Dialnorm - Broadcaster Compliance With The Calm Act

In December of 2011, Congress enacted and the President signed the "Calm Act" regulating perceived loudness of programming at broadcast facilities in the U.S. This regulation -based on the ATSC A/85 Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television – has caused a great deal of confusion in the market place. The key to understanding this legislation is the phrase "Perceived Loudness". Perceived loudness goes beyond simple VU gain values and AGC control. It also takes into account audio processing such as compression, expansion, gating, limiting, etc., that increases perceived loudness in program audio without increasing gain. The purpose of this legislation is to maintain consistent audio loudness to the consumer, both between channels and between programming and commercial content. Broadcasters, satellite providers, cable operators and other multi channel content providers have until December 2012 to comply. Ensemble Designs has products that can help broadcasters maintain compliance with the Calm Act.

Perceived loudness compliance is based on Dialog Normalization – dialnorm. Dialnorm is defined in ATSC A/85 as

"An AC-3 metadata parameter, numerically equal to the absolute value of the Dialog Level, carried in the AC-3 bit stream. This unsigned 5-bit code indicates how far the average Dialog Level is below 0 LKFS. Valid values are 1-31. (zero value is reserved) The dialnorm values of 1 to 31 are interpreted as -1 to -31 LKFS."

To many, LKFS is a relatively new term, which means "Loudness K-weighted relative to Full Scale." It's a scale for audio measurement similar to VU or Peak. However, rather than measuring gain, it's measuring loudness. LKFS is based on the ITU-R BS.1770 Loudness Measurement Method. The ITU-R group performed tests to groups of people around the world with various content and listening environments and were able to construct an algorithm that accurately measured the loudness of audio content. It was quite an undertaking, and the details of the process take up the majority of the BS.1770 document. Have a look sometime. It covers everything from speaker placement to noise in the environment, to the type of content used for the testing. The loudness unit of LKFS is dB and is used the same way as a dB of gain. For example, a -15 LKFS program can be made to match the loudness of a -22 LKFS program by attenuating 7 dB.

There is a defined process for determining the dialnorm value for a particular piece of content. It's a fairly involved process depending on whether dialog is present and the length of the sample that must be used to determine the dialnorm value for the entire program. It can be found in the ATSC A/85 document if you'd like more detail. Dialnorm relates to the level of an "Anchor Element" which is usually dialog. It disregards momentary intentional loud elements such as gunshots or car crashes. Here's another way to think about this. When you are home watching TV, what do you adjust the volume to hear? It's usually the talking part of the program. Sound effects may be louder, whispering softer. But when you adjust the TV volume, you're setting it to comfortably hear what's being spoken. This is essentially dialnorm. If there is no dialog present, as in music programming, it relates to the viewer focus element. For example: the level of the featured pianist rather than the level of the entire orchestra. A single dialnorm value is determined for the entire program content, and is embedded into the metadata bit stream of the SDI signal. To simplify the process, companies such as Dolby, Sony, Tektronix and others, have a box that will "listen" to the program content and then report a dialnorm value for that piece of programming. Another box is often used to encode this value into the SDI stream. File based content solutions are also available from companies such as Telestream and Masstech which analyze a particular file's content and then embeds a dialnorm into the metadata of the file. All this is well and good for content that is already produced, but what about live productions such as sports or news? For this type of programming, a dialnorm target is selected. This is usually specified by network specs or distributor standard practices. The audio technician then mixes the audio content to the target dialnorm using LKFS metering. The dialnorm metadata (value) is entered and encoded into the SDI stream on the fly.

The AC-3 audio system defined in the ATSC Digital Television Standard uses dialnorm metadata to control loudness and other audio parameters more effectively without permanently altering the dynamic range of the content. The content provider or DTV operator encodes metadata (the dialnorm value) along with the audio content into an AC-3 encoder. This metadata parameter, when extracted at the decoder, sets different content to a uniform loudness transparently. Basically, it provides results similar to the viewer using his remote control to set a comfortable volume between disparate TV programs, commercials and channels.

So, what does this mean to the average content distributor who has to comply with the Calm Act? The FCC is attempting to regulate differences in perceived loudness with particular attention being paid to inter-channel audio levels; the difference between programming content and commercial content. In an ideal world, all content providers would provide accurate dialnorm values in the audio bitstream of their programming. These values, sent to the AC-3 encoder before transmission would then be decoded at the consumer set top box or flat screen display, which automatically adjusts "volume" levels at the receiver based on the values received in the metadata. This works by Dynamic Range Control "gain words" calculated in the encoder and then applied at the decoder. These DRC calculations are relative to and based on the indicated loudness of content represented by the dialnorm metadata parameter. In other words, the encoder needs to know how loud the content is intended to be so it can determine when the content is either "too loud" or "too quiet". Dialnorm effectively sets this target. Therefore, it's very important that the dialnorm accurately indicates the loudness of the content.

There are three different methods of complying with the requirement for accurate dialnorm. Each can be used at the discretion of the operator:

Fixed Dialnorm Metadata: The AC-3 encoder Dialog Level is "fixed" to a single value and the content Dialog Levels are conformed to that setting.

