Dolby Digital Professional Encoding Guidelines

Issue 1

S00/12972

Dolby Laboratories Inc

Corporate Headquarters

Dolby Laboratories Inc
100 Potrero Avenue
San Francisco, CA 94103-4813
Telephone 415-558-0200
Facsimile 415-863-1373
www.dolby.com

European Licensing Liaison Office

Dolby Laboratories
Wootton Bassett
Wiltshire, SN4 8QJ, England
Telephone (44) 1793-842100
Facsimile (44) 1793-842101

Far East

Dolby Laboratories
International Services, Inc.
Japan Branch
Fuji Chuo Building 6F
2-1-7, Shintomi, Chuo-ku
Tokyo 104-0041 Japan
Telephone (81) 3-5542-6160
Facsimile (81) 3-5542-6158

Dolby Laboratories
Representative Office
7/Fl., Hai Xing Plaza, Unit H
1 Rui Jin Road (S)
Shanghai 200023 China
Telephone (86) 21-6418-1015
Facsimile (86) 21-6418-1013

S00/12972 Issue 1

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Purpose and Scope

This manual is intended to serve as a guide for performing Dolby Digital (formerly AC-3) professional audio encoding in applications other than film production. It contains information on the features of the Dolby Digital system, professional encoders and decoders, consumer decoders, distribution formats, and Dolby trademark use. The key features of the Dolby Digital system, including Dialog Normalization (also referred to as volume normalization), Dynamic Range Control (DRC), and downmixing (multiple channels through fewer outputs) are described and guidelines for their use are presented.

For detailed descriptions of the Dolby Digital algorithm, refer to the Advanced Television Systems Committee (ATSC) documents A/52*, Digital Audio Compression Standard (AC-3),* and A/54, *Guide to the Use of the ATSC Digital Television Standard*, which can be found on the ATSC web site at www.atsc.org. For a list of additional technical papers on Dolby Digital, refer to Appendix G, Bibliography.

1.2 General Information

Dolby Digital is a perceptual audio coding system that is based on the fundamental principles of human hearing. It was first developed in 1992 as a means to allow 35 mm theatrical film prints to carry multichannel digital audio directly on the film without sacrificing the standard analog optical soundtrack. Since its introduction the system has been adopted for use with laser disc, ATSC high definition and standard definition digital television, digital cable television, digital satellite broadcast, DVD-Video, DVD-ROM, DVD-Audio, and Internet audio distribution. It is intended for use as an emissions coder that encodes audio for distribution to the consumer, not as a multigenerational coder that is used to encode and decode audio multiple times. For applications that require multigenerational coding, refer to Section 7.2, Contacting Dolby Laboratories, to acquire information on Dolby E technology.

Dolby Digital divides the audio spectrum into narrow frequency bands using mathematical models derived from the characteristics of the ear and analyzes each band to determine the audibility of those signals. To maximize data efficiency, a greater number of bits represent more audible signals; fewer bits represent less audible signals. In determining the audibility of signals, one phenomenon that the system makes use of is known as *masking*. Masking refers to the fact that the ear is less sensitive to low-level sounds when there are higher-level sounds at nearby frequencies. When this occurs, the high-level sound masks the low-level one, rendering it either less audible or inaudible. By taking advantage of this phenomenon, audio can be encoded much more efficiently than in other digital coding systems with comparable audio quality, such as linear PCM. This makes Dolby Digital an excellent choice for systems where high audio quality is desired, but bandwidth or storage space is restricted. This is especially true for multichannel audio, where the compact Dolby Digital bitstream allows full 5.1-channel audio to occupy less space than a single channel of linear PCM audio.

The Dolby Digital system is designed to allow the encoder to continue evolving and improving. As more research is conducted, the encoding algorithm can be modified for improved accuracy. The Dolby Digital system is also designed to pass encoding improvements along to the decoder providing improved audio quality for all listeners.

1.2.1 More than a Coding Technology

Dolby Digital is more than just an audio coding technology. It is also a sophisticated audio delivery and reproduction system that allows both the program producer and the end listener to affect how the audio program will ultimately be heard. For the first time in the consumer audio industry, a program producer can deliver multichannel audio along with control parameters. These parameters can determine the relative playback level, using Dialog Normalization, the preferred dynamic range compression setting, and how the audio program will sound to consumers listening to a stereo downmix of it. These capabilities not only provide the program producer useful new tools that can enhance the listening experience, but they also create the possibility of undesirable results if encoding parameters are not set correctly.

When Dolby Digital encoding it is important to understand the options typical Dolby Digital decoders present to the consumer and how those options can affect the audio. Dolby Digital consumer decoders are used in a variety of listening situations with high-end products offering the listener a wealth of options not found in more basic implementations. Consumers can listen to multichannel audio in a number of ways including:

- Full dynamic range 5.1-channel state
- Reduced dynamic range 5.1-channel state (for apartment dwellers or late night listeners, etc.)
- Two-channel Dolby Surround compatible downmix (which can then be Dolby Surround Pro Logic decoded)
- Normal two-channel stereo downmix
- Mono downmix

It is important to understand the various options presented by both Dolby Digital professional encoders and consumer decoders, to be aware of how they interact with and affect the audio, and to ensure that encoder options are used properly for best results.

1.2.2 Features of Dolby Digital

- Delivers from one (mono) to 5.1 discrete channels of audio in a variety of configurations

- Efficiently reduces audio data with typical savings of 12:1 (5.1-channel Dolby Digital at 384 kb/s compared to a 16-bit, 48 kHz linear PCM source)
- Provides the capability for all main channels to deliver a frequency bandwidth from 3 Hz (DC Highpass Filter enabled) to 20.7 kHz
- Provides an optional Low-Frequency Effects (LFE) channel for additional bass with a frequency bandwidth from 3 Hz (DC Highpass Filter enabled) to 120 Hz
- Supports a wide range of data rates from 32 kb/s to 640 kb/s
- Accepts input sampling rates of 32, 44.1, and 48 kHz
- Fully accepts input word lengths of 16, 18, 20, or 24 bits
- Provides for uniform, level-matched playback for all sources
- Allows producer and/or user to control dynamic range during playback
- Provides for multichannel programs to be downmixed all the way to mono
- Allows the creation of a Dolby Surround compatible downmix for listeners with Dolby Surround Pro Logic systems
- Permits the producer to optimize Center and Surround channel levels for use in stereo downmix mode
- Provides a Karaoke Mode
- Permits the transmission and disk storage of SMPTE time code stamps along with the Dolby Digital data

- Enables the producer to indicate:
  - The reference Sound Pressure Level (SPL) used when mixing the audio program; allows for calibrated playback levels
  - The type of room (large or small) used to mix the audio program
  - That the two-channel audio program is Dolby Surround (Lt/Rt) encoded
- Provides an indicator for original bitstreams, to distinguish from a bitstream copy
- Provides a copyright control bit
- Provides a Bitstream Mode indicator (Main and Associated Services for broadcast)
- Produces improved sound quality for all listeners through refinements in the encoding algorithm

Chapter 2
Production Environment

2.1 System Configuration

There are many ways to configure a production environment for Dolby Digital professional encoding. The type of Dolby Digital encoder and the associated hardware determines how to connect the components.

As shown in Figure 2-1, the essential components for Dolby Digital professional encoding are an audio source (mono to 5.1-channel), a Dolby Digital encoder, a capture and storage device for the encoded output, and a professional reference decoder. Not shown, although just as important, is a properly calibrated audio reproduction system for monitoring the decoded output. Considering the wide variety and availability of these components, general guidelines are given on the setup of these systems. It is impossible to cover every configuration in this manual, therefore the reader is encouraged to contact the appropriate equipment manufacturer for details on each product.

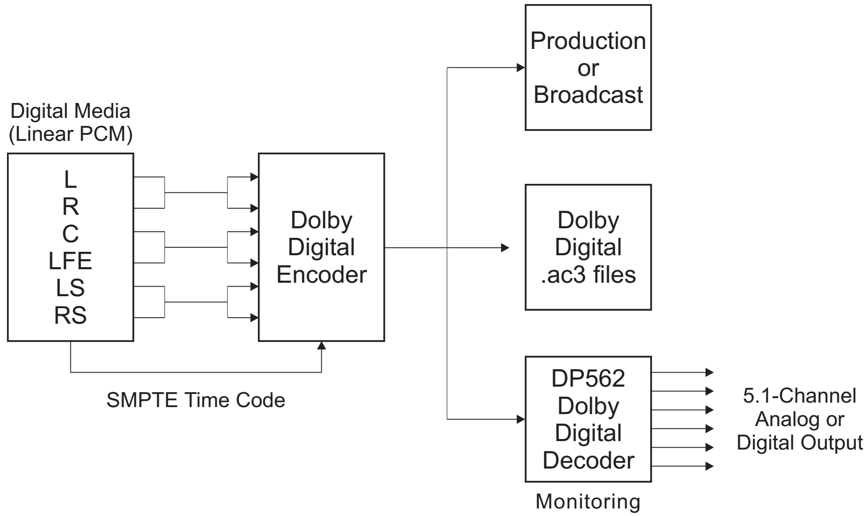


Figure 2-1 Generic Dolby Digital Encoding System

Following is a brief description of the components in Figure 2-1.

- The linear PCM source can be any digital audio storage device (e.g., Modular Digital Multitrack (MDM), DAT, CD, etc.). Many types of media and products are capable of storing digital audio content or files. Choosing the right medium and product depends on whether the project requires real-time or non-real time encoding. In addition, the number of channels to be encoded affects the choice of medium.

- The Dolby Digital encoder can be any professional encoder manufactured or licensed and approved by Dolby Laboratories and bearing the Dolby Digital logo. These encoders are available either from Dolby Laboratories or Dolby professional encoder licensees.
- The capture device for the Dolby Digital bitstream can be any computer with a digital audio interface or equivalent.
- It is important that the monitoring environment for Dolby Digital professional encoding includes a Dolby Laboratories Model DP562 professional reference decoder. For more information, refer to Section 2.2, Monitoring Through a Decoder, Section 4.11, Using the Dolby Model DP562 Professional Reference Decoder, and Section 6.2, Dolby Digital Professional Decoders.

Figure 2-2 depicts a configuration in which a real-time Dolby Digital encoder receives multichannel linear PCM audio and SMPTE time code from an MDM. The encoder delivers a Dolby Digital bitstream and SMPTE time code data to the digital audio input of a computer that is equipped with the appropriate hardware and software. The Dolby Digital Recorder program for Windows 95 and Windows NT, available from Dolby Laboratories, can capture and store the encoder output to a hard drive. The recorded .ac3 file is compatible with all DVD authoring systems. A DP562 professional reference decoder is used for real-time monitoring of the Dolby Digital bitstream.

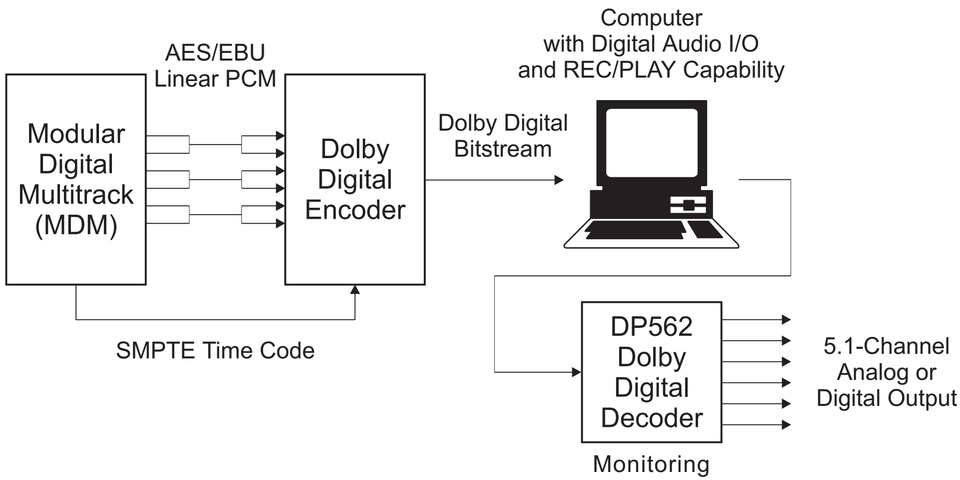


Figure 2-2 Dolby Digital Recorded to a Computer

The configuration in Figure 2-3 is similar to the one in Figure 2-2 except that the computer is not equipped with a digital audio input. In this case a DAT-Link+ (AES3/IEC 958 to SCSI) adapter converts the Dolby Digital bitstream to a SCSI interface and delivers it to the SCSI bus of the computer. The rest of the configuration is identical to that in Figure 2-2.

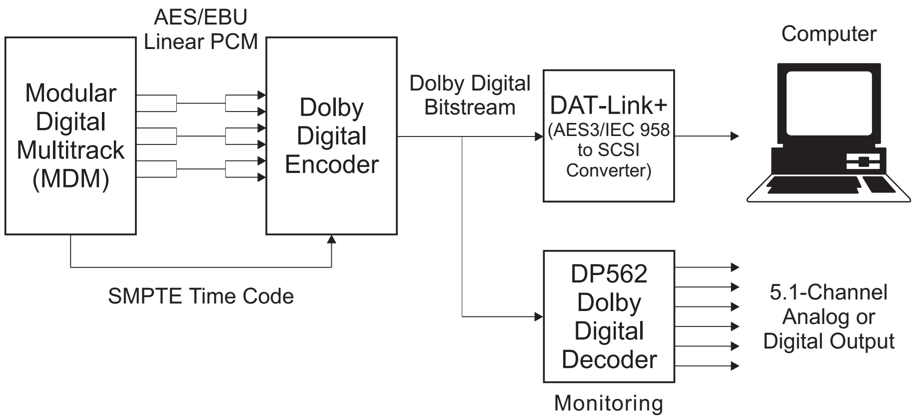


Figure 2-3 Dolby Digital Recorded to a Computer with DAT-Link+

The configuration in Figure 2-4 depicts a real-time PCI card Dolby Digital encoder installed in a computer. The PCI card (two- or 5.1-channel) can accept SMPTE time code as well as AES/EBU or S/PDIF linear PCM audio input. In addition to being stored to the hard drive of the computer for DVD authoring, the encoded output from the card can be monitored by the decoder in real-time. The stored .ac3 file is compatible with all DVD authoring systems. Dolby Digital professional encoder cards are available from Dolby professional encoder licensees.

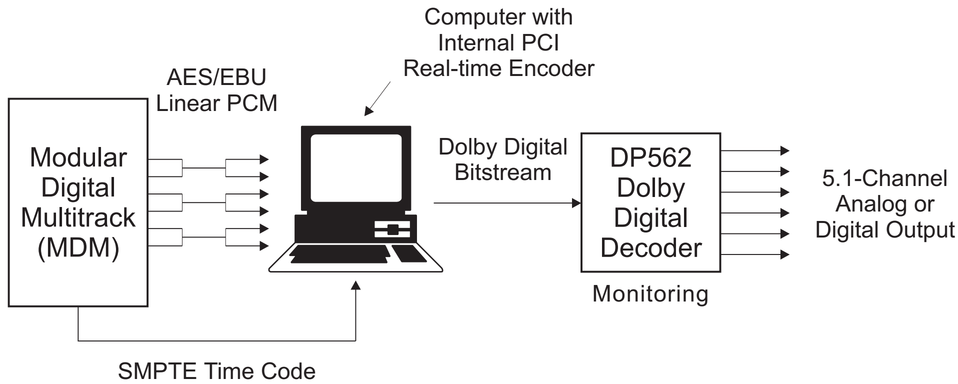


Figure 2-4 Dolby Digital Encoding Using a Licensed PCI Card

Figure 2-5 shows Dolby Digital professional encoder software running on a computer. Like the PCI card, this solution is available from Dolby Laboratories professional encoder licensees. Usually, the audio source material exists on digital media as mono or stereo files. The computer encodes the linear PCM files and outputs .ac3 files compatible with all DVD authoring systems. The software may be able to control digital devices such as an MDM to capture the linear PCM files to a hard drive for later encoding. Monitoring of the .ac3 files is done after the encoding session is complete. The .ac3 files played from the digital audio card in the computer are monitored using a DP562 decoder.

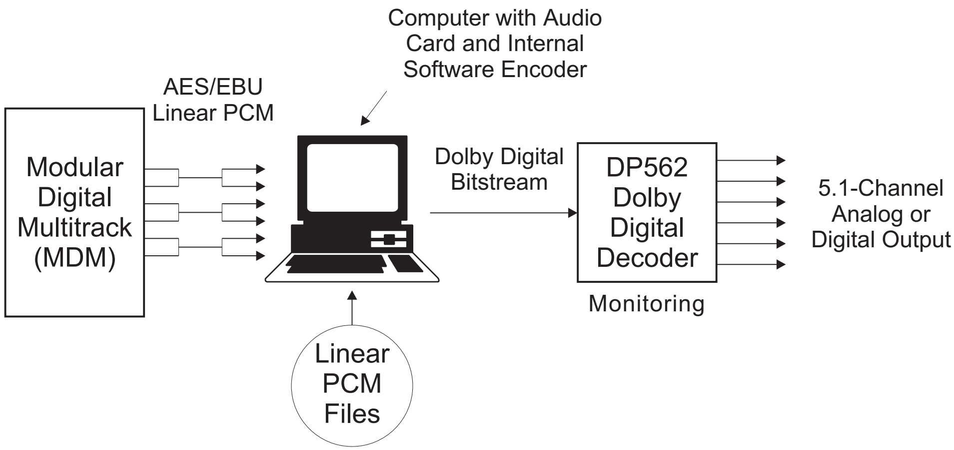


Figure 2-5 Dolby Digital Encoding Using Licensed Computer Software

2.2 Monitoring Through a Decoder

Dolby Digital decoders are divided into two fundamental categories: consumer and professional. It is important to monitor the Dolby Digital encoding process using a professional reference decoder such as the Dolby Model DP562, which also includes Dolby Surround Pro Logic decoding. Unlike consumer decoders, a professional reference decoder affords the user greater flexibility as well as the critical diagnostic capabilities essential in a production environment. One such feature is the ability to obtain real-time information on the effects of various parameter settings such as

Dialog Normalization and Dynamic Range Control (DRC). Another is the capability to emulate any type of consumer decoder on the market whether it is a DVD player, an A/V receiver, an HDTV, or a set-top box (STB). Since most consumer decoders have some Dolby Digital features (Dialog Normalization, Dynamic Range Control, downmixing, etc.) preset at the factory, it is critical that the encoding engineer use a professional reference decoder to allow monitoring of all possible decoding options. A professional reference decoder also offers a rack-mount style chassis, professional electrical connections (XLR, AES/EBU, etc.), and comprehensive bass management controls to accommodate various monitoring configurations. For further information, refer to Section 4.11, Using the Dolby Model DP562 Professional Reference Decoder.

