

What Happens to My Recording When it's Played on the Radio?

*Frank Foti, Omnia Audio
Robert Orban, CRL/Orban*

June, 2001

Few people in the record industry really know how a radio station processes their material before it hits the FM airwaves. This article's purpose is to remove the many myths and misconceptions surrounding this arcane art.

Every radio station uses a transmission audio processor in front of its transmitter. The processor's most important function is to control the peak modulation of the transmitter to the legal requirements of the regulatory body in each station's nation. However, very few stations use a simple peak limiter for this function. Instead, they use more complex audio chains. These can accurately constrain peak modulation while significantly decreasing the peak-to-average ratio of the audio. This makes the station sound louder within the allowable peak modulation.

Garbage In—Garbage Out

Manufacturers have tuned broadcast processors to process the clean, dynamic program material that the recording industry has typically released throughout its history. (The only significant exception that comes to mind is 45-rpm singles, which often were overtly distorted.) Because these processors have to process speech, commercials, and oldies in addition to current material, they can't be tuned exclusively for "hypercompressed," distorted CDs. Indeed, experience has shown that there's no way to tune them successfully for this degraded material.

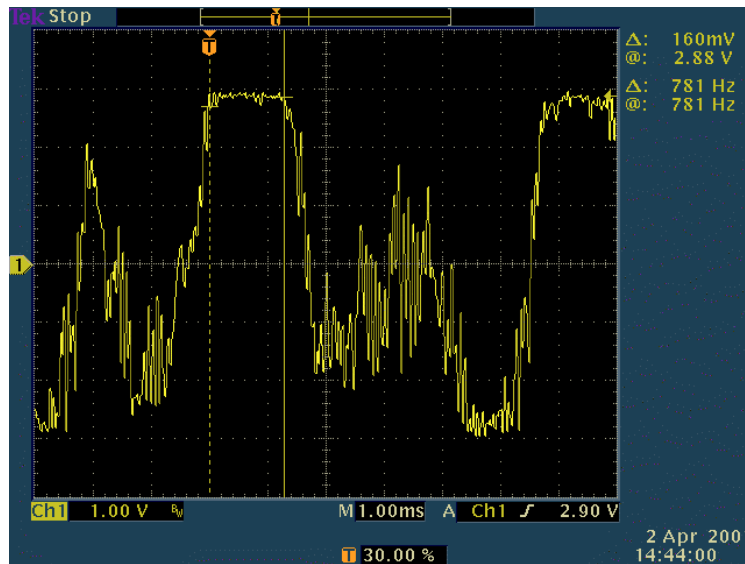
For 20 years, broadcast processor designers have known that achieving highest loudness consistent with maximum punch and cleanliness requires extremely clean source material. For more than 20 years, Orban has published application notes to help broadcast engineers clean up their signal paths. These notes emphasize that any clipping in the path before the processor will cause subtle degradation that the processor will often exaggerate severely. The notes promote adequate headroom and low distortion amplification to prevent clipping even when an operator drives the meters into the red.

About three years ago, we started to notice CDs arriving at radio stations that had been pre-distorted in production or mastering to increase their loudness. For the first time, we started seeing frequently reoccurring flat topping caused by brute-force clipping in the production process. Broadcast processors react to pre-distorted CDs exactly the same way as they have reacted to accidentally clipped material for more than 20 years—they exaggerate the distortion. Because of phase rotation, the source clipping never increases on-air loudness—it just adds grunge.

The authors understand the reasoning behind the CD loudness wars. Just as radio stations wish to offer the loudest signal on the dial, it is evident that recording artists, producers, and even some

record labels want to have a loud product that stands out against its competition in a CD changer or a music store's listening station.

In radio broadcasting this competition has existed for at least the last 25 years. 25 years ago, radio stations used simple clipping to get louder, and this 25-year-old technique has now migrated to the music industry. The following graphic shows a section of a severely clipped waveform from a contemporary CD. The area marked between the two pointers highlights the clipped portion. This is one of the roots of the problem as described in this paper; the other is excessive digital limiting that does not necessarily cause flat-topping, but still removes transient punch and impact from the sound.



The problem today is that we now have sophisticated and powerful audio processing for the broadcast transmission system and this processing does not coexist well with a signal that has already been severely clipped. Unfortunately, with current pop CDs, the example shown above is more the norm than the exception.

