

1 Dolby Digital Guidelines

This document contains guidelines from several Dolby Laboratories publications (available at www.dolby.com).

1.1 Dolby Digital Encoding for DVD-Video

When encoding audio content for DVD-Video authoring a few guidelines must be kept in mind:

- Always have at least two seconds of digital black silence at the beginning of the bitstream. This gives the large amount of digital circuitry in playback systems time to lock and start decoding before the real material starts playing. It is not necessary to leave digital black at the end of the file.
- Always encode at 448 kb/s data rate for multichannel material and at 192 kb/s for two-channel stereo material.
- Make sure the source material has a sample rate of 48 kHz, otherwise the sample rate of the material must be converted before the encoding begins. This is a necessary step since the specifications for the DVD-Video format allow only a 48 kHz sample rate.
- Start with the best quality material possible. Dolby Digital encoders can accommodate up to 24-bit audio.
- Select the most appropriate Dialog Normalization value and Dynamic Range Compression profile for the material. This is crucial since each type of material requires a different setting of these parameters.
- Select the appropriate values for each of the parameters. If the original material is Dolby Surround encoded, then enable the Dolby Surround flag. If mixing information exists fill-in the appropriate parameters. Many decoders react to these parameters so the correct values must be set.
- Do not enable the LFE Channel unless there is dedicated Low-Frequency Effects (LFE) material in the original audio source.
- Always monitor the encoded output with a professional reference decoder. Monitoring is the only way to check the integrity and accuracy of the encoded audio.

1.2 Music on DVD-Video

One exploding area in the DVD-Video format is the music disc. Since the introduction of DVD-Video there have been many titles authored exclusively with music content. The purpose of these discs is to provide the consumer with high quality multichannel audio without the need to watch the video or maneuver with software menus and buttons. The main goal is to have these discs behave as audio compact discs from an authoring standpoint. The user can insert the disc in the player without turning on the TV to listen to the content. There are many video options the authoring engineer can choose from. Some discs have been authored to display one video slide throughout an entire song, some display rolling credits and lyrics, and some display a live or studio produced music video.

1.3 Supported Data Rates

When encoding audio for DVD-Video one of the key parameters in the Dolby Digital encoder is the Data Rate. The DVD-Video specification requires every DVD-Video player to be able to decode Dolby Digital bitstreams up to and including 448 kb/s. Through its rigorous licensing program, Dolby Laboratories guarantees that every DVD-Video player on the market is capable of decoding a Dolby Digital bitstream at 448 kb/s.

Dolby Laboratories recommends that all multichannel (more than two-channel) material is encoded at 448 kb/s and that all two-channel stereo content is encoded at 192 kb/s.

1.4 Trademark Usage

Dolby Laboratories encourages use of the Dolby Digital trademark to identify soundtracks and other audio programs that are Dolby Digital encoded. This is an effective way to inform listeners of the audio format, and the use of a standard logo promotes easy recognition in the marketplace. As with any trademark, the Dolby Digital logo may not be used without permission. Dolby Laboratories provides a royalty-free Trademark and Standardization Agreement (TSA) for companies who wish to use Dolby trademarks. The company that owns the program material being produced must sign this agreement. Recording studios or production facilities that provide audio production, encoding, or manufacturing services for outside clients generally do not require a trademark license.

[For further information, please contact Dolby Laboratories at www.dolby.com.]

1.5 Dialog Normalization Value (dialnorm)

The Dialog Normalization value is a Dolby Digital parameter that describes the long-term average dialog level of the associated program. It may also describe the long-term average level of programs that do not contain dialog, such as music. The Dialog Normalization value indicates how far the average dialog level of the encoded program is below digital 100% full scale (0 dBFS). This level is specified on an absolute scale ranging from -1 dBFS to -31 dBFS. Dolby Digital decoders attenuate programs based on the Dialog Normalization value in order to achieve uniform playback level. This parameter determines the audio reproduction level and affects other parameters and decoder operation.

[For further information, please contact Dolby Laboratories at www.dolby.com.]

1.6 Dynamic Range Control (DRC)

An important feature of Dolby Digital is that it conveys audio unaltered in dynamics. Unlike almost any previous broadcast medium, it therefore gives the listener the option to hear the program as the mixer intended, even if that means that it goes from scarcely audible to extremely loud.

Present analog broadcast processors force the program level towards full modulation of the transmitter for a substantial portion of the time, eliminating most of the dynamic range; an incidental benefit being an approximate normalization of listening level. In other words, the same device that reduces the dynamic range determines the average volume. With Dolby Digital there are no technical pressures to reduce dynamic range, and mean or average volume is addressed by

dialog normalization. Thus in Dolby Digital, the need for dynamic range compression can be considered independently of average listening levels.