2.3 Room Layout, Monitoring, and Calibration

There are many standards and accepted practices as well as different opinions on room layout, monitoring, and calibration in multichannel production and encoding environments. Designs and implementations, therefore, can vary depending on application, speaker selection, and personal preference. Refer to the *5.1-Channel Production Guidelines* available on the Dolby web site at www.dolby.com for additional information.

Chapter 3
Consumer Decoder Products

It is important to note that parameter values set during the Dolby Digital professional encoding process have a direct correlation to decoder behavior. Simply put, encoding is essentially performed for the decoder. This chapter will help the encoding engineer to better understand this relationship.

3.1 Categories

Consumer products incorporating Dolby Digital decoders are classified in two fundamental categories: *Source* and *Decoder*. Requirements for a particular feature can differ depending on whether the product is a Source product or Decoder product.

3.1.1 Source Products

Products designed with the primary purpose of decoding signals from one delivery format, usually the particular medium and format supported by the product, are classified as *Source* products. Source products receive and decode only specific Dolby

Digital bitstream sources and have the source built into the product. Source products must support all bitstream parameters allowed by the particular delivery format.

3.1.2 Decoder Products

Products designed with the primary purpose of decoding bitstreams from external sources are classified as *Decoder* products. Because these products must accept bitstreams from many different sources, Decoder products are generally required to accept the full range of Dolby Digital bitstream parameters.

3.1.3 Channel Output Categories

Products are further categorized by the number of output channels provided: two-channel products and multichannel products.

Two-Channel Products

This category includes two-channel stereo DVD players, DTV sets, or set-top boxes for satellite, cable, or DTV conversion.

All two-channel decoders use Dialog Normalization and require Line Mode Dynamic Range Control (DRC) capability. Source products with RF modulation offer RF Mode processing. Other modes are optional to the product designer. All two-channel products offer Lt/Rt downmix mode; Lo/Ro downmix mode is optional.

Multichannel Decoders

This category includes multichannel A/V amplifiers, receivers, control centers, and preamplifiers. All multichannel decoders include basic bass management capabilities.

Multichannel Adapters

This is a simplified type of decoder for adding Dolby Digital capability to an existing Dolby Surround Pro Logic system. The end result meets all the same basic requirements as those for a complete multichannel system.

3.2 Features

Table 3-1 provides a summary of consumer decoder product features.

Table 3-1 Consumer Decoder Product Features

Feature	Two-Channel Decoder	Multichannel Decoder	Multichannel Adapter	Multichannel DVD Player	Comments
Line Mode	✓	✓	✓	✓	
Dialog Normalization	✓	✓	✓	✓	
Lt/Rt Downmix	✓	✓	✓	✓	
Lo/Ro Downmix	optional	Optional	optional	optional	
Bass Management		✓	✓	✓*	* Simplified design option
Dolby Surround Pro Logic		✓	optional	optional	

3.3 Supported Data Rates

Consumer sources of Dolby Digital bitstreams include NTSC laser disc (LD); digital cable and satellite; digital television (DTV), encompassing standard definition television (SDTV) and high definition television (HDTV); and digital versatile disc (DVD). The maximum data rates for these formats are shown in Table 3-2.

Table 3-2 Dolby Digital Bitstreams Available from Existing Formats

Format	Sample Rate	Data Rate (max)
Laser Disc	48 kHz	384 kb/s
DTV	48 kHz	384 kb/s*
Digital Cable	48 kHz	448 kb/s
Digital Satellite	48 kHz	448 kb/s
DVD-Video	48 kHz	448 kb/s

* A proposal to change the ATSC maximum rate to 448 kb/s has been made.
DVB systems can employ any sample rate with a maximum data rate of 640 kb/s.

Decoders built into any of these source formats are only required to support the sample rate and maximum data rate of that format. Decoders with an IEC 61937 (S/PDIF) input for Dolby Digital bitstreams must be able to accept data rates up to 640 kb/s, and sample rates of 48, 44.1, and 32 kHz, to allow for the possibility of new delivery formats. This requirement does not apply to ATSC-compliant DTV sets, which only need support data rates through 448 kb/s at the 48 kHz sample rate.

3.4 Compatibility

The same encoded multichannel content must play successfully on all decoders in the different product categories. Refer to Figure 3-1.

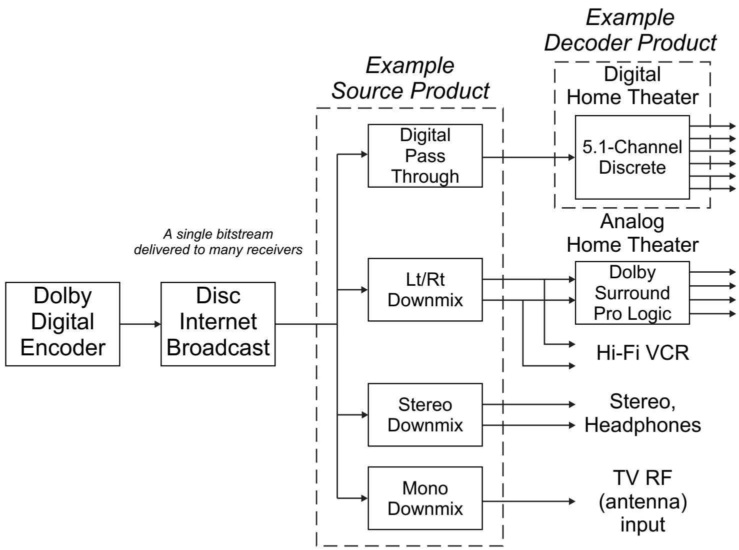


Figure 3-1 Audio Reproduction Hierarchy

Notes to Figure 3-1

(a) **Discrete Multichannel:** The encoded bitstream can be passed through to an A/V system with a 5.1-channel Dolby Digital decoder. It is also possible to find DVD players that provide full 5.1-channel decoding capability.

(b) **Surround Downmix (Lt/Rt):** The bitstream can be downmixed to a two-channel Dolby Surround Lt/Rt compatible format using a preset downmix formula. This downmix can be played over a stereo system, decoded by a Dolby Surround Pro Logic decoder, or recorded onto a VCR for later use.

(c) **Stereo Downmix (Lo/Ro):** The bitstream can be downmixed to a two-channel stereo format using a defined downmix formula with Center and Surround mixing level options. This downmix can be played over a stereo system or headphones.

(d) **Mono Downmix:** The bitstream can be downmixed to a mono format. This downmix can be output from, for example, an RF remodulator in a set-top box. It is equivalent to the Lo/Ro downmix summed to mono.

While all decoders must accept bitstreams made in any Audio Coding Mode (also referred to as Channel Mode) from mono through 5.1, there are options for how these audio programs are processed. In some cases the signals are presented exactly as created, in some cases the signals are downmixed, and in some cases the signals are upmixed by a matrix decoder such as in Dolby Surround Pro Logic or Dolby 3 Stereo. Table 3-3 shows how the various audio program configurations are processed with typical decoding modes. The letter designations correspond with the notes for Figure 3-1.

Table 3-3 Output Signals from Various Program Content and Decoding Modes

Delivered Program Content	Final Reproduction Mode				
	(a) Discrete Multichannel	(b) Dolby Surround Pro Logic	Dolby 3 Stereo	(c) Stereo	(d) Mono
3/2	→ 3/2	↓ 3/1 ↑	↓ 3/0 ↑	↓ 2/0	↓ 1/0
3/1	→ 3/1	↓ 3/1 ↑	↓ 3/0 ↑	↓ 2/0	↓ 1/0
3/0	→ 3/0	↓ 3/1 ↑	↓ 3/0 ↑	↓ 2/0	↓ 1/0
2/2	→ 2/2	↓ 3/1 ↑	↓ 3/0 ↑	↓ 2/0	↓ 1/0
2/1	→ 2/1	↓ 3/1 ↑	↓ 3/0 ↑	↓ 2/0	↓ 1/0
2/0	→ 2/0 or 3/1 ↑	3/1 ↑	3/0 ↑	→ 2/0	↓ 1/0
1/0	→ 1/0	1/0 ↑	1/0 ↑	m 2/0	→ 1/0

→ Denotes a delivered signal passed directly to the output channels.
↓ Denotes a delivered signal downmixed before reproduction.
m Denotes a mono program reproduced by two channels.
↑ Denotes a matrix decoder upmixed from a two-channel signal to derive the output channels.

The Discrete Multichannel column assumes that the Dolby Digital bitstream is being passed through to a multichannel Decoder product. Because a multichannel Decoder product also contains Dolby Surround Pro Logic, a 2/0 program can either be reproduced directly as 2/0 or matrix decoded to produce a 3/1 output.

The Dolby Surround Pro Logic and Dolby 3 Stereo columns assume that the Lt/Rt downmix from a two-channel Source product is being passed to a separate Dolby Surround Pro Logic or Dolby 3 Stereo decoder. No consumer product is required to downmix multichannel programs to Lt/Rt and apply Pro Logic decoding at the same time. In the case of a 1/0 program, the two-channel Source product upmixes the 1/0 program to a 2/0 output. When this 2/0 output is matrix decoded, the result is again a 1/0 output.

To distinguish stereo signals from Dolby Surround Lt/Rt signals, the Dolby Digital bitstream can carry a flag to indicate the format of 2/0 encoded programs. Decoders can use this to drive a Dolby Surround indicator, or it can be used to automatically control the Dolby Surround Pro Logic decoder. Decoders may ignore and/or override this flag, so the user has final control over the stereo or Dolby Surround listening modes.

Decoders output 1/0 mode (mono) signals to either the Center channel or Left and Right channels when the Center channel is not available. It is unnecessary and therefore not recommended to encode mono signals in 2/0 mode.

High-end home theater systems often reproduce the program with as much of the original quality as possible, and use neither downmixing nor Dynamic Range Control. Many listeners use a form of downmixing whenever the number of delivered channels exceeds the number of channels in the Dolby Digital decoder. When downmixing, stereo products operate in Line Mode. Refer to Section 3.6, Operational Modes, for more information.

It is crucial to set Dialog Normalization correctly to avoid the unwarranted application of Dynamic Range Control for peak overload protection. When selecting a Dynamic Range Compression profile (also referred to as a preset) during the encoding process, e.g., film, music, etc., it is important that it matches the program style and provides the intended result. Setting these parameters carefully ensures the average reproduction loudness is consistent, and that the Dynamic Range Control process operates with the correct program thresholds. Refer to Section 4.2, Features, and Appendix C, Dynamic Range Control (DRC), for more information.

When downmixing multichannel programs to mono or Lo/Ro stereo, the relative balance between the source channels is sometimes affected. This can be adjusted to some degree with the Center Mix Level and Surround Mix Level encoder settings.

When downmixing multichannel programs to Dolby Surround compatible format, any Surround signals that are correlated with the front signals, as in a front-back pan, must be phase shifted to ensure the Lt/Rt downmix will decode properly. A ninety-degree phase shift is provided in the encoder. In the vast majority of cases, this filter is inaudible with discrete reproduction.

Whenever downmixing takes place, the LFE signal is discarded. Essential LFE program content must be included in the main Left and Right channels to ensure that it will be heard by all listeners. The LFE channel is never required in a program, and is not an option in mono or stereo modes.

3.5 LFE and Bass Management

Multichannel Dolby Digital decoders offer bass management systems with many options including the ability to redirect bass from any channel where the speaker is unable to reproduce it to the subwoofer. The Low-Frequency Effects (LFE) signal is included in the total signal feeding the subwoofer. Many decoders offer great flexibility in the setting of the bass management options and crossover filters. This ensures that full-frequency content of all the channels in the audio program can be heard. Professional decoders include the same bass management features so it is possible to hear the signals reproduced in the same manner as the home theater listener does. There are two bass management configurations required for consumer products in the Decoder category. Figure 3-2 and Figure 3-3 illustrate these mandatory configurations.

Two-channel products can accept any valid Dolby Digital bitstream, including multichannel programs. Programs with more than two channels are automatically downmixed as described in Section 3.4. The LFE channel, if present in the encoded audio program, is always omitted during playback from a two-channel product.

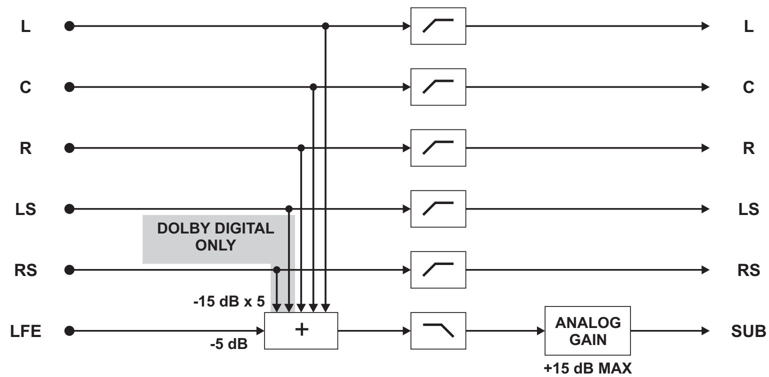


Figure 3-2 Decoder Product Bass Management Configuration One

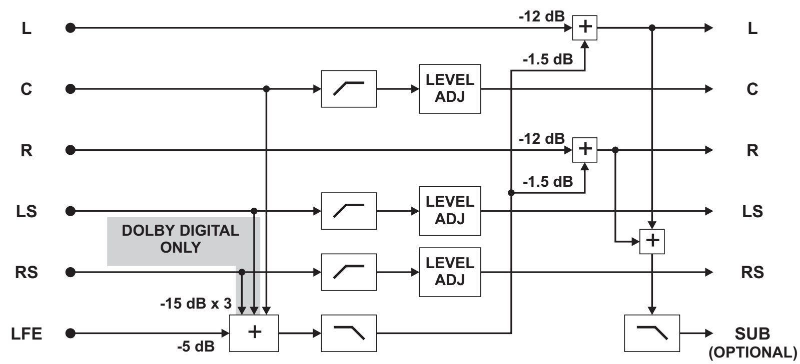


Figure 3-3 Decoder Product Bass Management Configuration Two

3.6 Operational Modes

To ease the design of consumer and professional decoder products, Dolby Digital integrated circuits (ICs) offer standard Operational Modes (also referred to as Dynamic Compression Modes) called Line Mode and RF Mode. This greatly simplifies the implementation of Dialog Normalization, Dynamic Range Control, and downmixing functions, all of which are necessary in Dolby Digital products.

Custom Modes are also available although rarely implemented. Custom Modes 1 and 0 can be used in addition to or instead of Line Mode to offer additional or enhanced functionality. Settings equivalent to the required settings must still be provided.

3.6.1 Line Mode

Summary of Line Mode features:

- Dialog Normalization enabled
- Dialog reproduced at a constant level of -31 dBFS LAeq (3 dB lower in each channel when downmixed to two-channel or mono); refer to Section 4.2.1, Dialog Normalization (*dialnorm*), for more information.
- *dynrng* compression variable used for Dynamic Range Control
- High-level cut compression scaling allowed when not downmixing
- Low-level boost compression scaling allowed

Line Mode is applicable in the widest range of products due to its flexibility and ease of use. All line level or power amplified outputs from two-channel set-top decoders, two-channel televisions, 5.1-channel televisions, A/V Surround decoders, and outboard Dolby Digital adapters should be derived from this mode.

Figure 3-4 shows the signal relationships of Line Mode under different conditions. Note that whether or not downmixing or Dynamic Range Control is active, the average program loudness remains constant.

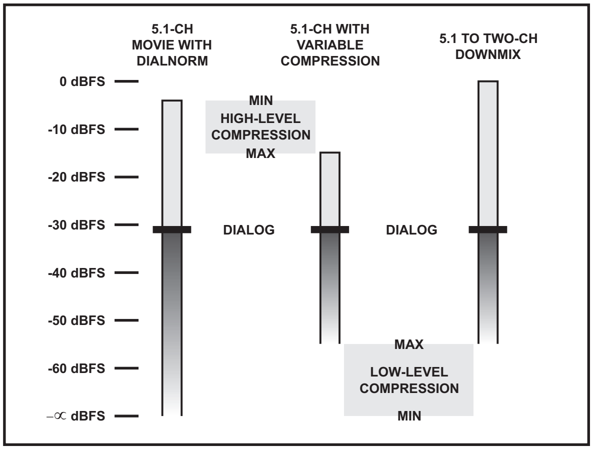


Figure 3-4 Signal Relationships in Line Mode

3.6.2 RF Mode

Summary of RF Mode features:

- Dialog Normalization enabled
- Dialog reproduced at a constant level of -20 dBFS LAeq (3 dB lower in each channel when downmixed to two-channel or mono); refer to Section 4.2.1, Dialog Normalization (*dialnorm*), for more information.
- *compr* compression variable used for Dynamic Range Control (*dynrng* used if *compr* does not exist)
- High- and low-level compression scaling not allowed (always fully compressed)
- +11 dB gain shift to raise overall program level

RF Mode is optimized for products (i.e., set-top boxes) that generate a downmixed signal for transmission to the RF (antenna) input of a television set. The overall program level is raised 11 dB, while the peaks are limited to prevent signal overload in the D/A converter. By limiting headroom to a maximum of 20 dB (3 dB greater in each channel when downmixed to two-channel or mono) above average dialog level, severe overmodulation of television receivers is prevented while providing a dialog RF modulation level that compares well with quality television broadcasts and premium movie channels. Figure 3-5 shows the signal relationships of RF Mode under different conditions. Note that whether or not downmixing is active, the average program loudness remains constant. Dynamic Range Control remains fully on at all times.

