

SomeJoe

Registered User

Join Date: Jan 2003

Posts: 315

How To Properly Encode Dolby Digital Audio (AC3)

How to Properly Encode Dolby Digital Audio (AC3)

Introduction

Many people on the forum have experienced problems when encoding audio using Dolby Digital. These problems are primarily volume-related, with some dynamic range compression issues as well. This guide aims to educate about how Dolby Digital audio should be encoded, and how to make it sound best.

References

The primary references for the information contained in this guide are two guides on Dolby's web site. The first is **Standards and Practices for Authoring Dolby Digital and Dolby E Bitstreams**, which has the best information on Dynamic Range Compression. The other is **Dolby Digital Professional Encoding Guidelines** which gives an excellent explanation of the dialogue Normalization parameter. You will need Adobe Acrobat Reader to view these .pdf documents.

Philosophy of Dolby Digital

Dolby Labs has been doing high-quality audio with cutting-edge techniques for a long time, using their past experience as a guide. As such, there is often confusion about their methods and philosophy to those of us who are not privy to that information. Of prime example is the current problem: *Why is Dolby Digital so much quieter compared to my original sound?*

Most audio destined for DVDs is audio originally recorded for use in the movie theater. The movie industry has a huge advantage when producing audio for the theater -- the theater has large speakers and amplifiers, and a quiet, near-ideal listening environment. Huge dynamic

ranges are possible, where the slightest whisper of dialogue is audible, yet gunshots and explosions can be earth-shattering. Dolby's dilemma was: "How do we bring this audio, with its huge dynamic range, into the home?" This is a major problem -- most homes don't have the speakers and amplifiers necessary to shake the living room. Further, background noise in the home can easily drown out those subtleties in the soundtrack.

Dolby's answer is to allow the decoder to modify the sound to compensate for these problems. Low-volume sounds are boosted automatically so they can be heard, whereas high-volume sounds are quieted down so that speakers aren't blown and other persons in the home are not disturbed. Further, Dolby Digital allows for different program material to be equalized, so that volume does not have to be adjusted when switching between inherently quiet programs and inherently loud programs.

Decoder Specifics

The methods I'm about to present here for encoding Dolby Digital are generic and do not apply specifically to any one encoder. All Dolby-certified encoders (and some non-certified ones) will have the appropriate parameters available to follow this procedure. I have personally tested the Sonic Foundry 5.1 Plug-In Pack for ACID Pro, as well as Sonic Foundry Soft Encode. These methods should also work for BeSweet, Vegas Video + DVD, Scenarist, and other software-based encoders.

Basic Parameters

Every Dolby Digital encoder has some basic parameters that need to be set.

The first is the channel combination, presented as (Number of front channels)/(Number of rear channels), with an optional ".1" added to represent a low-frequency-effects (LFE) channel if present. i.e. 2/0 represents normal left and right stereo sound. 3/2.1 represents a standard "5.1" setup, of Front Left, Front Right, Front Center, Rear Left, Rear Right, and LFE. This parameter should obviously be set to the number of channels of program material you will be encoding.

The other major basic parameter is the bitrate. Obviously, higher bitrates allow for less compression. Typical bitrates used are 192 kbps for 2/0 program material, and 448 kbps for 5.1 program material.

Referencing Volume to a Known Level - Dialogue Normalization

To meet the Dolby Digital requirement that different programs should have approximately the same listening level (thus the consumer does not have to adjust volume level between programs), Dolby Digital incorporates a parameter called dialogue Normalization. This metadata parameter tells the decoder how far away from the reference level the average sound pressure level of the material's dialogue is.

The movie industry masters their soundtracks in a specific way. The maximum rated sound level (where all amplifiers are putting out their rated power) is 0 dB. Sounds below that level are rated in terms of how many decibels (dB) they are down from that maximum level. As such, these values are negative. The movie industry typically masters the "normal" listening level of dialogue (where people are speaking in a normal voice) at -31 dBFS. In other words, a speaking voice is at an average of -31 dB when referenced to the 0 dB maximum sound level, hence the term decibels of full scale (dBFS).