After discrepancies in absolute or average volume have been reduced by the application of dialog normalization, many program items require no further processing for non-ideal listening conditions. Some program items, however, have too great a dynamic range for some listeners. An obvious example is the movie soundtrack; if the volume is set for satisfactory intelligibility of dialog, sounds such as explosions may be unacceptably loud when reproduced in the home. Another example is symphonic music; if the volume is set for comfort in loud passages, very quiet ones may be lost in the background noise. In contrast, news bulletins or rock music have little inherent dynamic range, and provided their absolute levels have been set appropriately there is no reason to apply dynamic range compression.

Dynamic Range Control incorporates both selectable dynamic range compression and automatic overload protection limiting. Dolby Digital encoders generate control words, `dynrng` and `compr`, which can be used in the decoder to compress and limit the dynamic range of a program. The Dynamic Range Compression profile algorithm is based on a simple audio loudness measurement. In contrast, overload protection limiting is based on the peak levels.

1.7 Dynamic Range Compression Profile

The Dynamic Range Compression profile (also referred to as a preset) determines the characteristic curve of the dynamic range compression algorithm. Dynamic Range Control generates `dynrng` and `compr` gain words during the encoding process. A Dolby Digital decoder uses these gain words to reduce the dynamic range of the audio program during playback. This feature can be disabled on the decoder (except when downmixing) by the user who desires program reproduction with the original dynamic range. Dynamic Range Compression can be set by selecting one of several built-in profiles. Six preset DRC profiles are available to content producers: Film Light, Film Standard, Music Light, Music Standard, Speech, and None. Each is applied in the pattern shown in Figure 1-1. In each case, the center of the null band is assigned to the dialogue level parameter setting, and the DRC profile is applied in relation to that level. Here are the details of the range for each profile.

Film Light

- Max Boost: 6 dB (below -53 dB)
- Boost Range: -53 to -41 dB (2:1 ratio)
- Null Band Width: 20 dB (-41 to -21 dB)
- Early Cut Range: -26 to -11 dB (2:1 ratio)
- Cut Range: -11 to +4 dB (20:1 ratio)

Film Standard

- Max Boost: 6 dB (below -43 dB)
- Boost Range: -43 to -31 dB (2:1 ratio)
- Null Band Width: 5 dB (-31 to -26 dB)
- Early Cut Range: -26 to -16 dB (2:1 ratio)

- Cut Range: -16 to +4 dB (20:1 ratio)

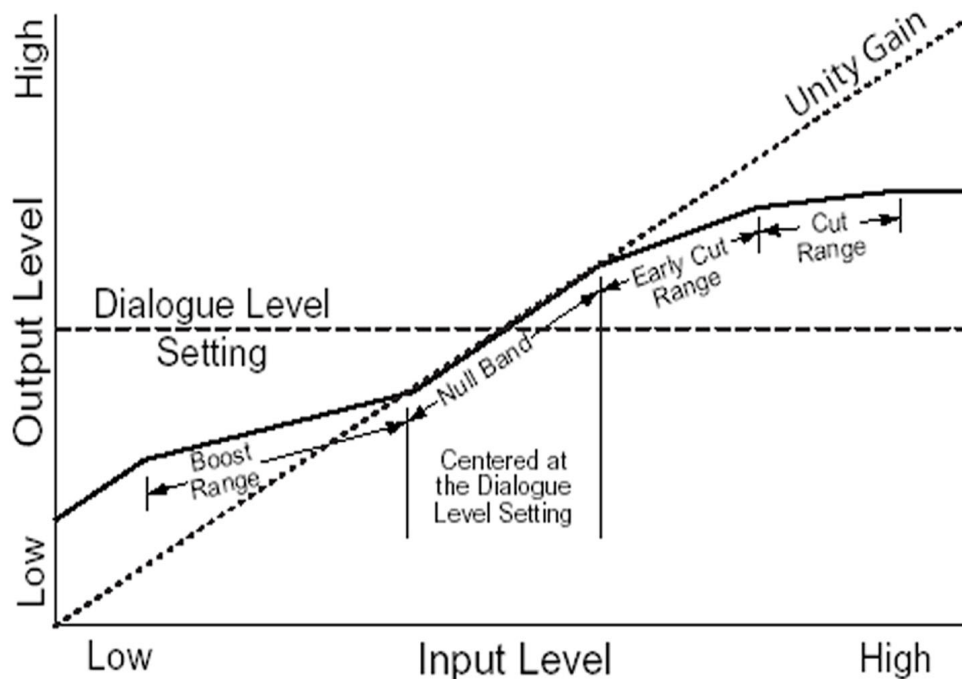


Figure 1-1

Music Light (No early cut range)

- Max Boost: 12 dB (below -65 dB)
- Boost Range: -65 to -41 dB (2:1 ratio)
- Null Band Width: 20 dB (-41 to -21 dB)
- Cut Range: -21 to +9 dB (2:1 ratio)
- Dolby Laboratories, Inc. Metadata Guide