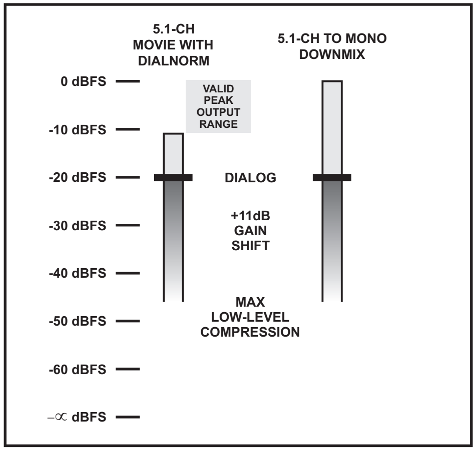


Figure 3-5 Signal Relationships in RF Mode

Refer to Figure 3-6, which shows an example of how a signal generated with RF and Line Modes can relate to the modulation index of a typical RF modulator circuit. While Line Mode can be used for this purpose, the improvement in program dynamics comes with a lower average loudness than other television signal sources.

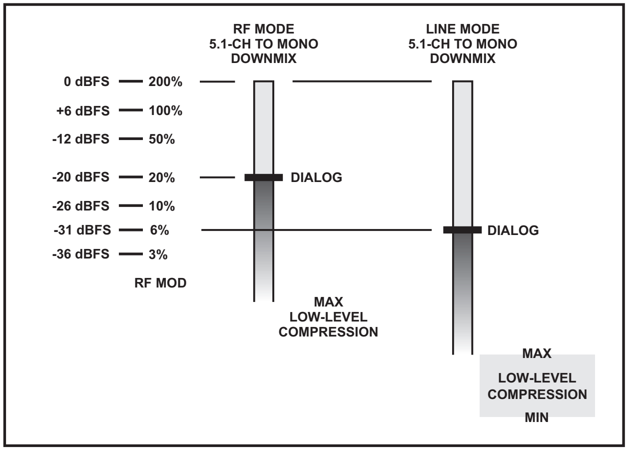


Figure 3-6 RF Modulator Signal Levels

3.6.3 Using Operational Modes in Products

More than one Operational Mode can be used in a product, and they can be used in different ways depending on the type of product. Line Mode is intended for products providing line level or power amplified outputs and is used in the majority of products. RF Mode is intended primarily for products generating a downmixed signal delivered to the RF (antenna) input of a television set, but can also be used in conjunction with Line Mode to provide additional listening options.

Dynamic Range Control (DRC)

Dynamic Range Control (DRC) is required in all Dolby Digital products, but the degree of user control differs depending on the product type. Some products require Dynamic Range Control to suit the particular listening situation for which it was designed. Other products support a variety of listening situations and offer various dynamic range compression settings with the ability to set and store preferences.

Dynamic Range Control (DRC) for a Source Product

Table 3-4 shows the Dynamic Range Control (DRC) requirements, recommendations, and options for a typical Source product.

Table 3-4 DRC for a Typical Source Product

Setting		Operational Mode	Scale Factors (high/low)	Gain Correction Required
Optional	Maximum dynamic range	Line	0.0/0.0	None
Required	**Standard dynamic range**	**Line**	**1.0/1.0**	**None**
Recommended	Minimum dynamic range	RF	no scaling allowed	None

A Source product, such as a cable or satellite receiver, or DVD player, needs the ability to conform its audio output to traditional signal references, such as VHS Hi-Fi tapes or broadcast television signals. The standard dynamic range setting is required in all Source products and results in audio signals with good dynamic range, much like a prerecorded VHS Hi-Fi movie. The majority of users who connect the audio outputs to a Dolby Surround Pro Logic system do so with this setting. For example, it may be the mode of choice for people who connect the audio to a Dolby Surround Pro Logic system.

Other users may feel that the average loudness of the signal is too low when switching between Dolby Digital programs and regular television programs, or that the loud portions of the program are too loud compared with the average dialog level. In these cases, the minimum dynamic range setting is recommended as it will raise the average loudness of the dialog and restrict the program peaks, much in the style of conventional television audio.

The optional maximum dynamic range setting is meant for use in multichannel Source products to provide the widest dynamic range when the product is not downmixing. Products that must downmix to two channels will not be able to reproduce the full dynamic range in Line Mode because there are restrictions to prevent digital overload.

For Source products that offer an RF modulated output in addition to line outputs, the minimum dynamic range setting based on RF Mode becomes a requirement instead of a recommendation, as shown in Table 3-5.

Table 3-5 DRC for a Source Product with RF Modulated Output

Setting		Operational Mode	Scale Factors (high/low)	Gain Correction Required
Optional	Maximum dynamic range	Line	0.0/0.0	None
Required	**Standard dynamic range**	**Line**	**1.0/1.0**	**None**
	Minimum dynamic range	**RF**	**no scaling allowed**	**None**

Source products that offer only an RF modulated output with no other outputs can reduce the DRC requirement to just the minimum dynamic range setting based on RF Mode.

Dynamic Range Control (DRC) in Decoder Products

Table 3-6 shows the Dynamic Range Control (DRC) requirements and recommendations for a typical Decoder product.

Table 3-6 DRC for a Typical Decoder Product

Setting		Operational Mode	Scale Factors (high/low)	Gain Correction Required
Required	**Maximum dynamic range**	**Line**	**0.0/0.0**	**None**
	Standard dynamic range	**Line**	**1.0/1.0**	**None**
Recommended	Minimum dynamic range	RF	no scaling allowed	-11 dB

A Decoder product used in a home theater setting must be able to adapt to different listening situations and user preferences. Both the standard and maximum dynamic range settings are required so that the user has the ability to reproduce the audio program with either the full or limited dynamic range intended by the producer.

The minimum dynamic range setting is also recommended for situations such as late night viewing at reduced volume levels wherein low-levels signals must be brought up to be heard, but peak level must be brought down so as to not disturb others. A -11 dB gain correction is required when using RF Mode in a Decoder product so that dialog is reproduced at a level consistent with the Line Mode output.

If a manufacturer wants to offer more flexible control for adjusting the amount of dynamic range compression that is applied, a linearly variable control that adjusts the scale factors from 0.0 to 1.0 can be used, as shown in Table 3-7.

Table 3-7 Alternative DRC for a Decoder Product

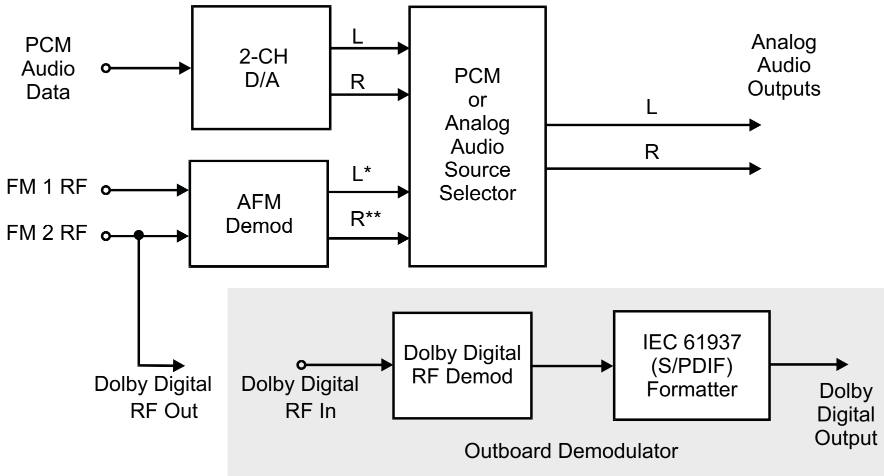
Setting		Operational Mode	Scale Factors (high/low)	Gain Correction Required
Required	**Variable compression**	**Line**	**(0.0-1.0)/(0.0-1.0)**	**None**
Recommended	Minimum dynamic range	RF	no scaling allowed	-11 dB

A single control can adjust both high- and low-level scaling in tandem, or there can be separate controls for each. A step size of no less than 0.1 is recommended. For example, with a step size of 0.2, a six-position control is possible. With a step size of 0.5, a three-position control is possible. This type of control must include both maximum dynamic range (0.0/0.0) and standard dynamic range (1.0/1.0).

3.7 Laser Disc

Laser disc is unique in that it is an established format already capable of delivering two-channel audio from the original analog FM tracks (AFM), and also from the two-channel 16-bit linear PCM digital audio tracks. Dolby Digital compatible laser disc players are able to provide both conventional stereo audio and Dolby Digital bitstreams without an internal Dolby Digital decoder. This is because both linear PCM tracks remain available and the Dolby Digital bitstream will usually represent the discrete multitrack version of the same audio program found in those tracks.

To add Dolby Digital bitstreams in a way that preserves as much of the existing format as possible, the Right channel AFM area carries the Dolby Digital RF signal. It is a QPSK modulated carrier at 2.88 MHz. The data is Reed-Solomon coded for error correction. This RF signal has a separate output from the player for external decoding, and is demodulated into a standard Dolby Digital bitstream in the S/PDIF format for connection to Dolby Digital A/V decoders, as shown in Figure 3-7.



*The Left channel of the AFM Demod ouput contains a mono mix delivered by FM 1 RF

**The Right channel of the AFM Demod contains a Dolby Digital bitstream delivered by FM 2 RF

Figure 3-7 Basic Laser Disc Player Structure

Chapter 4
Encoding

4.1 General Information

4.1.1 Preparing the Source Delivery Master

When preparing the source delivery master, adhere to accepted standards and practices to ensure proper Dolby Digital encoding.

One of the most common source delivery formats for Dolby Digital encoding is the Hi-8 mm tape used in many popular Modular Digital Multitracks (MDMs). Digital Audio Workstations (DAWs), open reel digital multitracks, and other formats are also used for this application, although less frequently.

Channel-to-Track Allocation

Dolby encourages adopting the channel-to-track allocation described in the ITU-R recommendation, *Parameters for Multichannel Sound Recording,* and in SMPTE standard 320M. Track layouts depend on channel complement although tracks 1, 2, and 3 are always channels Left (L), Right (R), and Center (C) respectively. Table 4-1 shows one possible configuration. Since inclusion of the LFE channel is optional and the listener determines its reproduction from a decoder, essential low-frequency information should not be mixed exclusively to the LFE channel. Alternative practices exist within various industries so it is important to check the source and accompanying documentation.

Table 4-1 Channel-to-Track Layout Example

Channel	L	R	C	LFE	LS	RS	Lt	Rt
Track	1	2	3	4	5	6	7	8

Channel Levels

The following text is in accordance with the ITU-R recommendation and SMPTE standard referred to in the above section.

For consumer and DVD production studios, relative channel levels assume each speaker delivers identical acoustic sound pressure levels to the listener. This excludes the LFE channel, which is intended for reproduction at +10 dB SPL (with respect to the main channels within the same 3 Hz to 120 Hz passband). Assuming that a Surround (S) signal is delivered to a single speaker and two Surround signals (LS, RS) are each delivered to individual speakers, Surround levels should be identical to those for the front channels.

In film sound practice, stereo Surround channel levels are typically recorded +3 dB relative to the front channels. This is done to compensate for the -3 dB Surround levels (relative to the front channel levels) encountered in cinema audio monitoring systems. Calibrating cinemas in this manner allows for compatibility with other soundtrack formats. Soundtracks mixed in film rooms require selecting the “3 dB Attenuation” option for the LS and RS channels in the Dolby Digital encoder to compensate for the difference in calibration.

When the Surround channel is mono, allocate it to both tracks 5 and 6 with 3 dB of attenuation applied to each signal. Use the following formula:

$$\text{Track 5} = \text{Track 6} = 0.707 * S$$

Follow this recommendation even when track 4 also contains the S signal, which should always be at normal level on this track. Label the tape clearly to indicate tracks 5 and 6 each contain the S signal at -3 dB relative to their normal levels.

Reference Levels

The standard reference level is -20 dBFS for digital recorders (0 VU for analog recorders). This level is typically +4 dBm from professional consoles and -10 dBV from semi-professional consoles. When transferring from analog 35 mm magnetic film, attenuation and/or peak limiting may be needed to avoid digital overload. These processes require selecting a different Dialog Normalization value in the encoder and necessitate complementary gain recovery in the reproduction chain.

A 30-second, 1 kHz alignment signal at -20 dBFS should appear on all channels at the beginning of the source delivery master prior to program start. The finished master should contain at least 30 seconds of "digital black" after the alignment signal and before each subsequent program. Each title should begin with at least two seconds of encoded digital black.

Documentation

Complete, clear, and accurate documentation should always accompany the source delivery master used for Dolby Digital encoding. This information is important not only when the master is in use but also for reference once it is archived. Dolby has created Mix Data and Mastering Information sheets to facilitate proper documentation or to use as a guide for creating similar documents. The two documents correspond with the stages at which they are used in production. These sheets are included in Appendix E, Mix and Mastering Data Sheets, and are available on the Dolby web site under the Technical Information heading at www.dolby.com.

The Mix Data sheet provides concise information about the source media to all the engineers on a project. Typically, it includes information on sampling frequency, bit resolution, time code, track assignment, titles, and program start and stop times. In addition to being documented on the form, all mix data information should be duplicated and attached to the source delivery master.

The Mastering Information sheet provides documentation relevant to the mastering engineer or authoring facility on source media, timing, and encoder settings as well as general notes. This sheet can be used as a setup guideline for proper Dolby Digital encoding.

These documents do not guarantee success, but are a starting point for customizing a document(s) specific to the task(s) you perform. If you have questions about these documents, refer to Section 7.2, Contacting Dolby Laboratories.

4.1.2 System Operation

There are many issues to consider before beginning Dolby Digital encoding. When setting parameters it is important to take into account the type of content and the distribution media. For example, DVDs and laser discs require different encoding parameter settings. Following is a brief description of some of the issues that production and authoring engineers face when generating Dolby Digital encoded content.

Whether an engineer uses a real-time encoder or a non-real-time encoder significantly affects the production process. Real-time encoders, although generally more expensive, offer the ability to check and monitor the Dolby Digital bitstream as it is being encoded, saving the engineer a significant amount of verification time. A non-real-time encoder generally offers the capability of batch encoding so the engineer can run multiple sessions overnight and automate the encoding process.

When encoding for non-real-time applications, the process results in a Dolby Digital file referred to as an .ac3 file. An .ac3 file adheres to the standard file format defined by Dolby Laboratories for Dolby Digital files. All DVD authoring systems and the majority of Dolby Digital encoding products on the market today are capable of processing the .ac3 file format.

Audio/video synchronization is a concern that engineers need to address with every production. Although the use of SMPTE time code can help synchronize the audio and video components of a production, occasionally slight timing adjustments need to be made after the material has been assembled. Although an engineer may author a DVD with an exact match in time code, differences in decoder latencies can produce a noticeable discrepancy in audio and video synchronization when playing the final production disc. Most authoring tools today offer a way to emulate a DVD before production in order to adjust for such problems.

Latency is an issue when encoding simultaneous audio and video in real-time for broadcast or DVD authoring. Since every real-time encoder has an inherent latency, it is important to analyze the latencies of both the audio and the video encoders and make appropriate delay adjustments to match the elements. When encoding in real-time for the purpose of storing a file to disk, latency is irrelevant.

Most real-time encoders allow the user to start and stop an encode session at predetermined time code points. For this to occur, the source material must carry SMPTE time code in addition to the audio tracks. Punching in and out using time code is advantageous during encoding since it eliminates the need to edit the .ac3 files later.

4.2 Features

An important aspect of Dolby Digital is that it caters both to the critical and to the casual listener. The former may wish to hear precisely what the mixer heard in the studio; the latter may want a processed form of audio resembling current broadcast practice. Neither would, however, voluntarily choose a system requiring continual adjustment of the volume control, as is demanded by the present mixed formats (CD, cassette, FM, AM, TV, DVD, etc.).

Dialog Normalization, Dynamic Range Control (DRC), and downmixing are interdependent features and thus during encoding they cannot be treated separately. Since Dolby Digital appears as a sound format in many different media and the listener will want to switch between these media without dramatic volume changes, it is necessary to consider all types of programming. In some applications, e.g., set-top boxes for cable and/or satellite distribution, these Dolby Digital features also permit matching loudness with present analog broadcast sources.

The effects of Dialog Normalization, Dynamic Range Control, and downmixing should be assessed during program origination, preferably by monitoring through a Dolby Digital encoder/decoder, so that the mixer can simulate worst-case as well as best-case listening conditions.

Dynamic Range Control cannot be properly applied or assessed without correctly setting Dialog Normalization, since some of the Dynamic Range Control parameters depend upon this value.