Preset Dialnorm Metadata: AC-3 encoder "presets" are programmed, each with different dialnorm values and engaged via a "GPI" or other control interface from automation.

Agile Dialnorm Metadata: The AC-3 encoder is configured to receive external metadata. An upstream "agile" dialnorm metadata system may be used to deliver dynamically changing dialnorm values to the encoder, corresponding to the changing loudness at the content boundaries.

The concept of Fixed metadata is simply to "fix" the AC-3 encoder dialnorm setting to a single value and to bring the loudness of the encoder audio input signal into conformance with this setting. This is the simplest method with no requirement for additional metadata equipment or data management. It is the only approach possible when using an encoder without metadata input or external GPI control.

The Preset metadata concept uses GPI triggers to set predetermined preset values to be loaded into the AC-3 encoder to accommodate known differences in content loudness. For example, known differences between network feed and local playout. Some AC-3 encoders however, reset and disrupt the audio bit stream output when a preset is changed. Depending on the encoder, this may result in an audible "glitch" on air. To avoid this potential problem, it may be necessary to provide a framesync for the output of the AC-3 encoder to stabilize the AC-3 source.

The Agile metadata system allows setting different dialnorm values for different content that has different loudness. This is accomplished by embedding the dialnorm parameter within the metadata bit stream accompanying the content at an upstream location. The metadata is dis-embedded just prior to the AC-3 encoder and then connected to the encoder's external serial metadata input. The encoder dialnorm setting then changes appropriately on the boundaries of the content. The downside of the Agile metadata system is the potential for a severe discrepancy in loudness between programs and between stations if metadata is lost. Encoders with external metadata input provide a "reversion" feature to mitigate the impact of metadata loss. It can be configured to either

retain the most recent metadata value, or revert to an operator-defined preset. While this feature can minimize the impact on the consumer, the error in loudness can still be significant. This method requires each piece of content submitted for broadcast to have its unique dialnorm value embedded into the audio bit stream.

In the real world, content provided to a broadcaster doesn't always contain a valid dialnorm value. Much of the commercial content received at the local level contains none whatsoever. In such cases, the target loudness value should be -24 LKFS (+/- 2dB). This equates to a dialnorm value of 24. Broadcasters should be using a BS.1770 metering system to determine proper LKFS values, and all content received needs to have the dialnorm embedded prior to AC-3 encoding.

Until the "ideal world" becomes a reality, it may be necessary to have a device that maintains audio levels at a particular dialnorm value, particularly when using a fixed dialnorm metadata AC-3 encoder. Enter the Ensemble Designs 9670 LevelTrack software key for the Avenue 9600, 9550, 7660 and 7555. This software uses the BS.1770 loudness algorithm to set a specific LKFS value to be maintained by the audio stream. As mentioned earlier, this translates directly to a dialnorm value.

Station output audio that is run through one of the Ensemble Designs products with the 9670 software key enabled, can be preset to a specific LKFS value before hitting the fixed dialnorm metadata EC-3 encoder. As an example, if the encoder has been set for a fixed dialnorm of 24, the 9670 software would be set to -24 LKFS – in effect feeding the encoder an audio loudness dialnorm equivalent of 24. This allows the broadcaster to use a lower cost encoder and still maintain consistent loudness levels required by the Calm Act.

The FCC mentions that enforcement of the Calm Act will be complaint driven. If stations show a consistent pattern of complaints related to audio level disparity, the FCC will investigate. This is where the Ensemble Designs 9690 Audio Compliance and Monitoring software along with any of the products mentioned earlier will provide a record of LKFS levels of up to four devices. This record can be used to prove compliance.

By installing and properly setting up a 7555, 7660, 9550 or 9600 with 9670 software key for LKFS AGC, and 9690 software key for compliance recording, broadcasters can rest assured that they are in compliance with the Calm Act, and limit audio loudness complaints by their viewers. That's a win-win.

Calm Act Update – NAB 2012

The word around NAB this year is that the ATSC A/85 committee will reconvene shortly to make modifications to the A/85 Recommended Practice for Maintaining Audio Loudness upon which the Calm Act is based.

Specifically, the committee is meeting to discuss standards for the use of "gating" in LKFS averaging. This process removes audio below a certain threshold, from the LKFS averaging equation. They will be meeting to determine exactly where and how the threshold will be implemented.

The problem with the existing spec is that silence is included in the LKFS averaging to determine dialnorm. This means that if a broadcaster were to air a 30 second spot that is silent – (has no audio), they are in effect, non-compliant with the CALM Act (+/- 2 dB from average). In addition, the silence will effect the averaging after audio resumes causing a louder than normal perceived loudness until proper averaging returns things to normal.

The committee will address these issues. It has an implication to broadcasters and manufacturers alike. LKFS metering will have to take into account the threshold of the gating before determining overall LKFS average. File based dialnorm solutions will have to rework their algorithms. Broadcasters of course, will still have to be in compliance with the Calm Act. All of this has to be done by the December of this year. Ensemble Designs will be keeping a close eye on these proceedings and will be looking to make enhancements to our AGC and Compliance software to reflect changes to the spec. Stay tuned.