The attack and release characteristics of broadcast multiband compression were tuned to sound natural with source material having short-term peak-to-average ratios typical of vinyl or pre-1990 CDs. Excessive digital limiting of the source material radically reduces this short-term peak-to-average ratio and presents the broadcast processor with a new, synthetic type of source that the broadcast processor handles less gracefully and naturally than it handles older material. Instead of being punchy, the on-air sound produced from these hypercompressed sources is small and flat, without the dynamic contours that give music its dramatic impact. The on-air sound resembles musical wallpaper and makes the listener want to turn down the volume control to background levels.

There is a myth that broadcast processing will affect hypercompressed material less than it will more naturally produced material. This is true in only one aspect—if there is no long-term dynamic range coming in, then the broadcast processor's AGC will not further reduce it. However, the broadcast processor will still operate on the short-term envelopes of hypercompressed material and will further reduce the peak-to-average ratio, degrading the sound even more.

Hypercompressed material does not sound louder on the air. It sounds more distorted, making the radio sound broken in extreme cases. It sounds small, busy, and flat. It does not feel good to the listener when turned up, so he or she hears it as background music. Hypercompression, when combined with “major-market” levels of broadcast processing, sucks the drama and life from music. In more extreme cases, it sounds overtly distorted and is likely to cause tune-outs by adults, particularly women.

A Typical Processing Chain—What Really Goes On When Your Recording is Broadcast:

A typical chain consists of the following elements, in the order that they appear in the chain:

Phase rotator

The phase rotator is a chain of allpass filters (typically four poles, all at 200Hz) whose group delay is very non-constant as a function of frequency. Many voice waveforms (particularly male voices) exhibit as much as 6dB asymmetry. The phase rotator makes voice waveforms more symmetrical and can sometimes reduce the peak-to-average ratio of voice by 3-4dB. Because this processing is linear (it adds no new frequencies to the spectrum, so it doesn't sound raspy or fuzzy) it's the closest thing to a “free lunch” that one gets in the world of transmission processing.

There are a few prices to pay. In the good old days when source material wasn't grossly clipped, the main price was a very subtle reduction in transparency and definition in music. This was widely accepted as a valid trade-off to achieve greatly reduced speech distortion, because the phase rotator's effects on music are unlikely to be heard on typical consumer radios, like car radios, boomboxes, “Walkman”-style portables, and table radios.

However, with the rise of the clipped CD, things have changed. The phase rotator radically changes the shape of its input waveform without changing its frequency balance: If you measured the frequency response of the phase rotator, it would measure “flat” unless you also measured phase response, in which case you would say that the “magnitude response” was flat and the phase response was highly non-linear with frequency. The practical effect of this non-linear phase response is that flat tops in the original signal can end up *anywhere* in the waveform after processing. It's common to see them go right through a zero crossing. They end up looking like little smooth sections of the waveform where all the detail is missing—a bit like a scar from a severe burn. This is an apt metaphor for their audible effect, because they no longer help reduce the peak-to-average ratio of the waveform. Instead, their only effect is to add unnecessary grungy distortion.

There has been a myth in the recording world that broadcast processing will modify these clipped, over-compressed CDs less it will modify clean, dynamic CDs. Thanks in part to phase rotation, this myth is absolutely false. In particular, any clipping in the source material causes nothing but added distortion without increasing on-air loudness at all.

AGC

The next stage is usually an average-responding AGC. By recording studio standards, this AGC is required to operate over a very wide dynamic range—typically in the range of 25dB. Its function is to compensate for operator errors (in live production environments) and for varying average

levels (in automated environments). Average levels vary mainly because the peak to average ratio of CDs themselves has varied so much in the last 10 years or so. Therefore, normalizing hard disk recordings (to use all available headroom) has the undesirable side effect of causing gross variations in average levels. Indeed, 1:1 transfers (which are also common) will also exhibit this variation, which can be as large as 15dB.

The price to be paid is simple: the AGC will eliminate long-term dynamics in your recording. Virtually all radio station program directors want their stations to stay loud always, eliminating the risk that someone tuning the radio to their station will either miss the station completely or will think that it's weak and can't be received satisfactorily. Radio people often call this effect "dropping off the dial."