A note about terminology: Encoding parameters affect the presentation of the audio program. During encoding, parameter values are embedded into the bitstream that contains the coded audio. SMPTE and other technical organizations have adopted a special term for data that is packaged or transmitted with program material, *metadata*. The term loosely means "data about the data" and is intended to distinguish program material, referred to as the audio or video essence, from the data that controls or describes it. In recent technical papers and presentations, Dolby Laboratories has begun referring to these encoding parameters (Dialog Normalization, Dynamic Range Control, etc.) using the general term metadata.

4.2.1 Dialog Normalization (*dialnorm*)

For the purpose of Dialog Normalization (also referred to as volume normalization) loudness is currently quantified using the equivalent loudness method LAeq, the long-term average of A-weighted sound pressure. The LAeq measurement correlates more closely with subjective loudness but yields figures lower than VU meter readings. The most useful measure for dialog level is the ratio of the LAeq measurement to digital full-scale. Readings taken in this manner are noted as dBFS LAeq.

The following examples assume that program material entering the Dolby Digital encoder has the dynamic range it received at its source: It is as the producer intended and has not been further processed. A further assumption is that Line Mode is employed with no downmixing.

Consider a listener switching between a news bulletin, a wide-range movie, rock music, and a symphony orchestra. These may be different TV channels or recordings, or successive items from one source. In order of magnitude, these items will have loudness values of about -14, -28 (the dialog), -8, and -25 dBFS LAeq. Thus if a listener sets playback level (using a volume control) to the news bulletin and then switches to the movie, its dialog will be about 14 (28–14) dB quieter than the voice of the newsreader, and probably unintelligible. Conversely if the listener sets a playback level appropriate for the movie, the rock music will be reproduced 20 (28–8) dB higher than the dialog, probably intolerably loud. In fact with these items the typical listener will need to adjust the playback level over at least a 15 dB range.

Note that the quietest source is the movie dialog. For many years, movie mixers have used a standardized acoustic level for dialog; with digital formats this is equivalent to between -25 and -31 dBFS LAeq. Since movies constitute an important part of the material to be conveyed by Dolby Digital, this standardization is retained, and all dialog should emerge from a Dolby Digital decoder at about -31 dBFS LAeq. The average level of programming other than dialog should be adjusted appropriately relative to dialog.

The Dolby Digital encoder sends a control word called *dialnorm* to command the decoder to adjust the playback level. In other words, *dialnorm* acts as an automatic volume control. In the example above, *dialnorm* needs to command about 17 dB of attenuation on the news bulletin, only 3 dB on the movie, perhaps 15 dB on the rock music (so that it is a little louder than speech), and about 6 dB on the symphony orchestra.

For speech, the *dialnorm* figure is the equivalent loudness level with respect to full-scale, and the attenuation introduced in the decoder is (31 + *dialnorm*) dB with *dialnorm* being negative. Thus speech with an LAeq of -31 dBFS should have -31 entered; this commands 0 dB of attenuation in the decoder. Similarly, the news bulletin with speech at -14 dBFS LAeq requires a *dialnorm* setting of -14, giving 17 dB of attenuation so that speech from the newsreader comes out of the decoder at -31 dBFS, matching the movie dialog.

If the source material is recorded at a lower level resulting in peaks that do not approach digital full-scale, less attenuation is needed. Thus if the news bulletin used in the example above had a loudness of -20 dBFS LAeq rather than -14 at its source, a *dialnorm* setting of -20 would yield the standard level (-31 dBFS) at the decoder output and a match other speech.

The desired attenuation (volume normalization) is performed at the decoder under the control of *dialnorm* sent from the encoder. The actual audio data is not modified, and therefore the original program level is available at the decoder. Most if not all Dolby Digital consumer decoders automatically normalize the playback volume using the Dialog Normalization feature.

In general there can be no default setting for *dialnorm*; the value depends on the nature of the program, and in the context of mixed programming it is essential for the setting to change from item to item. For a channel with uniform material, a fixed (but appropriate) setting may be acceptable. Generally a setting for *dialnorm* of -31 is unusual, required only for a few unprocessed wide-range movie soundtracks. For typical broadcast material (speech and popular music), the setting lies more often in the range of -15 to -20.

In the future, Dolby Laboratories expects that the program material arriving at the Dolby Digital encoder used for broadcast transmission will be accompanied by metadata containing appropriate values for *dialnorm* (*dynrng* and *compr* as well), so that no manual intervention at the emission stage will be required. Currently, parameter values such as *dialnorm* must be set at the Dolby Digital encoder.

For programs including speech, the correct figure for *dialnorm* is the dBFS LAeq value of that speech. Thus an objective setting involves measuring LAeq and entering the result at the encoder. For non-speech items, it may initially be desirable to ask a number of people to listen to speech at the standard level and then to adjust the music volume to their taste; the amount of the adjustment can then be used as an offset to the speech setting.

The above examples apply to the Line Mode output of the decoder. Dolby Digital decoders can also operate in RF Mode, delivering an output intended primarily to feed TV sets or VCRs via an RF modulator, typically mono. It is desirable that Dolby Digital sources reproduced in this manner match analog sources (broadcast and cable channels and VCR recordings), and the potential wide dynamic range of Dolby Digital is usually inappropriate. In RF Mode the decoder switches in an output boost of 11 dB relative to line output modes, and applies Dynamic Range Control. For more information, refer to Dynamic Range Compression for RF Outputs (*compr*), in the next section. The standard gain setting for the RF modulator (digital full-scale =

200% of nominal maximum broadcast modulation) and correct selection of *dialnorm* typically provide such a match.

4.2.2 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 *dialnorm*. 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 *dialnorm*, 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 (refer to Profiles later in this section) 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.

Dynamic Range Compression for Line Outputs (*dynrng*)

Dolby Digital encoders generate Dynamic Range Compression profile information in accordance with one of a number of selectable algorithms. In Line Mode, this information (along with overload protection limiting) is contained in control words called *dynrng*, accompanying the full dynamic range audio and used in the decoder to apply dynamic range compression at the listener's option. Some consumer decoders also offer the option of using partial dynamic range compression.

As mentioned above, some types of programming have little inherent dynamic range and therefore remain close to their average most of the time. The Dynamic Range Compression profile algorithms provided in Dolby Digital encoders have a *null band*, an intermediate range of input levels where no gain or loss is applied. Thus programming with narrow dynamic range is substantially unchanged. Above the null band, the algorithms command attenuation in the decoder; below, they command amplification. The width of the null band, the degrees of high- and low-level compression and the time constants depend on the selected algorithm.

Average program levels should lie within the null band. *Dialnorm* represents the average loudness of the input signals. Hence the thresholds of the various Dynamic Range Compression profile algorithms (for example, the bottom and top of the null band) are set relative to the *dialnorm* reference playback level. Refer to Appendix C, Dynamic Range Control (DRC), for more information.

Dynamic Range Compression for RF Outputs (*compr*)

RF Mode employs a different dynamic range compression control word, *compr*. This operates similarly to *dynrng*, apart from the overload protection. Refer to Appendix C, Dynamic Range Control (DRC), for more information.

Profiles

There are several Dynamic Range Compression *profiles* (also referred to as presets), each with a name denoting the most suitable application. They all share the property of a null band, a region in the middle of the dynamic range where the gain is fixed at unity, no boost or cut. The ends of this null band are referred to the value of *dialnorm* so that dialog or average program lies within the null band and is not subject to gain variation. Unless the feature is disabled at the decoder, sounds quieter or louder than the average and outside the null band are boosted or cut in accordance with the profile selected at the encoder. This provides a reduction in the dynamic range of the reproduced audio. The width of the null band determines the proportion of time that program level lies outside it, and therefore the actual amount of dynamic range compression.

One result of implementing dynamic range compression in this way is that program material with an already restricted dynamic range, whether inherent or because of prior processing, lies primarily within the null band and hence is not subjected to further compression. Another outcome is that dialog does not modulate background noises, at least between syllables. The resultant processing reduces dynamic range without the audible side effects often associated with broadcast processors, gain pumping and transient distortion.

Irrespective of the selected Dynamic Range Compression profile, the encoder assesses the possibility of peak overload in the decoder. If necessary, the encoder overrides and increases the high-level compression gain word, preventing output overload. Refer to the section on Downmixing and Overload Protection for more information.

Currently, there are five Dynamic Range Compression profiles in addition to "None." These profiles are divided into three groups that are described below. For more detailed information refer to Appendix C, Dynamic Range Control (DRC).

1. Film

Movie soundtracks contain dialog at a standardized level with respect to digital full-scale. Sounds are rarely much quieter than dialog (they would not be audible beneath the typical background noise of a movie theatre), so film soundtracks only call for modest degrees of low-level boost. In addition, raising low-level sounds excessively would sometimes reveal unwanted background disturbances, such as camera and traffic noise, which the film mixer did not intend to be audible in the theatre. Sounds are, however, frequently much louder than dialog, by 20 dB or more, so large amounts of gain reduction may be required.

There are two film profiles, standard and light, with null bands 10 and 20 dB wide respectively, straddling the *dialnorm* setting. In both cases, low-level boost is applied using a 2:1 compression ratio, with a maximum boost of 6 dB. Above the null band, the compression adopts a characteristic (20:1 ratio) close to limiting, except that it is based on RMS, not peak.

2. Music

The dynamic range of music varies according to type. Most popular music has an inherently limited dynamic range, and requires little or no compression. It does, however, demand an appropriate setting for *dialnorm* to ensure that its absolute loudness is not out of line with that of other programming. The *dialnorm* setting also determines whether the music will lie within the null band of the Dynamic Range Compression profiles.

As with film, there are two music profiles, standard and light, with null bands 10 and 20 dB wide respectively, straddling the *dialnorm* setting. Below the null band, both profiles offer up to 12 dB of boost with a 2:1 compression ratio. Above the null band, a standard profile gives 20:1 compression and light 2:1.

3. Speech

While any one source of speech usually has a limited dynamic range and can be easily accommodated inside a null band, some speech programming may include moments that are abnormally loud or soft, such as shouts or whispers. The speech profile uses a 10 dB null band for average speech. The correct setting of *dialnorm* usually ensures that average speech lies within the null band; outside this region fairly heavy compression is applied: 5:1 ratio up to 15 dB boost for low levels and 20:1 (near limiting) for high levels. This technique retains the impression of quieter or louder speech while ensuring that non-average voices remain intelligible and do not get excessively loud.

The speech profile may be inappropriate when large amounts of background audio accompanies the speech, as gaps in the speech would boost the background by 15 dB. In these circumstances, one of the film profiles may be more suitable.

Downmixing and Overload Protection (*dynrng* or *compr*)

When multiple sources are mixed together in a studio, the engineer adjusts the relative gain of each source for the desired subjective balance and the overall gain for the desired total output level.

In contrast, when a multichannel program is downmixed inside a Dolby Digital decoder, the mixing coefficients are in general fixed. The decoder performs downmixing in the digital domain (except from two-channel to mono), so there is the possibility that the downmix will overload the output digital-to-analog converters (DACs).

If the coefficients in the decoder were chosen to ensure that downmixes could never overload the DACs, many downmixes would sound somewhat quieter than the same program reproduced in a multichannel mode or a mono or stereo program that did not require downmixing.

The actual mixing coefficients are therefore chosen to give a more satisfactory match in output volume between downmixed and non-downmixed sources. As a result downmixes could lead to output overload on the comparatively rare occasions that a multichannel source approached digital full-scale on all channels simultaneously.

As described earlier, most programming demands volume attenuation within the decoder via dialnorm. This attenuation is applied in the digital domain prior to downmixing, and therefore reduces the probability of overload. In RF Mode, though, the 11 dB of boost provided by the decoder does increase the probability of downmix overload.

Dolby Digital prevents overload by invoking overload protection limiting during the encoding process. The encoder generates several possible downmixes, estimates the worst-case peak level at a decoder output, taking into account the setting of *dialnorm*, and separately calculates the gain reduction required in the decoder to prevent overload in Line Mode and RF Mode. Optionally, the gain reduction for RF Mode can take into account the pre-emphasis employed in RF modulators. Whenever one of these values of gain reduction exceeds the gain reduction demanded by a selected Dynamic Range Compression profile, if any, it is substituted for the compression value in the *dynrng* or *compr* control word.

To put overload protection in perspective, consider the extraordinarily improbable case of a five-channel source that reached full-scale simultaneously in all channels; this represents a sound roughly 30 dB higher than standard dialog level.

In Line Mode, due to the choice of fixed mixing coefficients, overload prevention would require 11 dB of gain reduction from *dialnorm* and *dynrng* combined. Such a source would obviously be very loud and would probably already have demanded volume reduction via *dialnorm* and dynamic range compression via *dynrng*. If such reduction was 11 dB or more, no protection limiting would be required. In practice, overload protection limiting operates rarely, if at all, on real programs.

In RF Mode, this source could demand 22 dB of total gain reduction (*dialnorm* plus *compr*). Considering that in analog broadcasting average speech is typically within a few dB of full modulation, it is not surprising that a sound 30 dB louder demands large degrees of dynamic range compression and overload protection limiting.

4.3 Metering

Metering is not available on all Dolby Digital encoders and is generally implemented in software. There can also be meters other than those described in this section.

4.3.1 Input Level Meter

The meter marked Input Level shows the input signal level that the compressor is using to determine the amount of boost or cut. The displayed signal level is derived from the largest of the individual channel signal levels (i.e., the input level meters) and incorporates the attenuation applied by the *dialnorm* value.

4.3.2 Line Mode Meter

The meter marked Line Mode shows the value of *dynrng* being sent in the bitstream. This is the amount of Dynamic Range Control that the decoder applies if it is configured in Line Mode, and no dynamic range compression scaling is applied. Green means boost, red means cut, and the meter shows a range of -24 dB to +24 dB.

4.3.3 RF Mode Meter

The meter marked RF Mode shows the value of *compr* being sent in the bitstream. This is the amount of Dynamic Range Control that the decoder applies if it is configured in RF Mode. Green means boost, red means cut, and the meter shows a range of -24 dB to +24 dB (although *compr* can take on values as large as +/- 48 dB).

Frequently the RF Mode meter reads about the same as the Line Mode meter, indicating that *dynrng* and *compr* are being set to comparable values. If signal conditions arise that would cause digital overload in the decoder, the protection circuits in the encoder activate and cause *dynrng* and *compr* to be further limited. In these conditions, *compr* tends to be more limited than *dynrng*, and thus the RF Mode meter shows a larger amount of cut than the Line Mode meter.

4.3.4 Calibration of Dialog Normalization (*dialnorm*)

The preferred method for determining the Dialog Normalization (*dialnorm*) value for speech is to use an LAeq meter. When this is impractical or impossible, the Line and RF Mode meters can be used as a starting point to calibrate *dialnorm*. With only speech being sent through the encoder, and using the appropriate Dynamic Range Compression profile (also referred to as a preset), an operator can adjust *dialnorm* until the meters read 0 dB or show *reasonable amounts of boost and cut*. The same procedure can also be used for programs that do not contain speech, such as music. Subjective listening should then be used to set the final value. It is hoped that in the future a more practical and effective solution for metering will be available for this application.

Dolby Laboratories has created a *Dolby Digital Dialog Normalization Disc* with examples that can be used as references to assist in subjectively determining the appropriate *dialnorm* value. For more information on this disc, refer to Section 7.2, Contacting Dolby Laboratories.

4.4 Parameter Default Values

Although Dolby Digital professional encoders typically have preset default values, parameter settings nearly always require adjustment to appropriate values for specific content and applications.

In addition to the Data Rate and Audio Coding Mode (also referred to as Channel Mode), the encoding engineer should pay particular attention to the Dialog Normalization, Dynamic Range Compression profile, Surround Channel 3 dB Attenuation, and 90-Degree Phase-Shift parameter settings. In addition, Input/Output Control should be verified to ensure the appropriate clock source, input channel assignment mode, and output format selections.

Dolby Laboratories recommends the default values given in Table 4-2 for the parameters that are required and therefore common to all licensed encoders. This is not to say that the same default values appear in all encoders, even if they are from the same manufacturer. For some parameters, two-channel encoder defaults differ from those for a 5.1-channel encoder. Refer to Chapter 5, Applications and Formats, for more information on appropriate parameter values for various applications.

Table 4-2 Recommended Parameter Default Values

Audio Service	
Data Rate — two-channel	192 kb/s
Data Rate — 5.1-channel	448 kb/s
Audio Coding Mode — LFE Channel, two-channel	2/0, LFE disabled or off
Audio Coding Mode — LFE Channel, 5.1-channel	3/2, LFE enabled or on
Bitstream Mode	Complete Main
Dialog Normalization	-27 dB
Bitstream Information	
Center/Surround downmix levels	-3 dB
Dolby Surround Mode	Not indicated
Audio Production Information Exists	No or 0
Audio Production Information — Mixing level	25
Audio Production Information — Room type	Small room
Copyright Bit	Copyright protected or 1
Original Bitstream	Original or 1

Table 4-3 Recommended Parameter Default Values - *continued*

Preprocessing	
DC Highpass Filter	On
Channel Bandwidth Lowpass Filter	On
LFE Lowpass Filter	On
Surround Channel 90-Degree Phase-Shift — two-channel	N/A
Surround Channel 90-Degree Phase-Shift — 5.1-channel	On
Surround Channel 3 dB Attenuation — two-channel	N/A
Surround Channel 3 dB Attenuation — 5.1-channel	Off
Dynamic Range Compression profile	On or Film Standard
RF Overmodulation Protection (RF Pre-emphasis filter)	On

4.5 Audio Service

The parameters in the Audio Service specify the fundamental aspects of the Dolby Digital encoded program. They include the Data Rate, Sampling Rate, Audio Coding Mode, Bitstream Mode, and the Dialog Normalization value.