Music Standard

- Max Boost: 12 dB (below -55 dB)
- Boost Range: -55 to -31 dB (2:1 ratio)
- Null Band Width: 5 dB (-31 to -26 dB)
- Early Cut Range: -26 to -16 dB (2:1 ratio)
- Cut Range: -16 to +4 dB (20:1 ratio)

Speech

- Max Boost: 15 dB (below -50 dB)
- Boost Range: -50 to -31 dB (5:1 ratio)
- Null Band Width: 5 dB (-31 to -26 dB)
- Early Cut Range: -26 to -16 dB (2:1 ratio)

- Cut Range: -16 to +4 dB (20:1 ratio)

None

- No DRC profile selected. The dialogue level parameter (dialnorm) is still applied.

The content producer chooses which of these profiles to assign to each mode; when the consumer or decoder selects a DRC mode, the profile chosen by the producer is applied.

In addition to the DRC profile, metadata can limit signal peaks to prevent clipping during downmixing. This metadata, known as overload protection, is inserted by the encoder only if necessary. For example, consider a 5.1-channel program with signals at digital full scale on all channels being played through a stereo, downmixed linelevel output. Without some form of attenuation or limiting, the output signal would obviously clip. Correct setting of the dialogue level and DRC profiles normally prevents clipping and unnecessary application of automatic overload protection.

Note: DRC profile settings are dependent on an accurate dialogue level setting. Improper setting of the dialogue level parameter may result in excessive and audible application of overload-protection limiting.

1.8 LFE Channel

The status of the LFE Channel parameter indicates to a Dolby Digital encoder whether an LFE Channel is present within the bitstream. Channel mode determines whether the LFE Channel parameter can be set. You must have at least three channels to be able to add an LFE channel.

1.9 Bitstream Mode

- **Complete Main (CM):** CM flags the bitstream as the main audio service for the program and indicates that all elements are present to form a complete audio program. Currently, this is the most common setting. The CM service may contain from one (mono) to six (5.1) channels.
- **Main M&E (ME):** The bitstream is the main audio service for the program, minus a dialogue channel. The dialogue channel, if any, is intended to be carried by an associated dialogue service. Different dialogue services can be associated with a single ME service to support multiple languages.
- **Assc. Visual Imp. (VI):** This is typically a single-channel program intended to provide a narrative description of the picture content to be decoded along with the main audio service. The VI service may also be a complete mix of all program channels, comprising up to six channels.
- **Assc. Hear Imp. (HI):** This is typically a single-channel program intended to convey audio that has been processed for increased intelligibility and decoded along with the main audio service. The HI service may also be a complete mix of all program channels, comprising up to six channels.
- **Assc. Dialogue (D):** This is typically a single-channel program intended to provide a dialogue channel for an ME service. If the ME service contains more than two channels, the D service is limited to only one channel; if the ME service is two channels, the D service can be a stereo pair. The appropriate channels of each service are mixed together (requires special decoders).

- **Assc. Commentary (C):** This is typically a single-channel program intended to convey additional commentary that can be optionally decoded along with the main audio service. This service differs from a dialogue service because it contains an optional, rather than a required, dialogue channel. The C service may also be a complete mix of all program channels, comprising up to six channels.
- **Assc. Emergency (E):** This is a single-channel service that is given priority in reproduction. When the E service appears in the bitstream, it is given priority in the decoder and the main service is muted.
- **Assc. Voice Over (VO):** This is a single-channel service intended to be decoded and mixed to the Center channel (requires special decoders).
- **Main Sv Karaoke (K):** The bitstream is a special service for karaoke playback. In this case, the Left and Right channels contain music, the Center channel has a guide melody, and the Left and Right Surround channels carry optional backing.

1.10 Copyright Bit

This parameter indicates whether the encoded Dolby Digital bitstream is copyright protected. It has no effect on Dolby Digital decoders and its purpose is purely to provide information.

1.11 Original Bitstream

This parameter indicates whether the encoded Dolby Digital bitstream is the master version or a copy. It has no effect on Dolby Digital decoders and its purpose is purely to provide information.

1.12 Center Downmix Level

When the encoded audio has three front channels (L, C, R), but the consumer has only two front speakers (left and right), this parameter indicates the nominal downmix level for the Center channel with respect to the Left and Right channels. Dolby Digital decoders use this parameter during downmixing in Lo/Ro mode when Extended Bitstream Information parameters are not active.

1.13 Surround Downmix Level

When the encoded audio has one or more Surround channels, but the consumer does not have surround speakers, this parameter indicates the nominal downmix level for the Surround channel(s) with respect to the Left and Right front channels. Dolby Digital decoders use this parameter during downmixing in Lo/Ro mode when Extended Bitstream Information parameters are not active.