Wheatstone
WHITE PAPER

Audio Transfer Through The Internet

By Jeffrey A. Keith, CPBE, NCE
Senior Product Development Engineer

Wheatstone Corporation
May 2015

Podcasting Audio

If you are podcasting audio content, chances are that audio is going to go through a codec, and there's a lot you can do to optimize your signal for a successful, good sounding transfer through the internet, which stands between you and your audience.

The Codec

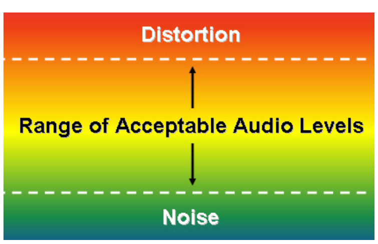
The codec's job is to convert the audio signal into a manageably sized file. It does this by reducing the amount of data normally present in a full spectrum audio signal and passing on the remaining information. This is accomplished by analyzing the audio data and throwing away things your listeners probably wouldn't hear anyway. The codec keeps on doing this until there's just enough of the original data left to reconstitute manageable, listenable audio. One way you can help the codec do its job is to never hand it things to code that either aren't relevant to your audio, or might be difficult for it to turn into good sounding data.

My own "Top Five" for making sure a codec is ingesting only the things it should be are:

- **Don't try to send 'Hi-Fi' content to 'Lo-Fi' listening environments**
- **Avoid sending the codec sustained noises such as hiss and hum**
- **Minimize any errors in left/right channel balance**
- **Avoid distortions due to badly clipped audio**
- **Avoid overly boosted high frequencies**

Even though the codec discards most of the original audio data, things should still sound pretty good if our expectations remain reasonable and we don't do anything to make its job even harder than it already is.

FIGURE 1



Prepping The Audio For Your Codec

A good signal processor can greatly help a codec do its job. The first task that a signal processor feeding a codec must do is manage the dynamic range of the program content—but do so in ways that listeners won't notice. For this to work, you must ensure that audio levels going into the processor remain within a certain range, in order to minimize the influence of noise and distortion. This range is illustrated in Figure 1, where dashed lines represent the minimum and maximum audio levels that should be allowed.

Next, the audio is processed to remove long-term level variations. In Wheatstone signal processing this is accomplished by an intelligent level control algorithm specifically designed to hide any hint to the listener that audio level modifications have occurred. Done with care and finesse, this step virtually eliminates any possibility that your podcast audio will be a source of listener fatigue.

Next, specialized algorithms operating in several frequency bands condition the audio to enhance the consistency of its loudness and intelligibility. This task isn't trivial and involves several important steps that shouldn't be skipped. These include an automatic rebalancing of short-term levels and corrections to the medium-term spectral balance, to create audio that is consistent and easy to listen to. As in the long-term level control previously mentioned, the processor should be able to perform these tasks on all kinds of program material and do so without a listener ever noticing that something has been done. In point of fact, this is the processing stage that does most of the heavy lifting to ensure that a podcast audio file or web stream remains listenable, even at very low volume settings or in noisy environments.

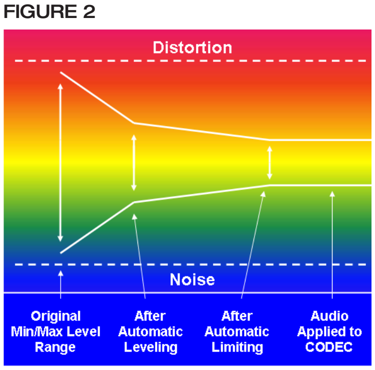
The final two stages of the processor are the most sophisticated. Though these two stages usually operate independently, on some program material they behave as a unified algorithm. The nuances of this processing section are critically important to the quality of the final audio feed, because there are two mutually exclusive tasks at work: controlling peak audio levels to prevent overload within the codec, and predicting and mitigating certain kinds of audio content that would prove troublesome for the codec.

Note it is this final processing section that separates 'stream' processing from traditional 'on-air' processing. Pre-codec processing is a completely different animal than on-air processing. The importance of this difference cannot be overstated!

It may seem that the front ends of both processors are similar, and to an extent this is true. However, what may not be intuitive is that the two processors' back ends are, and need to be, completely different in order to appropriately perform their tasks.

The primary purpose of on-air processing is to control modulation. The primary purpose of pre-codec processing is to manage what the codec 'sees' audio-wise, as it turns your program content into a drastically reduced number of digital bits that become the audio your audience ultimately hears.

FIGURE 2



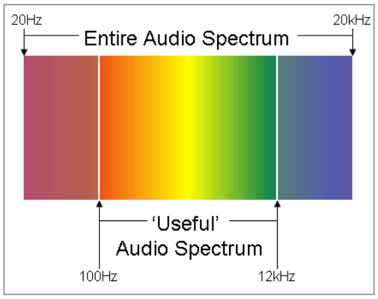
Looking Deeper

Audio consistency and minimal coding artifacts help make listeners comfortable listening to your audio content (and may even entice them to come back for more). Although other processing functions are involved, we have described in detail only the management of audio levels. Figure 2 is a simplified view of how these levels are managed inside the pre-codec processor.

Figure 3 (next page) hints that audio frequencies beyond the reproduction range of a listener's equipment can consume codec resources better reserved for coding audio that everyone can hear. Several informal surveys we have performed reveal that most internet audio listening is done on devices or equipment with 'Lo-Fi' specifications, often in environments with substantial background noise. Does it make sense then to hand the codec audio content that will never be heard by the majority of your listeners? In a word: No.

Removing program content that can't be heard on the devices and systems listeners use can dramatically improve the subjective quality of your audio.

FIGURE 3



As Figure 3 shows, audio frequencies below about 100 Hz and above about 12 kHz can usually be reduced (or in some cases completely removed) with no negative impact on the perceived quality of your content. In fact, doing so can actually help your audio sound better, due to reduced or removed 'codec teasers' such as hiss and hum.

Finally, the lower the codec's bitrate (i.e., the smaller the final audio file) the fewer bits there are for coding your program's audio. Seemingly unimportant issues such as left/right channel balance can become critical at low codec bitrates. Therefore, below about 48kb/s I recommend using mono operation, because although the stream may no longer be 'in stereo' it will certainly have a much higher apparent quality to your listeners.

Beyond the Codec...

Beyond choosing the codec to use you must also choose an audio processor capable of preconditioning your content so it always sounds great to your listener. Plus, you'll need a mixer designed for broadcast - one that will give you the control you need in a live production environment. Naturally, I feel Wheatstone's own designs are ideal for this.

Audioarts 08 Audio Mixer

http://audioartsengineering.com/index.php/audioarts-08-console-overview

In one compact frame, the Audioarts 08 covers all the basics: a single stereo mixing bus (balanced or unbalanced output); two mic inputs (one for guest, and one for host); USB input to play in audio from a PC; USB output to record directly to PC recording software; mix-minus telephone output for interfacing to a telephone hybrid; monitor, headphone and cue for off-air monitoring; automatic speaker mute for silencing monitor speakers when the mic is on (to prevent feedback); and unbalanced or balanced inputs and outputs for interface with consumer or professional grade equipment.

Wheatstone Audio Processing

http://wheatstone-processing.com

Wheatstone audio processing combines intelligent multiband AGC with advanced final limiter technologies specifically designed for their target mediums. The synergistic combination of these technologies automatically adapts to incoming source material, creating consistent, spectrally balanced program density while also managing loudness. Our designs also offer smart stereo enhancement, sophisticated bass management and high-performance distortion mitigation to give listeners a pleasing and compelling audio experience.