AGCs can be either single-band or multiband. If they are multiband, it's rare to use more than two bands because AGCs operate slowly, so "spectral gain intermodulation" (such as bass' pumping the midrange) is not as big a potential problem as it is for later compression stages, which operate more quickly.

AGCs are always gated in competent processors. This means that their gain essentially freezes if the input drops below a preset threshold, preventing noise suck-up despite the large amount of gain reduction.

Stereo Enhancement

Not all processors implement stereo enhancement, and those that do may implement it somewhere other than after the AGC. (In fact, stand-alone stereo enhancers are often placed in the program line in front of the transmission processor.)

The common purpose of stereo enhancement is to make the signal stand out dramatically when the car radio listener punches the tuning button. It's a technique to make the sound bigger and more dramatic. Overdone, it can remix the recording. Assuming that stereo reverb, with considerable L-R energy, was used in the original mix, stereo enhancement, for example, can change the amount of reverb applied to a center-channel vocalist. The moral? When mixing for broadcast, err on the "dry" side, because some stations' processors will bring the reverb more to the foreground.

Because each manufacturer uses a different technique for stereo enhancement, it's impossible to generalize about it. The only universal constraints are the need for strict mono compatibility (because FM radio is frequently received in mono, even on "stereo" radios, due to signal-quality-triggered mono blend circuitry), and the requirement that the stereo difference signal (L-R) not be enhanced excessively. Excessive enhancement always increases multipath distortion (because the part of the FM stereo signal that carries the L-R information is more vulnerable to multipath). Excessive enhancement will also reduce the loudness of the transmission (because of the "interleaving" properties of the FM stereo composite waveform, which we won't further discuss).

These constraints mean that recording-studio-style stereo enhancement is often incompatible with FM broadcast, particularly if it significantly increases average L-R levels. In the days of vinyl, a similar constraint existed because of the need to prevent the cutter head from lifting off the lacquer, but with CDs, this constraint no longer exists. Nevertheless, any mix intended for airplay will yield the lowest distortion and highest loudness at the receiver if its L-R/L+R ratio is low. Ironically, mono is loudest and cleanest!

Equalization

Equalization may be as simple as a fixed-frequency bass boost, or as complex as a multi-stage parametric equalizer. EQ has two purposes in a broadcast processor. The first is to establish a signature for a given radio station that brands the station by creating a “house sound.” The second purpose is to compensate for the frequency contouring caused by the subsequent multiband dynamics processing and high frequency limiting. These may create an overall spectral coloration that can be corrected or augmented by carefully chosen fixed EQ before the multiband dynamics stages.

Multiband Compression and Limiting

Depending on the manufacturer, this may occur in one or two stages. If it occurs in two stages, the multiband compressor and limiter can have different crossovers and even different numbers of bands. If it occurs in one stage, the compressor and limiter functions can “talk” to each other, optimizing their interaction. Both design approaches can yield good sound and each has its own set of tradeoffs.

Usually using anywhere between four and six bands, the multiband compressor/limiter reduces dynamic range and increases audio density to achieve competitive loudness and dial impact. It’s common for each band to be gated at low levels to prevent noise rush-up, and manufacturers often have proprietary algorithms for doing this while minimizing the audible side effects of the gating.

An advanced processor may have dozens of setup controls to tune just the multiband compressor/limiter. Drive and output gain controls for the various compressors, attack and release time controls, thresholds, and sometimes crossover frequencies are adjustable, depending on the processor design. Each of these controls has its own effect on the sound, and an operator needs extensive experience if he or she is to tune a broadcast multiband compressor so that it sounds good on a wide variety of program material without constant readjustment. Unlike mastering in the record industry, in broadcast there’s no mastering engineer available to optimize the processing for each new source!

Pre-Emphasis and HF Limiting

FM radio is pre-emphasized at 50 microseconds or 75 microseconds, depending on the country in which the transmission occurs. Pre-emphasis is a 6dB/octave high frequency boost that’s 3dB up at 2.1kHz (75 μ s) or 3.2kHz (50 μ s). With 75 μ s pre-emphasis, 15kHz is up 17dB!