4.5.1 Data Rate

The total Dolby Digital bitstream data rate is set using the bitstream Data Rate parameter. The bitstream Data Rate value determines which of the 19 pre-defined Dolby Digital data rates will be used. In order to maintain high audio quality, data rates that are supported by a Dolby Digital encoder depend on the selected Audio Coding Mode parameter. In general, Audio Coding Modes that include fewer channels in the bitstream have lower data rate limits. Table 4-4 shows examples of the supported and suggested data rates as a function of the Audio Coding Mode. The suggested data rate assumes normal high-quality audio. For speech, low-bandwidth audio, or when constrained by standards, it may be appropriate to use lower data rates than those suggested.

Table 4-4 Supported and Suggested Data Rates According to Audio Coding Mode

Audio Coding Mode	Supported Data Rates	Suggested Data Rate
1/0	56–640 kb/s	96 kb/s
2/0 or 1+1	96–640 kb/s	192 kb/s
3/1 or 2/2	192–640 kb/s	384 kb/s
3/2	224–640 kb/s	448 kb/s

4.5.2 Audio Coding Mode

The Audio Coding Mode (also referred to as Channel Mode) parameter defines the number of main audio channels within the encoded bitstream and also indicates the channel format. The Audio Coding Mode is designated as two numbers, *m/n*, with *m* indicating the number of front channels, and *n* indicating the number of rear (Surround) channels. If the mode is set to 1+1, then two completely independent program channels (dual-mono), referenced as Ch1 and Ch2, are encoded into the bitstream. *Note that the 1+1encoding mode is not allowed in ATSC DTV format, nor in the DVD-V format*. If the program material is encoded as 1/0 (mono), decoders output the signal to either the Center channel or both Left and Right channels depending on the system configuration. Table 4-5 lists all eight modes and defines which input channel is used for encoding based on the selected mode.

Table 4-5 Audio Coding Mode

Audio Coding Mode	Channel Format
1+1	L/Ch1, R/Ch2
1/0	C
2/0	L, R
3/0	L, C, R
2/1	L, R, S
3/1	L, C, R, S
2/2	L, R, LS, RS
3/2	L, C, R, LS, RS

4.5.3 LFE Channel

The LFE Channel parameter enables or disables the Low-Frequency Effects (LFE) channel. Use of the LFE channel is optional with multichannel programs, but is not available for mono, stereo, or surround-encoded programs.

Two-channel Dolby Digital products, or multichannel products operating in a two-channel downmix mode, omit the LFE signal. Therefore, low-frequency content essential to the program should never be mixed exclusively to the LFE channel. Refer to Section 3.5, LFE and Bass Management, for more information on decoder handling of LFE signals.

4.5.4 Bitstream Mode

The Bitstream Mode parameter indicates the type of audio service that the bitstream conveys. Complete Main (CM) is the normal mode of operation and contains a complete audio program including dialog, music, and effects. The CM and ME Main Services can be further enhanced by means of Associated Services. The Bitstream Modes and audio service types are listed in Table 4-6.

Table 4-6 Bitstream Mode/Audio Service Type

Bitstream Mode/Audio Service Type
Main Service: Complete Main (CM)
Main Service: Music and Effects (ME)
Associated Service: Visually-Impaired (VI)
Associated Service: Hearing-Impaired (HI)
Associated Service: Dialog (D)
Associated Service: Commentary (C)
Associated Service: Emergency (E)
Associated Service: Voice Over (VO) / Karaoke

4.5.5 Dialog Normalization (*dialnorm)*

The Dialog Normalization (*dialnorm*) value indicates how far the average dialog level of the encoded program is below digital 100% full scale (0 dBFS). Valid settings are -1 dB to -31 dB. This parameter determines the audio reproduction level and affects other parameters and decoder operation. Refer to Section 4.2, Features, for detailed information on this subject. A thorough definition of the *dialnorm* parameter can be found in ATSC document A/52, *Digital Audio Compression Standard (AC-3)* as well.

4.6 Bitstream Information

The parameters in this group directly relate to the Dolby Digital Bitstream Information (BSI) fields. Definitions for BSI parameters follow.

4.6.1 Center Downmix Level

The Center Downmix Level parameter indicates the nominal Lo/Ro downmix level of the Center channel with respect to the Left and Right channels. This parameter setting does not affect Lt/Rt downmixes. Table 4-7 lists the valid values for Center downmix level. This parameter *appears in the bitstream* only when three front channels are in use, i.e., only when the Audio Coding Mode is set to 3/0, 3/1, or 3/2.

Table 4-7 Center Downmix Level

Center Downmix Level	Mix Coefficient
-3.0 dB	0.707
-4.5 dB	0.596
-6.0 dB	0.500

4.6.2 Surround Downmix Level

The Surround Downmix Level parameter indicates the nominal Lo/Ro downmix level of the Surround channel(s) with respect to the Left and Right channels (consistent with the ITU BR specification). This parameter setting does not affect Lt/Rt downmixes. Table 4-8 lists the valid values for Surround downmix level.

This parameter *appears in the bitstream* only when a Surround channel is in use, i.e., only when the Audio Coding Mode is set to 2/1, 2/2, 3/1, or 3/2. It is recommended that the parameter be user-adjustable only when one of these modes has been selected.

Table 4-8 Surround Downmix Level

Surround Downmix Level	Mix Coefficient
-3.0 dB	0.707
-6.0 dB	0.500
-999 dB	0.000

4.6.3 Dolby Surround Mode

The Dolby Surround Mode parameter indicates whether or not a two-channel Dolby Digital bitstream is conveying a Dolby Surround encoded program. This information is not used by the Dolby Digital decoding algorithm, but can be used by other portions of the audio reproduction equipment, such as a Dolby Surround Pro Logic decoder. Table 4-9 lists the valid values for Dolby Surround Mode.

Table 4-9 Dolby Surround Mode Indications

Dolby Surround Mode Indications
Not indicated
Not Dolby Surround encoded
Dolby Surround encoded

In some cases, an operator finds this parameter user-adjustable even when an Audio Coding Mode other than 2/0 has been selected. The Dolby Surround Mode indicator, however, *appears in the bitstream* only when operating in the two-channel mode, i.e., only when the Audio Coding Mode is set to 2/0.

4.6.4 Language Code

The Language Code was intended to represent the language of the Dolby Digital audio service. Typically, there are MPEG system elements that are used for indicating the service language (language descriptors, for example). At this time there is no known use for this code and consequently there may be no reference to it on a Dolby Digital encoder user interface.

4.6.5 Audio Production Information Exists

The Audio Production Information Exists flag indicates whether the Mixing Level and Room Type parameters explained below exist within the Dolby Digital bitstream.

4.6.6 Mixing Level

The Mixing Level informational parameter indicates the absolute Sound Pressure Level (SPL) of the audio program as heard by the original mixing engineer. This information makes it possible to replay the program at exactly the same loudness, or at a known difference in loudness. By knowing how much lower a program is played at home, for example, it is now possible to apply the correct degree of *loudness compensation*.

The value for Mixing Level represents the theoretical loudness of a full-scale (0 dBFS) tone in one channel.

There are two kinds of encoders in use: The newer ones with the Windows interface provide a range of adjustment from 80 to 111, whereas the older ones range from 0 to 31. The actual data encoded into the bitstream is exactly the same in either case. The examples below are for the newer style encoders. If the older one is used, subtract 80 from the result.

Example A: Measure the SPL (C-weighted) of pink noise at -20 dBFS in the center channel. The reading is 85 dB. Set the encoder Mixing Level value to (85 + 20) = 105.

(For the older type of encoder, the setting will be 105 – 80 = 25)

Example B: Measure the SPL (C-weighted) of pink noise at -30 dBFS in the left channel. The reading is 72 dB. Set the encoder Mixing Level value to (72 + 30) = 102.

(For the older type of encoder, the setting will be 102 – 80 = 22)

A second value for this parameter type becomes active when the Audio Coding Mode is 1+1 and Audio Production Information Exists is set to 1, or yes. This second item applies to the second independent channel (Ch2) residing within the bitstream. This parameter appears in the bitstream only when the Audio Production Information Exists parameter is set to 1, or yes.

4.6.7 Room Type

The Room Type informational parameter indicates the type and calibration of the mixing room used for the final audio mixing session. The Room Type value is not normally used within the Dolby Digital decoder but can be used by other elements in the audio system. Table 4-10 lists the valid values for Room Type. This parameter appears in the bitstream only when the Audio Production Information Exists parameter is set to 1, or yes.

Table 4-10 Room Type

Room Type
Not indicated
Large room
Small room

4.6.8 Copyright Bit

The Copyright Bit informational parameter sets the value of a single bit within the Dolby Digital bitstream. If this bit has a value of 1, the information in the bitstream is indicated as protected by copyright. If it has a value of 0, the information is not copyright protected.

4.6.9 Original Bitstream

The Original Bitstream informational parameter sets the value of a single bit within the Dolby Digital bitstream. This bit has a value of 1 if the bitstream is an original. If it is a copy of an original bitstream, it has a value of 0.

4.7 Preprocessing Options

The Processing Options parameters listed in this group are used to precondition the audio input signals before they are encoded. Refer to Appendix D, Dolby Digital Time-Domain Filters, for further information.

4.7.1 Digital De-emphasis

Dolby Digital encoders can allow activation of digital de-emphasis applied to the linear PCM input signals whenever it is detected that the input has been pre-emphasized. Detection is typically achieved by monitoring the pre-emphasis flags within the channel status data of the incoming digital audio signal (e.g., AES/EBU or S/PDIF). Since the value of this parameter depends on some other parameter(s) or condition(s), it does not require explicit user control and can be adjusted automatically by the encoder. This parameter is not available on the Dolby Model DP561 encoder. When using the Dolby Remote software and the DP561 the Digital De-emphasis status will always be indicated as "Off."

4.7.2 DC Highpass Filter

This parameter can be used to activate the DC Highpass filter for all input channels. The DC Highpass filter should always be enabled unless the encoding engineer is absolutely sure that there is no DC in the input audio.

4.7.3 Channel Bandwidth Lowpass Filter

The Channel Bandwidth Lowpass Filter parameter can be used to activate a low-pass filter with a cut-off near the specified audio bandwidth that is applied to the main input channels. If the digital signal fed to the main input channels does not contain information above the specified audio bandwidth, this filter can be disabled. This parameter is not available on the Dolby Model DP561. It has also been disabled in the Dolby Remote software for the DP561.

4.7.4 LFE Lowpass Filter

The LFE Lowpass Filter parameter can be used to activate a 120 Hz low-pass filter applied to the LFE input channel. If the digital signal fed to the LFE input does not contain information above 120 Hz, this filter can be disabled. This parameter is user-adjustable only when the LFE channel is enabled.

4.7.5 Surround Channel 90-Degree Phase-Shift

The Surround Channel 90-Degree Phase-Shift feature is useful for generating multichannel Dolby Digital bitstreams that can be downmixed in an external two-channel decoder to create a true Dolby Surround compatible output. This parameter is user-adjustable only when Surround channels are present in the bitstream, i.e., only when Audio Coding Mode is set to 2/1, 2/2, 3/1, or 3/2.

The 90-Degree Phase-Shift parameter should always be left enabled except under specific conditions. These include, but are not necessarily limited to, system calibration, encoding of certain test signals, and in the *extremely rare* case when the discrete playback of highly coherent program material may be compromised.

4.7.6 Surround Channel 3 dB Attenuation

The Surround Channel 3 dB Attenuation function is useful for applying a 3 dB attenuation to the Surround channels of a multichannel soundtrack created in a room with film style calibration, when encoding it for consumer home theater playback. Cinema soundtrack Surround channels are mixed +3 dB relative to the front channels in order to account for cinema calibration standards. Home theater Surround channel gains are calibrated differently, and so a -3 dB adjustment to the Surround tracks is necessary. This parameter is user-adjustable only when Surround channels are present in the bitstream, i.e., only when Audio Coding Mode is set to 2/1, 2/2, 3/1, or 3/2.

4.7.7 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. Depending on encoder implementation, Dynamic Range Compression can be set by merely enabling the parameter or by selecting one of several built-in profiles. These profiles include Film Standard, Film Light, Music Standard, Music Light, and Speech. Refer to Section 4.2 Features, and Appendix C, Dynamic Range Control (DRC), for detailed information on this subject. A thorough definition of Dynamic Range Control can be found in ATSC A/52, *Digital Audio Compression Standard (AC-3)*.

4.7.8 RF Overmodulation Protection (RF Pre-emphasis Filter)

The RF Overmodulation Protection parameter determines whether or not an RF pre-emphasis filter is used in the overload protection algorithm to prevent RF overmodulation in set-top box decoders. It is primarily used for broadcast applications. This parameter is not available on the Dolby Model DP561 Dolby Digital encoder. It has been disabled in the Remote Control software for the DP561.

4.8 Automatic Parameters

The values for the Automatic Parameters are adjusted automatically by the Dolby Digital professional encoder.

4.8.1 Audio Bandwidth

In general, audio bandwidth is adjusted automatically within the encoder. Under most circumstances, the encoder maintains the full audio spectrum bandwidth. As the data rate is decreased below a certain value, however, it becomes useful to decrease the audio bandwidth in order to maintain high audio quality. The optimum choice for a bandwidth value depends on the selected Data Rate and the Audio Coding Mode.

4.8.2 Coupling

In general, coupling is adjusted automatically within the encoder. It is used in the cases where more than one audio channel is being encoded in order to increase coding efficiency by sharing information across the channels. Dolby Digital encoders employ coupling primarily at lower data rates and only when the audio signals meet the proper criteria.

4.9 Input/Output Control

The parameters in the Input/Output Control group relate to physical and electrical characteristics of the input and output connections for a Dolby Digital professional encoder. Controls of this type are by nature dependent on the particular encoder application because there are many possible ways to connect audio signals to and from an encoder product. The controls listed here are suggestions; there may be other input and output controls needed for a given Dolby Digital encoder.

4.9.1 Sampling Rate

Dolby Digital supports three standard sampling rates: 48 kHz, 44.1 kHz, and 32 kHz. The linear PCM audio signals that are input to the Dolby Digital encoder must be sampled at one of these rates. Since the value of this parameter depends on other parameters or conditions, it does not require explicit user control and can be adjusted automatically by the encoder. It is important that the sampling rate conform to that of the final release medium, which is typically 48 kHz.

4.9.2 Input Channels

Dolby Digital professional encoders typically implement a means to configure the mapping of the input signals to the proper Dolby Digital encoded channel assignment. The number of input channels that are actually encoded depends upon the Audio Coding Mode and the LFE Channel setting.

4.9.3 Output Format

Some Dolby Digital encoders may include an Output Format control. These parameters determine the output bitstream configuration and include format, audio bit, and SMPTE time code selections. Refer to Appendix B, Bitstream Format, for more information.

4.10 Processing State Control

The parameters in the Processing State Control group influence the processing state of the Dolby Digital encoder. The simplest example is a control for starting and stopping the encoder. Some of the controls explained here are only relevant for specific encoder applications. One such example is a user control for recording the Dolby Digital bitstream to a disk file. This function can be useful for an application such as DVD authoring, where the encoded data is to reside as a file on a computer disk, but would not be useful in broadcast encoder applications.

4.10.1 Start/Stop Encoding

Many encoders have a control for starting and stopping the encoding process.

4.10.2 Configuration Presets

Some encoders may offer configuration presets (not to be confused with Dynamic Range Compression profiles that are also sometimes referred to as presets) as a way to save and recall parameter settings for encoder applications. Due to the large number of possible parameter combinations, it can be useful to store a particular combination that can be recalled later, thus saving setup time and reducing the possibility of an erroneous setting.

4.10.3 Time Code Control

Some applications may require the use of SMPTE time code for controlling certain aspects of the Dolby Digital encoder operation. An example would be DVD authoring, where it is necessary to resolve synchronization between the separate audio and video elements before they are multiplexed into a single system bitstream.

Note that this example does not require the SMPTE time code values to be transmitted as part of the elementary bitstream. Typically, there are MPEG system elements that are used for audio/video synchronization at decode time (e.g., presentation time stamp, PTS). At this time there is no known use for the syntactic time code elements, which are optional in the BSI portion of the elementary Dolby Digital bitstream. Instead, the time stamp values that are associated with each Dolby Digital synchronization frame can be outside of the elementary bitstream.

4.10.4 Record/Play Bitstream

Some applications may require the capability of recording the Dolby Digital bitstream as a computer disk file. In these applications controls for destination file name, mode selection and record/playback are useful.

4.11 Using the Dolby Model DP562 Professional Reference Decoder

It is important to use a professional reference decoder to monitor content for Dolby Digital encoding. Dolby Digital offers many features to maintain backward compatibility as well as to allow consumers the ability to customize their listening environment. Features such as downmixing and Dynamic Range Control need to be checked at various stages of content creation and delivery to ensure they meet the intent of the content provider as well as the needs of the consumer. The Dolby Model DP562 professional reference decoder provides monitoring capabilities for these parameters in addition to being able to simulate almost any listening environment.

4.11.1 Downmixing

Downmixing has two, frequently interrelated applications: format compatibility and channel redirection.

1. Format Compatibility

Dolby Surround-compatible, stereo, and mono mixes are often created when multichannel material is downmixed to fewer channels. It is important to check a number of aspects of the downmix to confirm that it translates as closely as possible to the original intent of the mix.

Many consumers listen to Dolby Digital sources such as DVD or DTV without a full 5.1-channel Dolby Digital playback system. These consumers hear the two-channel analog or linear PCM outputs of their DVD players or DTV set-top boxes through stereo or Dolby Surround Pro Logic systems. All DVD video players and DTV set-top boxes have the ability to create and deliver a Dolby Surround compatible, stereo, or mono downmix from the two-channel analog or linear PCM outputs. The DP562

Professional Reference Decoder can simulate what the consumer hears when listening in these modes.