Since movie content is the largest class of programs to go on DVD, Dolby chose -31 dBFS as the reference level for audio on DVD, where 0 dB represents the maximum encodable digital sound level (full scale).

The dialogue normalization parameter needs to be set to the LAeq level of your program material's dialogue. LAeq stands for the long-term A-weighted sound pressure level. Loosely, this is the average volume level of your source material's dialogue. Us lowly consumers really don't have a tool that can measure this parameter, but we can get close. Sonic Foundry's Sound Forge has a "Normalization" feature that can measure the RMS level of a .wav file (or the portion thereof containing dialogue). CoolEdit may also have a feature like this. To use it in Sound Forge, open your .wav file containing the movie audio. Select a section containing dialogue (no sound effects or music). Go to "Process"/"Normalize". Select the "Average RMS Power (Loudness)" radio button. Then click the "Scan Levels" button. The displayed "RMS" level is very close (within 1-2 dB) to the LAeq level.

That RMS level is the number that the dialogue normalization parameter should be set to. In other words, if the RMS level in Sound Forge shows as -17.6 dB, set the dialogue normalization parameter in your Dolby Digital encoder to -18 dBFS.

The decoder will perform an attenuation of $(31 + \text{dialnorm})$ dB to the program material when played back. So, in this case, the decoder will attenuate by $(31 + -18) = 13$ dB. This will bring the average sound level of the material to $(-17.6 - 13) = -30.6$ dBFS. The program is now played back at approximately -31 dBFS, the reference level.

-31 dBFS is a lower average volume level than what is typical from other sources. It will be noticeable that you will have to turn the volume up on your system when playing a DVD versus playing broadcast, tape, or other non-Dolby Digital program material.

Allowing Comfortable Listening - Dynamic Range Compression

Meeting the other end of the requirement, that the consumer should be able to listen to quiet and loud sections of the program without having to adjust volume levels, requires a decrease in the dynamic range of the program. A movie, with whispers at -50 dBFS and explosions at -5 dBFS can't be comfortably listened to in the average home. The whisper is drowned out by extraneous background noise, and if the explosion is played at a tolerable level that doesn't wake up the neighbors, regular dialogue at -31 dBFS requires straining to adequately hear.

Dolby solves this problem by compressing the dynamic range of the program material. Quiet sounds are automatically boosted in volume so that they're audible, and loud sounds are automatically cut down in volume to tolerable levels.

There are several dynamic range compression profiles available that are custom tailored to the particular flavor of program material. However, all of them share the same basic features. All of the compression profiles can be thought of as an input-output "black box", where certain input volume levels are mapped to certain output volume levels. Observe this graph, which is a graph of one of the Dolby Digital compression profiles (Film Light).

The blue line is the "unity gain" line, also referred to as the "no compression" line. This line represents that the dynamic range compressor feature of the decoder is essentially turned off, and no boost or cut of the program material is done.

The purple line is the compression profile for "Film Light". It is divided into 5 different sections, as are the other Dolby Digital compression profiles:

Unity Gain = Volume neither boosted nor cut

Variable Boost = Beginning of increasing volume of soft sounds

Constant Boost = Increase volume by a fixed amount for very soft sounds

Early Cut = Beginning of attenuating volume for loud sounds

Cut = Very loud sounds almost clamped to a maximum volume level

The application of a compression profile like this allows the soft sounds to be heard while preventing speaker overdrive and disturbances by the loud sounds.

The Dolby Digital encoder offers 5 different compression profiles that can be specified depending on the nature of the program material being encoded. This graph illustrates all of the available profiles. The profiles range from no compression ("None"), to fairly light compression ("Music Light") all the way to extremely aggressive compression ("Speech").

For the exact dB numbers where each range of dynamic range compression is located on the graph, see the Dolby documents cited in the references at the beginning of this guide.

