics as does “The Day That Never Comes” by Metallica. Instead, the dynamics detectors use a psychoacoustic model that takes into account both levels and spectral content to mimic how the human ear and brain perceive sound.

The dynamics detectors inversely set the ratio of multiband upward expanders (causing more expansion when less dynamics are present at the input, to a point) while the threshold of the expanders are kept at a fixed offset from the current peak level of the original audio in each band. The result is that the upward expanders essentially “de-compress” the music, undoing peak- limiting and compression on more dense material while mostly leaving well-recorded material alone.

BEFORE AND AFTER EXAMPLES

Below are some before and after examples to illustrate the difference between the original waveforms and those processed by Undo:

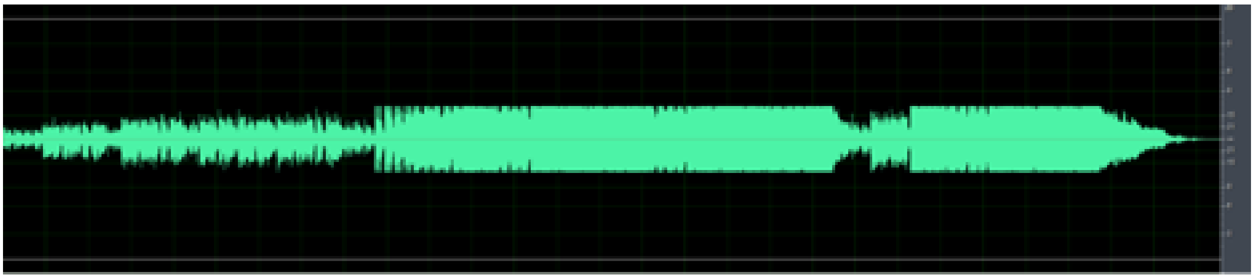

Figure 7 – “Wake Me Up When September Ends” – Green Day – 2004 – CD Release

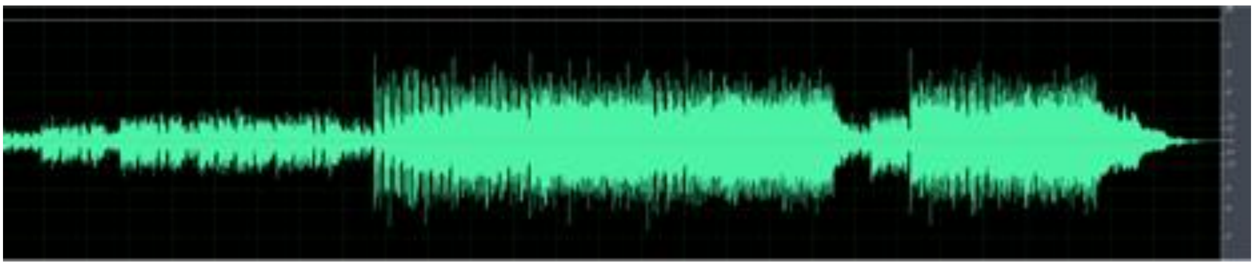

Figure 8 – “Wake Me Up When September Ends” – Green Day – 2004 – after being repaired by Undo

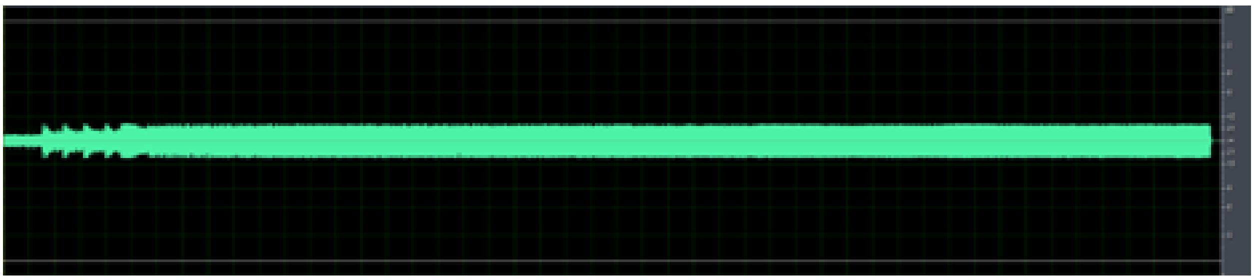

Figure 9 – “The Day That Never Comes” – Metallica – 2008 – CD Release

Figure 10– “The Day That Never Comes” – Metallica – 2008 –
after being repaired by Undo

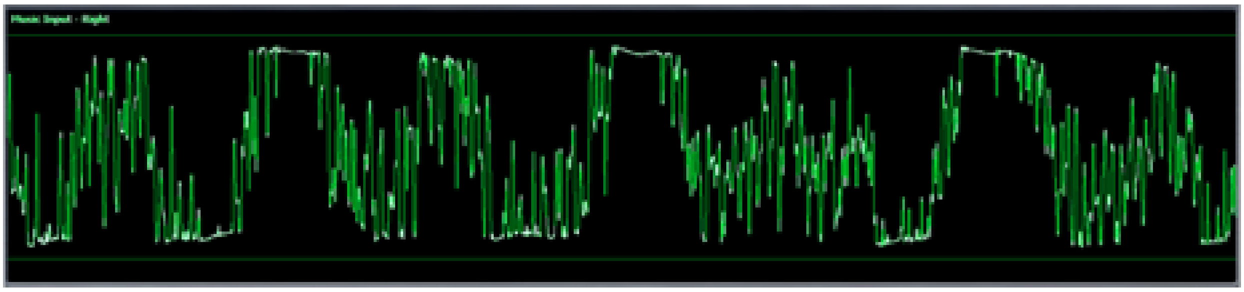

Figure 11– “The Day That Never Comes” – Metallica – 2008 –
Waveform detail CD release

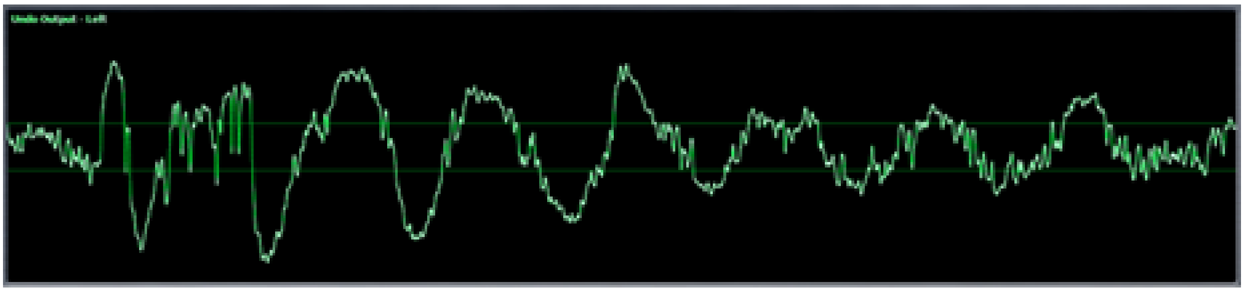

Figure 12– “The Day That Never Comes” – Metallica – 2008 –
Waveform detail after being repaired by Undo (note lines for scale

SUMMARY

In conclusion, while the ideal solution lies in a return to mixing and mastering techniques that take advantage of the headroom provided by the digital domain without compromises to quality, Undo offers a highly effective means by which to repair existing recordings, whether intentionally or unintentionally distorted.

ACKNOWLEDGEMENTS

This work was supported by Jim Kuzman, Omnia Applications Engineer, and Mary Ann Seidler, Omnia Sales.