Example 1: Using a properly calibrated 5.1-channel monitoring system (incorporating appropriate bass management), set the DP562 to “Dolby Digital” and “Full.” In this configuration, a 5.1-channel bitstream reproduces all channels as a consumer with a Dolby Digital 5.1-channel system hears them. Pressing “Pro Logic” on the DP562 downmixes the five main channels (discarding the LFE channel) to a Dolby Surround-compatible bitstream. The downmix is then Dolby Surround Pro Logic decoded resulting in Left, Center, Right, and mono Surround channels at the outputs. Monitoring in this mode simulates how a consumer hears the 5.1-channel bitstream when downmixed and then reproduced through a Dolby Surround Pro Logic system.

Example 2: With “Pro Logic” still engaged, select “Stereo” instead of “Full” in the “Listening Mode” section. This mode allows monitoring the 5.1-channel bitstream as a consumer hears it when downmixed and then reproduced through a two-channel stereo system. Refer to Figure 4-1, Dolby Surround Compatible Lt/Rt Downmix.

Example 3: In addition to Dolby Surround (Lt/Rt) compatible downmixes, Lo/Ro downmixes can be checked. Selecting “Stereo” mode without “Pro Logic” engaged creates an Lo/Ro downmix at the outputs. Refer to Figure 4-2, Stereo Compatible Lo/Ro Downmix.

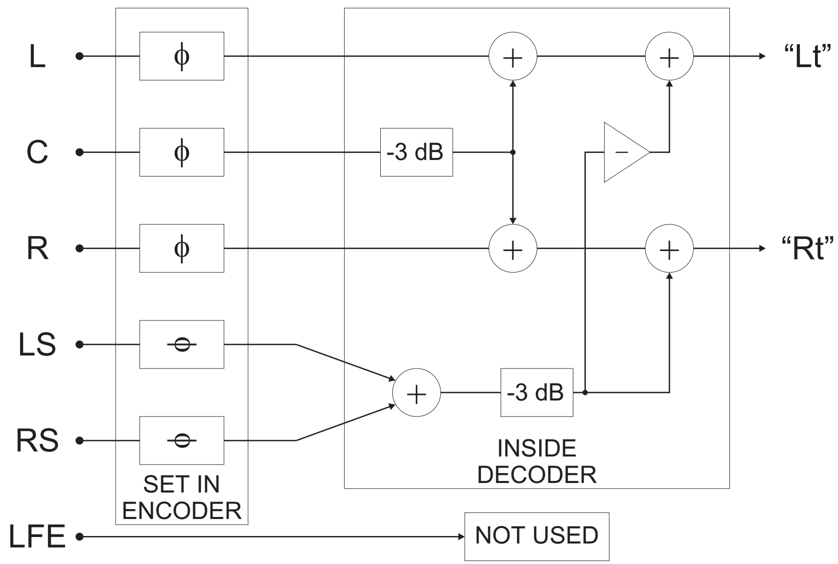


Figure 4-1 Dolby Surround Compatible Lt/Rt Downmix

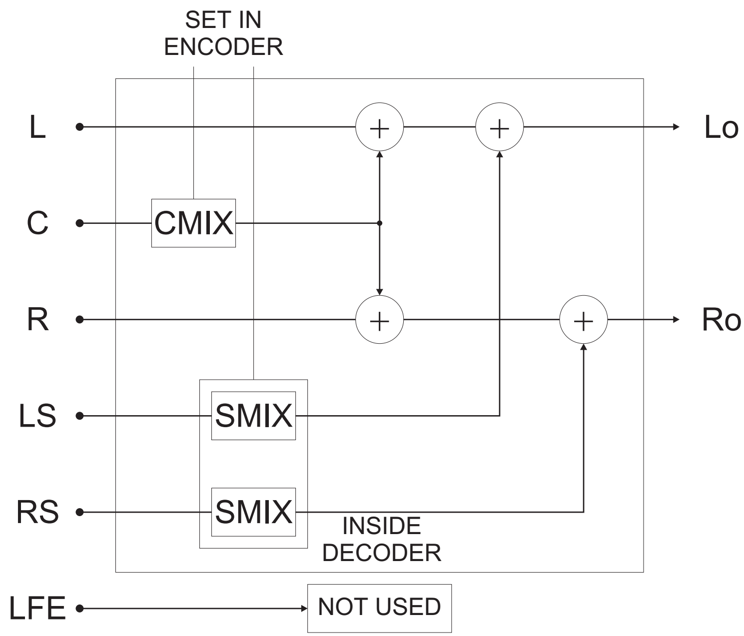


Figure 4-2 Stereo Compatible Lo/Ro Downmix

2. Channel Redirection

The ability to redirect channel information provides a means to account for the design of and number of speakers in the listening environment.

Some consumers may not have or cannot use all 5.1 speakers with their Dolby Digital decoder. Dolby Digital consumer decoders can redirect or downmix decoded multichannel information such as a 5.1-channel audio program.

Example 1: Using a properly calibrated 5.1-channel monitoring system (incorporating appropriate bass management), set the DP562 to “Dolby Digital” and “Full.” In this configuration, a 5.1-channel bitstream reproduces all channels as a consumer with a Dolby Digital 5.1-channel system hears them. Pressing any of the other “Listening Modes” causes the DP562 to redirect audio to the outputs of the selected speaker configuration.

Example 2: Select “3 Stereo” instead of “Full” in the “Listening Mode” section and the Surround channel information is redirected to the Left and Right speakers to simulate a monitoring system with no Surround speakers. Select “Phantom Center” and the Center channel information is redirected to the Left and Right speakers to simulate a monitoring system with no Center speaker.

Technical Note: Summing multiple channels of audio, as occurs in downmixing, has the potential for overloading the channel outputs. The DP562 applies the necessary level scaling to prevent overload in the DSP processing. This results in an 11dB attenuation of the AES/EBU audio outputs in “Custom” and “None” Dynamic Compression Modes (also referred to as Operational Modes) when downmixing. The attenuation is recovered in the analog outputs.

4.11.2 Dynamic Range Control (DRC)

Dynamic Range Control (DRC) incorporates both selectable Dynamic Range Compression profiles and automatic overload protection limiting. Many consumer products allow the user to select reduced dynamic range when listening to a Dolby Digital multichannel audio program. During downmixing, protection limiting can be automatically applied to prevent overload. The DP562 has the capability to monitor the DRC encoded into the Dolby Digital bitstream.

Example: Selecting "Line" from the "Dynamic Compression" section applies DRC either fully or partially as determined by the configuration. Selecting "RF" implements both Dynamic Range Control and Dialog Normalization as would be applied for outputs that are connected to the RF (antenna) input of a television set. "Custom" offers the same options for monitoring DRC as "Line" along with the ability to defeat Dialog Normalization. "None" is strictly a professional mode that defeats both DRC and Dialog Normalization. "Line" and "RF" always have Dialog Normalization applied.

4.11.3 Bass Management

Bass management allows the user to redirect low-frequency information from any of the five main speakers to the subwoofer or conversely, if there is no subwoofer the LFE information can be redirected to the left and right speakers. This is important as the vast majority of consumer home theater speaker systems require some degree of bass management since typically none of the five main speakers are designed to reproduce frequencies below 80 Hz. The DP562 provides the same bass management functions as a consumer Dolby Digital decoder. Even when monitoring with full-range main speakers that require no bass management, this function is useful for checking how low frequencies redirected from any of the main channels may interact with the LFE channel information. Proper bass management is necessary to emulate a consumer home theater system, as most consumers use some form of it. Two bass management configurations are required for consumer products in the Decoder category (refer to Section 3.5, LFE and Bass Management).

4.11.4 LFE Monitor Mode

The LFE Monitor Mode in the setup menu provides a means to monitor the use of the LFE channel in a Dolby Digital downmix. The DP562 defaults to the “AUTO” mode, which does not allow the LFE channel to be included in the downmix where it is not appropriate.

Example: Whenever Dolby Surround compatible downmixing occurs, the LFE channel is not included in the downmix. With the LFE Monitor Mode set to “AUTO,” the DP562 automatically mutes the LFE channel whenever “Pro Logic” is selected. This is also true for a “Stereo” (Lo/Ro) or “Mono” Dolby Digital downmix without “Pro Logic.” Conversely, in “AUTO” mode, the DP562 does not mute the LFE channel when in “Dolby Digital” and either “Full,” “Phantom Center,” or “3 Stereo.” Selecting LFE Monitor Mode “ON” allows the DP562 to pass LFE channel information regardless of the downmix mode. LFE Monitor Mode “OFF” is essentially a mute switch. “AUTO” mode should always be used unless there is a specific requirement unrelated to checking downmixes.

Chapter 5
Applications and Formats

5.1 DVD-Video

Dolby Digital is one of the standard audio formats for DVD-Video in all NTSC and PAL countries. Refer to Section 5.1.1, DVD-Video Specification, for more information. It is crucial for all engineers and technicians who author DVDs to be familiar with the basic Dolby Digital audio requirements for DVD-Video before starting the audio encoding process. This section covers the details of audio encoding for DVD-Video. Any questions regarding DVD authoring or video encoding should be addressed to the appropriate companies.

DVD players come in a few form factors: stand-alone home and portable units, incorporated into personal computers, and convergence products. All these players are capable of playing the same DVD-Video discs on the market today. In NTSC and PAL countries all DVD players are equipped with Dolby Digital decoders that are capable of playing DVD-Video discs with Dolby Digital tracks. The consumer usually has the choice of listening to the DVD player decoded line outputs or connecting the Dolby Digital data line to an external decoder (A/V receiver, etc.). This gives the consumer more flexibility in the decoding process.

The following guidelines apply when encoding audio material for the DVD-Video format. Contact manufacturers of DVD authoring software and video encoding equipment for more information in these areas.

5.1.1 DVD-Video Specification

The DVD Forum has issued a specification outlining the accepted audio formats for DVD-Video in PAL countries. Any PAL DVD-Video disc can now have Dolby Digital audio tracks *only*—without the need for linear PCM or MPEG audio to be present anywhere on the disc. In NTSC countries, a DVD-Video disc conforms to the specification if it carries either a Dolby Digital audio track or a linear PCM audio track.

5.1.2 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.

In some cases, video encoder inefficiencies may require a higher video data rate than normal. In these cases the audio data rate for multichannel content could be lowered to 384 kb/s. This lower data rate should only be used when it is clear that the video encoder requires a higher data rate.

5.1.3 Bit Resolution

The Dolby Digital algorithm is capable of encoding up to 24-bit audio. In addition, all DVD authoring systems can accept Dolby Digital (.ac3) files that were created from 24-bit audio sources. Dolby Digital decoders offer bit resolutions from 16 to 24 bits.

5.1.4 Audio/Video Synchronization

One of the issues facing DVD-Video authoring engineers is audio/video synchronization. An important step to follow in the encoding stages of DVD-Video authoring is to include SMPTE time code in both the video file and the Dolby Digital .ac3 file. Most DVD authoring systems today accommodate audio/video synchronization using the imbedded SMPTE time code in both the .ac3 file and in the MPEG-2 video file. Although using SMPTE time code greatly reduces synchronization issues, fine-tuning may still be required before the final production disc shows perfect synchronization. Experience has shown that verifying correct synchronization prior to encoding can save time in the DVD-Video production process.

5.1.5 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. Refer to Chapter 4, Encoding, for more information. 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.
- Save the SMPTE time code, if present in the source material, with the Dolby Digital output: It is crucial for synchronization with video.
- 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.

5.1.6 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.

5.1.7 Karaoke DVD

The key issues when creating a karaoke DVD-Video are the audio channel configuration, the correct setting of the Dolby Digital bitstream, and the actual authoring technique. Each karaoke audio element must be assigned to the correct channel of the Dolby Digital encoder. The karaoke format is designed for two-channel reproduction, therefore route the stereo music mix for a karaoke DVD-Video to the Left and Right channels of the encoder. If a guide melody to assist the singer is available, assign it to the Center channel. Allocate main and/or backing vocals to the Left and Right Surround channels. If there is only one backing vocal, use the Left Surround channel.

It is important to set the Dolby Digital encoder Bitstream Mode to Karaoke. When authoring, the engineer must be familiar with the karaoke capability of the software tools since various settings determine how the disc behaves in the DVD-Video player.

Although a karaoke DVD-Video will play back on standard DVD-Video players, a karaoke-capable player is required for the consumer to use all the features of this mode. The karaoke capable player allows the user to mix his voice with the main music as well as the guide and backing vocals, if present. Depending on the player model, other features can include pan controls and “echo” or reverb units. For those without karaoke capable players, a second audio track on the disc with either a suitable two-channel or even five-channel mix can be useful. The Bitstream Mode for the second audio track should be set to Complete Main.

5.1.8 Miscellaneous Issues

There are areas of DVD authoring that could be covered in this manual but are best answered by the DVD authoring companies. Issues such as seamless audio branching and multiple language files are handled differently on various authoring systems. Contact the respective DVD authoring company for more information.

5.2 DVD-Audio

Dolby Digital can be used as an alternative to linear PCM in the video zone of a DVD-Audio disc. Dolby Laboratories recommends that all DVD-Audio discs include a Dolby Digital bitstream with program content identical to the multichannel linear PCM tracks. This ensures playback compatibility with all DVD-Video players.

5.3 DVD-ROM

Traditionally, when preparing material for multimedia, heavy peak limiting has been used to maximize the volume of each sound element. When working with both .ac3 files and .wav files, examine the .wav files to determine the relative volume as compared to the .ac3 files. If adjustment is required, either increase the level of the .ac3 files to near 0 dBFS and encode them with a *dialnorm* setting of -31 dB, or attenuate the level of the existing .wav files.

Be advised that both program levels and Dolby Digital encoder settings are factors in determining the need for peak overload protection. Dolby Digital encoders automatically generate overload protection words (calculated by the worst-case downmix) when the selected Dynamic Range Compression profile (also referred to as a preset) is inadequate in preventing digital overload. This protection is applied to the audio in the form of peak limiting when either downmixing or dynamic range compression is active. Material with signals near 0 dBFS in multiple channels and a *dialnorm* setting of -31 dB induces more overload protection than that of similar material with a single channel near 0 dBFS and a higher *dialnorm* value. Dolby Laboratories recommends monitoring Dolby Digital encoding using various decoding modes during production to ensure that the audio will be reproduced as intended.

Currently, .ac3 files as elementary bitstreams can be played only through DirectShow compatible decoders. Only a few DVD hardware decoder boards recognize the .ac3 file format when using MCI driver commands for a DVD-ROM title.

Many software decoders, however, now have the ability through DirectShow to access the decoded linear PCM channels out of the Dolby Digital decoder and mix them to L, R, LS, and RS and output them through a standard four-channel audio card. This capability can be used to replace Redbook audio with 5.1-channel Dolby Digital music or background tracks and retain the ability to use run-time sound effects concurrently. Contact multimedia@dolby.com for more information.

5.4 Digital Television (DTV)

Dolby Digital is widely used in digital television (DTV) broadcasting systems. Dolby Digital was first standardized by the Advanced Television Systems Committee (ATSC) in document A/52, *Digital Audio Compression Standard (AC-3)*. The ATSC DTV Standard, approved by the FCC for use in the United States, is ATSC document A/53, *ATSC Digital Television Standard*. ATSC document A/54*, Guide to the Use of the ATSC Digital Television Standard* is a useful tutorial. All of the ATSC documents are available from the ATSC web page at www.atsc.org.

5.4.1 ATSC DTV Constraints

Certain Dolby Digital parameters are constrained in the ATSC DTV application. The most significant constraints are:

- The sample rate is fixed at 48 kHz (which must be locked to the picture rate)
- The data rate for Main or Associated Services containing all necessary program elements currently cannot exceed 384 kb/s.
- The 1+1 (or dual mono) mode is not allowed for emission. This mode is only for use in professional production or distribution links where it is necessary to place two completely independent programs into a single audio elementary bitstream.

Other constraints apply to the maximum data rates for Associated Services (see below). The constraints are summarized in Table 5-1, excerpted from ATSC document A/53, *ATSC Digital Television Standard, Annex B*.

Table 5-1 ATSC DTV Audio Constraints

Dolby Digital Syntactical Element	Comment	Allowed Value
fscod	Indicates sampling rate	‘00’ (indicates 48 kHz)
frmsizecod	Main or Associated Service containing all necessary program elements	≤ ‘011100’ (indicates ≤ 384 kb/s)
frmsizecod	Single-channel Associated Service containing a single program element	≤ ‘010000’ (indicates ≤ 128 kb/s)
frmsizecod	Two-channel Dialog Associated Service	≤ ‘010100’ (indicates ≤ 192 kb/s)
frmsizecod	Combined data rate of a Main and an Associated Service intended to be simultaneously decoded	(total ≤ 512 kb/s)
acmod	Indicates number of channels	≥ ‘001’

5.4.2 Implementation

Broadcasting typically employs hardware Dolby Digital professional encoders working in real-time at the point of emission to the consumer. In two-channel stereo, the encoder may be a Dolby Laboratories product such as the Model DP567, or may be an encoder that is embedded inside the video encoding system. In 5.1-channel broadcasting, the encoder is a Dolby Laboratories product such as the Model DP569. Dolby encoding products such as the DP567 and DP569, and some encoders built by Dolby Laboratories licensees can pass through encoded bitstreams. Encoders that encode linear PCM into Dolby Digital elementary bitstreams, when presented with an encoded Dolby Digital bitstream at their input simply pass through the Dolby Digital bitstream. In some broadcast systems it is thus possible for pre-encoded Dolby Digital bitstreams to be transmitted.