Many authors, when compressing audio to Dolby Digital, are turned off by the idea of dynamic range compression. You have this well-mixed audio with nice dynamic range and are then going to kill it by compressing that dynamic range. This is a valid concern, but should be answered by looking at what the listening environment is going to be. If you are authoring a DVD only for yourself, and you have a home theater room that can deliver theater-like sound, perhaps a compression profile of "None" is suitable for you. However, this profile may not sound good in a more mundane living room. Some experimentation may be in order to determine what compression profile will sound best for you. Most Hollywood DVDs use "Film Light" or "Film Standard".

The following point, however, cannot be stressed enough: **In order for the Dynamic Range Compression to work as designed, the Dialogue Normalization parameter MUST be properly set first!**

All of the dynamic range compression profiles assume that the average volume level of the program material's dialogue being fed to the dynamic range compressor is -31 dBFS. If that is not the case, boost or cut will be applied to the material when it shouldn't be!

A prime example is the situation where an average volume .wav file (with an LAeq of -16 dBFS, for example) is fed to Soft Encode using Soft Encode's default dialogue Normalization and Dynamic Range Compression settings. The default dialogue Normalization setting is -27 dB, and the default Dynamic Range Compression is set to "Film Standard". Because of the misadjusted dialnorm parameter, only $(31 + -27) = 4$ dB of attenuation is applied to the audio, so the average volume level is $(-16 - 4) = -20$ dBFS instead of the expected -31 dBFS. This places the audio on the DRC graph at the incorrect position, and now most of the audio is being played back in the Early Cut and Cut ranges. This causes the audio to sound flat and dull, with a possible audible "pumping" of the volume up and down as the decoder changes between Early Cut and Cut based on average volume level. Here is a representative [graph](#). If dialnorm had been properly set to -16 dB, the audio would be centered at -31 dBFS, and would sound like it is supposed to.

Line Mode and RF Mode

The Dolby Digital compressors have the ability to further alter the compression profile to compensate for different transport mediums. Most of the time, audio is transported between devices in "Line Mode", where a line-level is used. There is also "RF Mode", meant for broadcasting of Dolby Digital and devices that send audio via RF cables to a TV set. RF mode sound from the decoder uses a higher average volume level (-20 dBFS vice -31 dBFS) in order to correlate volume level well with other, non-Dolby broadcast audio, and also can use a more aggressive Dynamic Range Compression to prevent overmodulating the signal. There is an option in most Dolby Digital encoders to turn on that overmodulation limiter (in Soft Encode, it is labeled as "RF Overmodulation Protection"). Since we are primarily interested in authoring for DVD which will operate in Line Mode, we do NOT want to insert the additional Dynamic Range Compression that RF Overmodulation Protection will add. Therefore, for DVD authoring, the RF Overmodulation Protection option should be turned off.

Example Compression Settings

For this example, I will use an audio file that was captured from analog material. It is a plain stereo .wav file, 48 kHz, 16-bit. (Note that this audio was from a filmed seminar, and thus the entire audio file was dialogue. Because of this, I did not select a particular range to compute the RMS level, but I rather selected the whole file. In your file, you should select only a portion that is representative dialogue.) I loaded it into Sound Forge and [measured the RMS level of](#)

the entire file as -20 dBFS.

Knowing that, here are some screen shots of the proper settings to encode this file (Sonic Foundry's 5.1 Plug-In Pack for ACID Pro was used as the encoder. Your encoder may not have all options available.)

The first page in ACID is the Audio Service Configuration, where the coding mode (2/0), the data rate (192 kbps), and dialnorm (-20 dBFS) are set.

The second page is the Bitstream Information. Set these parameters are appropriate for your source material.

The third page is the Extended Bitstream Information page. Your encoder may not allow these options to be set. I typically do not set any options here.

The final page is the Preprocessing page. Here is where you set the Dynamic Range profile you want to use for the material. The Line Mode setting is the one that will actually be used by the DVD player or Dolby Digital receiver, since the sound will be coming from the line outs. The RF Mode generally won't be used, but I typically set it to the same profile as the Line Mode. Make sure RF Overmodulation Protection is not checked.

Conclusion

I hope that this guide has given some insight into the proper methods that should be used to get the most out of your Dolby Digital encoding.