Depending on the processor’s manufacturer, pre-emphasis may be applied before or after the multiband compressor/limiter. The important thing for mixers and mastering engineers to understand is that putting lots of energy above 5kHz creates significant problems for any broadcast processor because the pre-emphasis will greatly increase this energy. To prevent loudness loss, the processor applies high frequency limiting to these boosted high frequencies. HF limiting may cause the sound to become dull, distorted, or both, in various combinations. One of the most important differences between competing processors is how effectively a given processor performs HF limiting to minimize audible side effects. In state-of-the-art processors, HF limiting is usually performed partially by HF gain reduction and partially by distortion-cancelled clipping.

Clipping

In most processors, the clipping stage is the primary means of peak limiting. It's crucial to broadcast processor performance. Because of the FM pre-emphasis, simple clipping doesn't work well at all. It produces difference-frequency IM distortion, which the de-emphasis in the radio then exaggerates. (The de-emphasis is flat below 2-3kHz, but rolls off at 6dB/octave thereafter, effectively exaggerating energy below 2-3kHz.) The result is particularly offensive on cymbals and sibilance ("esses" become "effs").

In the late seventies, one of the authors of this article (R.O.) invented distortion-cancelled clipping. This manipulates the distortion spectrum added by the clipper's action. In FM, it typically removes the clipper-induced distortion below 2kHz (the flat part of the receiver's frequency response). This typically adds about 1dB to the peak level emerging from the clipper, but, in exchange, allows the clipper to be driven much harder than would otherwise be possible.

Provided that it doesn't introduce audibly offensive distortion, distortion-cancelled clipping is a very effective means of peak limiting because it affects only the peaks that actually exceed the clipping threshold and not surrounding material. Accordingly, clipping does not cause pumping, which gain reduction can do, particularly when gain reduction operates on pre-emphasized material. Clipping also causes minimal HF loss by comparison to HF limiting that uses gain reduction. For these reasons, most FM broadcast processors use the maximum practical amount of clipping that's consistent with acceptably low audible distortion.

Real-world clipping systems can get very complicated because of the requirement to strictly band-limit the clipped signal to less than 19kHz despite the harmonics that clipping adds to the signal. (Bandlimiting prevents aliasing between the stereo main and subchannel, protects subcarriers located above 55kHz in the FM stereo composite baseband, and protects the stereo pilot tone at 19kHz). Linearly filtering the clipped signal to remove energy above 15kHz causes large overshoots (up to 6dB in worst case) because of a combination of spectral truncation and time dispersion in the filter. Even a phase-linear lowpass filter (practical only in DSP realizations) causes up to 2dB overshoot. Therefore, state-of-the-art processors use complex overshoot compensation schemes to reduce peaks without significantly adding out-of-band spectrum.

Some chains also apply composite clipping or limiting to the output of the stereo encoder. The stereo encoder is the circuit that encodes the left and right channels into the single multiplex signal that drives the transmitter, and it's actually the peak level of this signal that government broadcasting authorities regulate. Composite clipping or limiting has long been a controversial technique, but the latest generation of composite clippers or limiters has greatly reduced the interference problems characteristic of earlier technology.

Conclusions

Broadcast processing is complex and sophisticated, and was tuned for the recordings produced using practices typical of the recording industry during almost all of its history. In this historical context, hypercompression is a short-term anomaly and does not coexist well with the "competitive" processing that most pop-music radio stations use. We therefore recommend that record companies provide broadcasters with radio mixes. These can have all of the equalization, slow compression, and other effects that producers and mastering engineers use artistically to achieve a desired "sound." What these radio mixes should not have is fast digital limiting and clipping. Leave the short-term envelopes unsquashed. Let the broadcast processor do its work.

The result will be just as loud on-air as hypercompressed material, but will have far more punch, clarity, and life.

A second recommendation to the record industry is to employ studio or mastering processing that provides the desired sonic effect, but without the undesired extreme distortion component that clipping creates. The alternative to brute-force clipping is digital look-ahead limiting, which is already widely available to the recording industry from a number of different manufacturers (including the authors' companies). This processing creates lower modulation distortion than clipping and also avoids blatant flat-topping of waveforms. Compared to clipping, it is therefore substantially more compatible with broadcast processing. Nevertheless, even digital limiting can have a deleterious effect on sound quality by reducing the peak-to-average ratio of the signal to the point that the broadcast processor responds to it in an unnatural way, so it should be used conservatively. Ultimately, the only way to tell how one's production processing will interact with a broadcast processor is to actually apply the processed signal to a real-world broadcast processor and to listen to its output, preferably through a typical consumer radio.