5.4.3 Main, Associated, and Multilingual Services

The following information is excerpted from ATSC document A/54*, Guide to the Use of the ATSC Digital Television Standard*.

A Dolby Digital elementary bitstream contains the encoded representation of a single audio service. Multiple elementary bitstreams provide multiple audio services. Each elementary bitstream is conveyed by the MPEG-2 transport multiplex with a unique *PID*. There are a number of audio service types that can (individually) be coded into each elementary bitstream. Each elementary bitstream is tagged as to its service type using the *bsmod* bit field. There are two types of Main Services and six types of Associated Services. Each Associated Service can be tagged (in the AC-3 audio descriptor in the transport PSI data) as being associated with one or more main audio services.

Associated Services can contain complete program mixes, or can contain a single program element. Associated Services that are complete mixes can be decoded and used without additional services. They are identified by the *full svc* bit in the AC-3 descriptor. Refer to ATSC document A/52, *Digital Audio Compression Standard (AC-3), Annex A*. Typically, Associated Services that contain a single program element are combined with the program elements from a Main Service.

In general, a complete audio program is presented to the listener over a set of loudspeakers. This may consist of a Main Service, an Associated Service that is a complete mix, or a Main Service combined with one Associated Service. The capability to simultaneously decode one Main Service and one Associated Service is required in order to form a complete audio program in certain service combinations described in this section. This capability may not exist in some receivers.

Summary of Audio Service Types

The audio service types that correspond to each value of *bsmod* are defined in the ATSC documents A/52, *Digital Audio Compression Standard (AC-3)* and A/53, *ATSC Digital Television Standard, Annex B*. The information is reproduced in Table 5-2. The paragraphs that follow briefly describe the service types.

Table 5-2 Audio Services

bsmod	Audio Service Types
000 (0)	Main Service: Complete Main (CM)
001 (1)	Main Service: Music and Effects (ME)
010 (2)	Associated Service: Visually-Impaired (VI)
011 (3)	Associated Service: Hearing-Impaired (HI)
100 (4)	Associated Service: Dialog (D)
101 (5)	Associated Service: Commentary (C)
110 (6)	Associated Service: Emergency (E)
111 (7)	Associated Service: Voice-Over (VO)

Multilingual Services

Each audio bitstream can be in any language. To provide audio services in multiple languages, individual Main Services can be created for each language. This is the artistically preferred method because it allows unrestricted placement of dialog along with the dialog reverberation. The disadvantage of this method is that each language requires as much as the maximum data rate for a full 5.1-channel service (currently 384 kb/s). One way to reduce the data rate is to restrict the number of audio channels for languages with a limited audience. For instance, alternate language versions in two-channel stereo could be provided at a data rate of 128–192 kb/s. A mono version could be supplied at a data rate of 64–96 kb/s.

Another way to offer multiple-language service is to provide Music and Effects service (ME), which does not contain dialog. Multiple single-channel Dialog services (D) can then be provided, each at a data rate of 64–96 kb/s. Formation of a complete audio program requires that the appropriate language D service be simultaneously decoded and mixed into the ME service. This method allows a large number of languages to be efficiently provided, but with artistic limitations. The single channel of dialog would be mixed into the Center reproduction channel, and could not be panned. Also, reverberation would be confined to the Center channel, which is not optimal. This method results in a substantial data rate savings, making it ideal for some types of programming. Some receivers may not have the capability to simultaneously decode a ME and a D service.

When transmitting a two-channel stereo ME service along with two-channel stereo D services, multiple languages are delivered efficiently without compromising artistic content. The D service and appropriate language ME service are combined in the receiver into a complete two-channel stereo program. Dialog can be panned, and reverberation can be included in both channels. A two-channel stereo ME service can be sent with high quality at 192 kb/s, while the two-channel stereo D services (voice only) can make use of lower data rates, such as 128 or 96 kb/s per language. Some receivers may not have the capability to simultaneously decode a ME and a D service.

During those times when dialog is not present, the D services can be momentarily removed, and their data capacity used for other purposes.

5.4.4 Detailed Description of Service Types

Complete Main (CM)

The Complete Main service (CM) is the normal mode of operation. It contains a complete audio program, with dialog, music, and effects. This is the type of audio service typically provided. The CM service can contain from one to 5.1 audio channels. It can be further enhanced with the VI, HI, C, E, or VO services described below. To provide audio service in multiple languages, individual CM services can be created for each language.

Music and Effects (ME)

The Music and Effects (ME) type of Main Service contains the music and effects for an audio program, but not the dialog. The ME service can contain from one to 5.1 audio channels. The primary program dialog is missing and, if any exists, is supplied by providing a D service. Multiple D services in different languages can be associated with a single ME service.

Visually-Impaired (VI)

The Visually-Impaired type of Associated Service typically contains a narrative description of the visual program content. In this case, the VI service is a single audio channel. Simultaneous reproduction of the VI service and the Main Service allows the visually-impaired user to enjoy the main multichannel audio program, as well as to follow the on-screen activity. The VI service can be mixed into one of the main reproduction channels; the choice of channel can be left to the listener or be provided as a separate output. The separate output might then be delivered to the VI user via open-air headphones or other means.

The Dynamic Range Control (DRC) data in this type of VI service is intended for use by the audio decoder to modify the level of the main audio program. Thus the level of the Main Service is under the control of the VI service provider. The provider can signal the decoder by altering the DRC words embedded in the VI audio elementary bitstream to reduce the level of the main audio service by up to 24 dB assuring that the narrative description is intelligible.

Besides being provided as a single narrative channel, the VI service can be provided as a complete program mix containing music, effects, dialog, and narration. In this case, the service can be coded using any number of channels, up to 5.1, and the Dynamic Range Control word applies only to this service. The fact that the service is a complete mix is indicated in the Dolby Digital descriptor. Refer to *ATSC A/52, Digital Audio Compression Standard (AC-3), Annex A*, for more information.

Hearing-Impaired (HI)

The Hearing-Impaired type of Associated Service typically contains only a single channel of dialog and is intended for use by those whose hearing impairments make it difficult to understand the dialog in the presence of music and sound effects. The dialog can be processed for increased intelligibility by the hearing impaired. The hearing-impaired listener may wish to listen to a mixture of the single-channel HI dialog track and the main program audio. Simultaneous reproduction of the HI service along with the CM service allows the HI listener to adjust the mixture to control the emphasis on dialog over music and effects. The HI channel is typically mixed into the Center channel. An alternative is to deliver the HI signal to a discrete output that could be fed to a set of open-air headphones or other device, worn only by the HI listener.

Besides being provided as a single narrative channel, the HI service can be provided as a complete program mix containing music, effects, and dialog with enhanced intelligibility. In this case, the service can be coded using any number of channels, up to 5.1. The fact that the service is a complete mix is indicated in the AC-3 descriptor. Refer to ATSC A/52, *Digital Audio Compression Standard (AC-3), Annex A, for more information*.

Dialog (D)

The Dialog type of Associated Service is employed to most efficiently offer multichannel audio in several languages simultaneously when the program material is such that the restrictions (no panning, no multichannel reverberation) of a single dialog channel can be tolerated. When the D service is used, the Main Service is type ME. If the D service contains a single channel, simultaneously decoding the ME service allows a complete audio program to be formed by mixing the D channel into the Center channel. Typically, when the Main audio Service is of type ME, there are several different language D services available. The transport demultiplexer can be designed to select the appropriate D service to deliver to the audio decoder based on the listener's language preference as defined by data stored in the receiver memory.

Or, the listener can override the default selection by instructing the receiver to select a particular language D service.

If the ME service contains more than two audio channels, the D service is monophonic (1/0 mode). If the ME service contains two channels, the D service can contain two channels (2/0 mode). In this case, a complete audio program is formed by simultaneously decoding the D service and the ME service. The Left channel of the ME service is mixed with the Left channel of the D service, and the Right channel of the ME service is mixed with the Right channel of the D service. The result is a two-channel stereo signal containing music, effects, and dialog.

Commentary (C)

The Commentary type of Associated Service is similar to the D service, except that instead of conveying primary program dialog, the C service conveys optional program commentary. When C service(s) are provided, the receiver can notify the listener of their presence. The listener should be able to inquire about the various available C services, and select one for decoding along with the Main Service. The C service can be added to any loudspeaker channel under listener control. Typical uses for the C service are optional added commentary during a sporting event, or different levels (novice, intermediate, and advanced) of commentary available to accompany documentary or educational programming.

The C service can be a single audio channel containing only the commentary content. In this case, simultaneous reproduction of a C service and a CM service allows the listener to hear the added program commentary.

The Dynamic Range Control data in the single-channel C service is intended for use by the audio decoder to modify the level of the main audio program. Thus, the level of the Main Service is controlled by the C service provider. The provider can signal the decoder, by altering the Dynamic Range Control words embedded in the C service audio elementary bitstream, to reduce the level of the Main Service by up to 24 dB in order to ensure intelligible commentary.

Besides providing the C service as a single commentary channel, the C service can be provided as a complete program mix containing music, effects, dialog, and the commentary. In this case the service can be provided using any number of channels (up to 5.1). The fact that the service is a complete mix is indicated in the Dolby Digital descriptor. Refer to ATSC A/52, *Digital Audio Compression Standard (AC-3), Annex A, for more information*.

Emergency (E)

The Emergency type of Associated Service is intended to allow the insertion of emergency announcements. The normal audio services do not necessarily have to be replaced to present the emergency message. The transport demultiplexer gives priority to this type of audio service. Whenever an E service is present, it is delivered to the audio decoder by the transport subsystem. When the audio decoder receives an E-type Associated Service, it stops reproducing any Main Service being received and only reproduces the E service. The E service may also be used for non-emergency applications. It may be used whenever the broadcaster wishes to force all decoders to quit reproducing the main audio program and substitute a higher priority single channel.

Voice-Over (VO)

It is possible to use the E service for announcements, but the use of the E service leads to a complete substitution of the voice-over for the main program audio. The Voice-Over type of Associated Service is similar to the E service, except that it is intended for reproduction along with the Main Service. The systems demultiplexer gives priority to this type of associated service, second only to an E service. The VO service is intended to be simultaneously decoded and mixed into the Center channel of the main audio service that is being decoded.

The Dynamic Range Control data in the VO service is intended for use by the audio decoder to modify the level of the main audio program. Thus the level of the Main Service is under the control of the broadcaster. The broadcaster may signal the decoder by altering the Dynamic Range Control words embedded in the VO audio bitstream, to reduce the level of the Main Service by up to 24 dB during the voice-over. The VO service allows typical voice-overs to be added to an already encoded audio bitstream, without requiring the audio to be decoded back to baseband and then re-encoded. Space however, must be available within the transport multiplex for the insertion of the VO service.

5.4.5 Splicing Bitstreams

In some broadcast applications it is likely that encoded bitstreams will be spliced. The ideal place to splice encoded audio bitstreams is at the boundary of a sync frame. If a bitstream splice is performed at the sync frame boundary, the audio decoding proceeds without interruption. If a bitstream splice is performed randomly, an audio interruption results. The frame that is incomplete does not pass the error detection test within the decoder and causes it to mute. The decoder does not find sync in its proper place in the next frame, and enters a sync search mode. Once the sync code of the new bitstream is found, synchronization is achieved, and audio reproduction can begin once again. The outage may be on the order of two frames, about 64 milliseconds at the 48 kHz sample rate. The actual outage depends on the specific audio decoder implementation, and the implementation of the demultiplexing system that precedes the audio decoder. In some implementations the actual mute could be longer than 64 milliseconds. When the audio goes to mute, there may be a gentle fade down over a period of 2.6 milliseconds due to the windowing process of the filter bank. When the audio is recovered, it may fade up over a period of 2.6 milliseconds. Except for the approximately 64 milliseconds (or longer) of time that the audio is muted, the effect of a random splice of a Dolby Digital elementary bitstream can be relatively benign.

SMPTE is developing a standard for splicing MPEG-2 transport bitstreams. Splices performed according to this standard will occur on Dolby Digital frame boundaries and on picture frame boundaries. Since audio frame boundaries and picture frame boundaries are not synchronous, however, splicing inevitably leaves a gap in the audio, and causes an interruption in the audio frame sequence. The first frame after the splice has an MPEG-2 presentation time stamp (PTS) value that is greater than 32 milliseconds relative to the frame prior to the splice. The behavior of equipment in this situation is not well defined. Ideally, a receiver would fade down the old audio, immediately resynchronize to the new framing sequence, and fade up the new audio. In the worst case a receiver may go into a frame repeat type of error concealment, followed by an extended mute, before recovering and beginning to reproduce the new audio.

Given the uncertain behavior of receivers, it is good practice to provide a moment of silence at anticipated splice points, i.e., at the end of any program segment.

5.5 Laser Disc

5.5.1 Track Layout

When preparing the D2 master for the pressing plant, the audio and video tracks should be conformed onto the D2 master as they will appear on the laser disc itself. The D2 masters provided for conforming audio tracks should have the finished video as it will appear on the laser disc with a separate D2 for each side of the disc. Depending on the laser disc format, this is a maximum of one hour per side for CLV discs and one half hour per side for CAV discs. These tapes are typically delivered to the audio facility with the two-channel Lt/Rt or stereo PCM tracks already on channels 3 and 4 of the D2 tape and a mono composite or commentary track on channel 1. Channel 2 is reserved for the Dolby Digital bitstream. The stereo or mono tracks should already be on the D2 tape before the Dolby Digital bitstream is recorded. Refer to Table 5-3 for the typical audio track layout for the D2.

Table 5-3 Typical D2 Track Layout

Track Layout	D2	Laser disc
Mono or Commentary (+12 dB analog)	PCM Channel 1	Left Analog Channel
Dolby Digital Bitstream	PCM Channel 2	Right Analog Channel
PCM Lt or Left (+20 dB analog)	PCM Channel 3	PCM Left (+20 dB analog)
PCM Rt or Right (+20 dB analog)	PCM Channel 4	PCM Right (+20 dB analog)

5.5.2 Audio/Video Synchronization

When a Dolby Digital bitstream is encoded onto the laser disc, the bitstream is given a six-frame advance on the disc itself. The advance accounts for the access time required to retrieve the data from the disc in the laser disc player. The common method for preparing the master is to record the Dolby Digital bitstream in synchronization with the picture and have the six-frame advance done at the laser disc pressing plant. This method allows the audio editor to hear the Dolby Digital bitstream in synchronization with picture as the disc is checked for quality control (QC). It is acceptable to have the six-frame advance performed at the time of encoding onto the D2 tape. If this is done, care must be taken to properly label the tape: “The Dolby Digital bitstream on this tape has a six-frame advance.” Digital delays can then be added to the decoded Dolby Digital audio channels to place the audio back in sync with the picture for QC purposes. The approximate delay time for six frames at a frame rate of 29.97 is 200 milliseconds.

Figure 5-1 depicts a typical setup for performing a Dolby Digital laser disc encoding session.

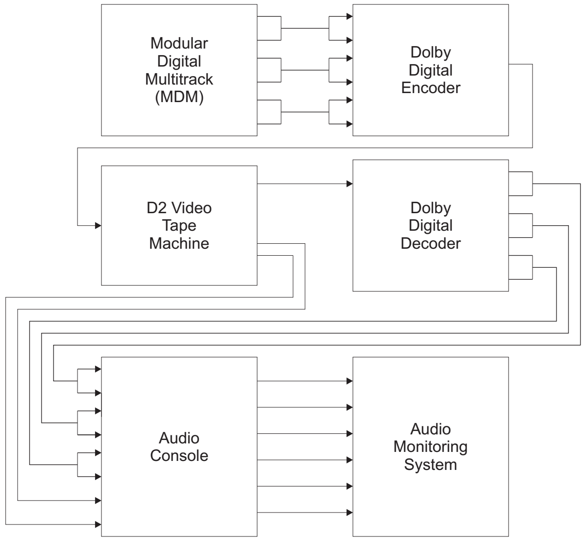


Figure 5-1 Typical Dolby Digital Laser Disc Encoding Setup

The Lt/Rt or stereo tracks on the D2 channels 3 and 4 can be used as guide tracks to confirm synchronization while listening to the decoded Dolby Digital bitstream.

5.5.3 Important Considerations

The output format of the Dolby Digital encoder must be set to “Pro 16-bit” to encode the Dolby Digital bitstream to a single channel of an AES/EBU pair. Pro 16-bit Ch1 is used for the odd number channels 1 and 3 while Pro 16-bit Ch2 is used for the even numbered channels 2 and 4.

Dolby Digital laser discs are always encoded with a sample rate of 48 kHz and a data rate of 384 kb/s.

During production the Dolby Digital bitstream is stored on a linear PCM channel. Since it is data and not linear PCM audio, the bitstream must therefore be recorded onto the linear PCM channel using certain guidelines.

- Use unity gain at all stages of input and output.
- Do not apply signal processing of any kind (this includes dithering).
- Do not edit the bitstream and record in one continuous pass.
- Use proper digital referencing to lock all digital audio and video equipment.
- Only use digital pathways in the audio and data chain beginning with the input of the Dolby Digital encoder through to the input of the Dolby Digital decoder.
- Record a pre- and post-roll of Dolby Digital silence at the beginning and end, respectively, of the encoded material. Standard practice is to begin the bitstream 30 seconds prior to the start of program and continue it 30 seconds after the end of program.

Chapter 6
Professional Encoders and Decoders

6.1 Dolby Digital Professional Encoders

A licensed and approved professional encoder must be used when Dolby Digital encoding content for DVD-Video, DVD-ROM, DVD-Audio, laser disc, and broadcast applications. Any content created with a Dolby Digital professional encoder is eligible to carry the Dolby Digital logo and will be compatible with any consumer decoder bearing the same Dolby Digital logo. Refer to Section 7.3, Trademark Usage, for more information on Dolby trademarks.

6.1.1 Software vs. Hardware

There are two fundamental types of Dolby Digital professional encoders on the market today: software encoders and hardware encoders. Each type of encoder has certain characteristics and advantages. The choice of encoder depends on the type of application and intended use. Hardware encoders operate in the real-time domain, which means that the input audio linear PCM bitstream is encoded into a Dolby Digital bitstream with very little delay (a few hundred milliseconds) in the process. A real-time encoder is ideal for live broadcast applications where the signal must be encoded and transmitted on the fly. Real-time encoders are also convenient for content generation for DVDs and laser discs since they allow real-time monitoring of the output using a professional Dolby Digital decoder. Software encoders, although usually non-real-time, are ideal for batch processing and encoding where a series of linear PCM files or bitstreams are to be encoded in a single session. It is reasonable to expect software encoders to reach real-time performance in the near future as the processing power of personal computers increases.

6.1.2 Licensed Dolby Digital Encoders and Quality

In addition to offering its own manufactured products, many different Dolby Digital professional encoders are licensed and approved by Dolby Laboratories. All licensed and approved encoders produce the same high quality audio signals. Whether an encoder is software or hardware, stand-alone or integrated, all materials generated with these products are eligible to carry the Dolby Digital logo indicating professional quality content.

6.1.3 Dolby Laboratories Encoders

Dolby Laboratories designs, manufactures, and sells Dolby Digital stand-alone hardware reference encoders, such as the Model DP569, for multichannel applications, and the Model DP567, for two-channel applications. These real-time reference encoders have been designed with professional features making them ideal for DTV broadcast applications and for DVD and laser disc content generation. The DP569 can encode from one to 5.1 channels in Dolby Digital, while the DP567 can encode one or two channels. For more information on the Dolby Laboratories manufactured encoders, refer to Section 7.2, Contacting Dolby Laboratories.

6.1.4 Dolby Laboratories Licensed Encoders

Dolby Laboratories licensees have produced both software and hardware Dolby Digital professional encoders for a variety of applications. Hardware encoders are available from Dolby licensees in many form factors such as PCI cards and integrated audio/video encoders.

6.1.5 Software Updates

Dolby Laboratories is committed to providing its customers and licensees with the highest quality encoders on an ongoing basis. Dolby Laboratories will continue to furnish all its customers and licensees with upgrades of encoder software and routines. To acquire the latest version of Dolby Digital professional encoder software, refer to Section 7.2, Contacting Dolby Laboratories.

6.2 Dolby Digital Professional Decoders

While a consumer decoder product may have many exciting and useful features, it usually does not have the flexibility and capability that is essential in a production environment. A professional decoder offers unique and useful features, such as the ability to monitor Dolby Digital bitstream parameters. This feature is essential anytime a production engineer is encoding audio material and needs to monitor the encoder output for correct parameter settings. Another important feature is the ability of a professional decoder to emulate any type of decoder on the market whether it is a DVD player, an A/V receiver, an HDTV, or a set-top box. Since most consumer decoders have some Dolby Digital features (Dialog Normalization, Dynamic Range Control (DRC), downmixing, etc.) preset at the factory, it is critical that the encoding engineer use a professional reference decoder to allow monitoring of all possible decoding options. Other features that differentiate a professional decoder from a consumer decoder are rack-mount style chassis and professional electrical connections (XLR, AES/EBU).

6.2.1 Dolby Laboratories Decoder

The Model DP562 is a Dolby Digital multichannel reference decoder manufactured by Dolby Laboratories. It is designed for high-quality monitoring of Dolby Digital bitstream parameters in real-time, and for keeping track of any errors or faults for quality assurance. The DP562 can emulate any Dolby Digital consumer decoder in addition to decoding Dolby Surround material using a built-in, digitally-implemented

Dolby Surround Pro Logic decoder. This unit is designed with professional features for DVD and laser disc production and DTV broadcast applications. For more information on the DP562, refer to Section 7.2, Contacting Dolby Laboratories.

6.2.2 Licensed Dolby Digital Professional Decoders

There are two types of licensed Dolby Digital professional decoders on the market today, broadcast baseband decoders for fixed-mode broadcast monitoring, and confidence decoders that are limited in features and are integrated with many Dolby Digital professional encoders.

Chapter 7
Miscellaneous Information

7.1 Technical Assistance

Dolby Laboratories provides technical support to content creators and encoder users in a variety of ways. Many technical documents are available for viewing or downloading on the Dolby web site at www.dolby.com. Printed copies of documents can also be obtained by sending e-mail to info@dolby.com with a description of the desired documents and a complete mailing address.

Dolby has a staff of engineers who can assist with audio production, encoding, and trademark usage. Dolby engineers are also available to provide on-site assistance with room configuration and calibration, audio production, and encoding. Telephone support is available free of charge, and local on-site support can often be provided without cost. In situations where extensive on-site support or long distance travel is required, standard engineering rates may apply.

If you would like technical support, please contact the nearest Dolby office at any of the locations listed in the following section.

7.2 Contacting Dolby Laboratories

In addition to its headquarters in San Francisco, Dolby has several offices around the world. All offices can provide information on audio production and encoding.

You may contact Dolby from anywhere in the world by e-mail using the addresses in Table 7-1.

Table 7-1 Dolby Email Contact Addresses

Address	Use
info@dolby.com	General information and inquiries
dvd@dolby.com	Questions on audio encoding for DVD
hdtv@dolby.com	Questions on audio production and encoding for DTV
Multimedia@dolby.com	Questions on multimedia applications
EncoderLicensing@dolby.com	Questions about encoder licensing
EncoderImplementations@dolby.com	Questions about encoder implementations
tsa@dolby.com	Applications for Dolby trademark agreements (TSA)

In addition, a wide variety of technical and trademark information can be found on Dolby's web site at www.dolby.com.

Information on local Dolby offices follows. Please contact the nearest office for direct assistance.

Corporate Headquarters

Dolby Laboratories Inc
100 Potrero Avenue
San Francisco, CA 94103-4813
Telephone 415-558-0200
Facsimile 415-863-1373

Los Angeles

Dolby Laboratories Inc
3375 Barham Boulevard
Los Angeles, CA 90068-1446
Telephone 323-845-1880
Facsimile 323-845-1890

New York

Dolby Laboratories Inc
1350 Avenue of the Americas
New York, NY 10019-4703
Telephone 212-767-1700
Facsimile 212-767-1705

UK Headquarters

Dolby Laboratories Inc
Wootton Bassett
Wiltshire, SN4 8QJ, England
Telephone (44) 1793-842100
Facsimile (44) 1793-842101

Shanghai Office

Dolby Laboratories Representative Office
7/Fl. Hai Xing Plaza, Unit H
Rui Jin Road (S)
Shanghai 2000023 China
Telephone (86) 21-6418-1015
Facsimile (86) 21-6418-1013

Tokyo Office

Dolby Laboratories International
Services Inc
Japan Branch
Fuji Chuo Building 6F
2-1-7, Shintomi, Chuo-ku
Tokyo 104-0041 Japan
Telephone (81) 3-5542-6160
Facsimile (81) 3-5542-6158

7.3 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. We do ask that these facilities refer their clients to us for trademark licensing information.

If you would like to use the Dolby Digital logo you can apply for a Dolby TSA by sending e-mail to tsa@dolby.com or by contacting Dolby Laboratories at any of the locations given in Section 7.2, Contacting Dolby Laboratories. When sending written requests please indicate that you would like a Dolby Digital trademark license and include your name, your company name, mailing address, and the type of media that your soundtracks or other audio programs will be distributed on (such as DVD, DVD-ROM, DTV broadcast, etc.).

For detailed information on Dolby trademark licensing, please refer to the document *Use of Dolby Trademarks on Audio and Video Media*, available on the Dolby web site at www.dolby.com. We are also planning to make our license application form available on-line, so check the Dolby web site in the future for the on-line version of the *Media Licensing Questionnaire*.

If you are already a Dolby licensee and would like more information on trademark use, please contact Dolby Laboratories. We are always happy to review artwork and assist with the proper use of our trademarks. Information on trademark licensing plus instructions for using the Dolby Digital trademark and marking audio features on DVD can also be found on the Dolby web site.

Appendix A
The Dolby Digital Algorithm - Theory of Operations

A.1 Introduction

In the broadest sense, Dolby Digital is a complete multichannel audio system. Dolby Digital encoders and decoders provide controls for important system features such as Dialog Normalization, Dynamic Range Control (DRC), channel downmixing, copyright notification, and many others discussed earlier in this manual. The most important feature of a Dolby Digital encoder, however, is its ability to reduce the data rate required to store or transport high-quality multichannel digital audio. Without data rate reduction, multichannel audio would simply not be a viable option for many applications in which data rate is scarce.

In order to accomplish this data rate reduction, Dolby Digital encoders make use of a sophisticated audio data compression algorithm called Dolby AC-3. Dolby AC-3 is a perceptual audio coder, also known as a *lossy* coder. The term *lossy* is used to indicate that the audio that comes out of the decoder is not identical to the source material that went into the encoder—some of the original information is lost. The goal of a perceptual audio coder is to ensure that whatever is lost is not perceptible. In other words, the output of the decoder, although numerically different, sounds the same as the original source material.

While a complete explanation of the inner workings of Dolby AC-3 is beyond the scope of this document, this section introduces the basic principles of the encoding and decoding algorithm. For more detailed information, please consult Document A/52 of the Advanced Television Systems Committee, *Digital Audio Compression Standard (AC-3)*, available at www.atsc.org, or on the Dolby Laboratories web site at www.dolby.com.

A.2 Perceptual Coding Principles

The primary task of a perceptual audio coder is to reduce the data rate necessary to represent a digital audio signal without introducing any audible differences. To achieve this, perceptual audio coders take advantage of several physiological limitations of the human hearing system. In other words, a perceptual coder predicts which sounds your ears will and will not hear, and only encodes audible sounds.

The human ear is a remarkably sensitive and accurate instrument, however, it does have limitations. One example is the well-known hearing threshold phenomenon. Simply stated, the ear is not equally sensitive at all frequencies. Our ears are easily able to detect quiet signals in the 2 kHz–4 kHz midrange, while they are much less sensitive to quiet signals at very low or very high frequencies. To varying degrees, this phenomenon occurs across the entire audible dynamic range including the threshold of hearing. The absolute hearing threshold varies as a function of frequency and sounds below the threshold are inaudible. Refer to Figure A-1.

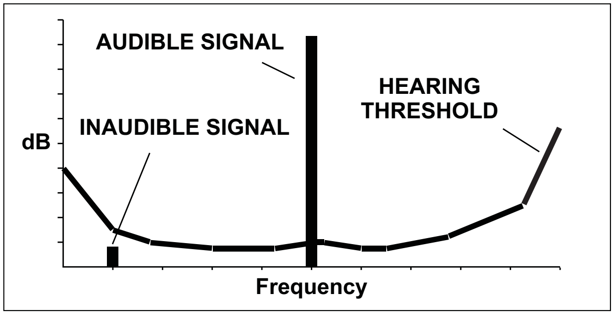


Figure A-1 Hearing Threshold

When the human hearing mechanism is stimulated by a complex acoustical waveform, regions along the basilar membrane within the inner ear are excited. Behaving as a filter bank, these regions correlate to frequency bands of varying widths known as critical bands. All aural processing within the brain is performed on the neural output from these regions.

The basilar membrane vibrates such that individual frequency components are not localized to a single point. Areas near the point of excitation also vibrate, resulting in a phenomenon known as frequency-domain masking. Frequency-domain masking occurs when two different signals are located close to one another in frequency. If one of the signals is much louder then the other, it is possible that the softer signal is not audible at all. Rather, the loud signal can obscure, or *mask*, any softer signals that are nearby in frequency. Refer to Figure A–2.

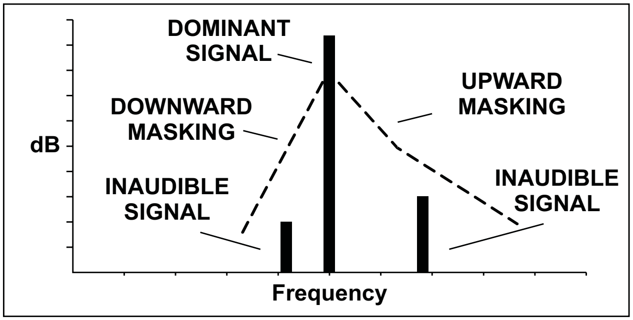


Figure A–2 Effect of Masking

Perceptual coders take advantage of these and other limitations of human hearing in order to remove inaudible information from the coded audio signal. By predicting how the ear will react to a complex audio signal, these coders are able to identify and remove a substantial amount of unnecessary information, and as a result achieve substantial data rate reduction while maintaining very high quality. Coding (quantization) noise is minimized through constraint to near the dominant frequency components in the audio signal. As a result, the subjective audio quality of the original signal is preserved. Refer to Figure A-3.

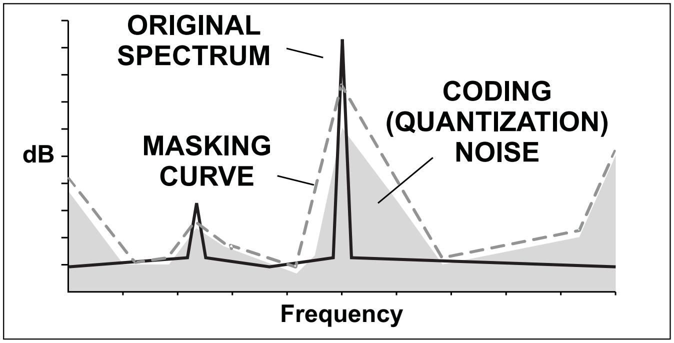


Figure A-3 Coding (Quantization) Noise Below the Masking Curve

Appendix B
Bitstream Format

Dolby Digital information can be conveyed as a bitstream when formatted onto S/PDIF (also referred to as IEC-958 or in the case of data, such as Dolby Digital, IEC 61937), or ATSC *A/52 Annex B* in one of several ways. In addition, a file may be recorded to a hard drive in either *big or little endian* format, and a *byte-reversal* process must be used to convert from one to the other.

The fundamental difference between a Dolby Digital bitstream transmitted through an S/PDIF connection and a bitstream saved as an .ac3 file is that the S/PDIF transmitted bitstream is padded with "zero" data. This is done to fill in the difference between the Dolby Digital data rate (e.g., 384 or 448 kb/s) and the serial transmission data rate (i.e., 48 kHz * 16 bits *2 channels = 1.536 Mb/s per channel).

B.1 Output Mode

There are four ways of placing the Dolby Digital data within the S/PDIF bitstream (Output Mode):

1. **Pro 32-bit**—Place alternate 16-bit words of Dolby Digital data in left and right channels of the audio bitstream, and pad with zeros until the next frame is available (Professional).
2. **Pro 16-bit Ch1**—Place consecutive 16-bit words of Dolby Digital data in the left channel only, leaving the right channel available for conveying PCM data (Professional, Left).
3. **Pro 16-bit Ch2**—Same as 2, except Dolby Digital data is packed in the right channel. (Professional, Right).
4. **Consumer Mode**—Same as 1, except used in consumer applications.

The Dolby Digital Recorder program available from Dolby Laboratories works with any of these formats. The program accepts the 1.536 Mb/s bitstream and ignores the zero padding, saving only the relevant data at the Dolby Digital data rate (e.g., 384 kb/s). The resulting file consists of 1536 16-bit data words per Dolby Digital frame. These 16-bit words may be written to disc either high-order byte first, or low-order byte first, and this depends on the processor and operating system being used.

A Dolby Digital bitstream can also be recorded into a computer using a digital soundcard and audio capture software. This results in a data file marked “Audio,” which if played back as audio, results in a *bursty* noise.

Figure B-1 shows a screen capture of Dolby Digital data with zero padding.

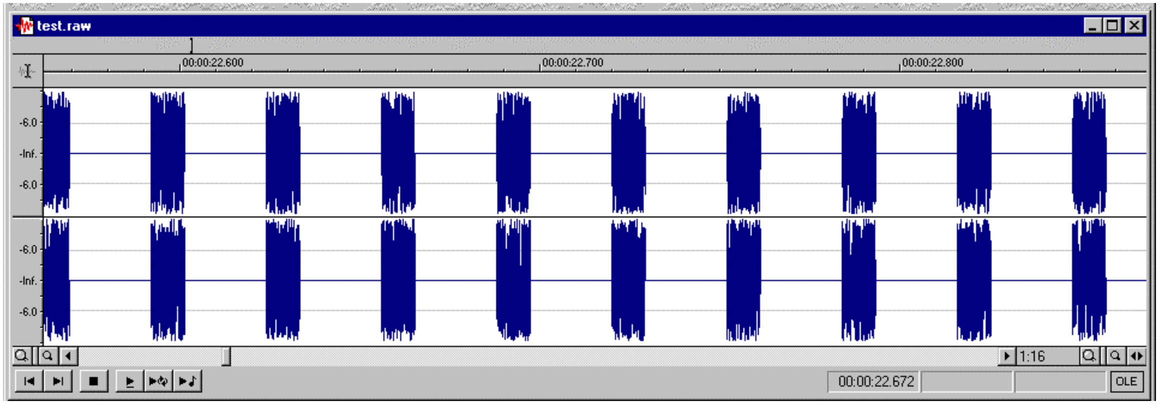


Figure B-1 Screen Capture of Dolby Digital Data with Zero Padding

A file saved on a personal computer may also have to be *byte-swapped*, or the order of the high-order and low-order bytes reversed, in order to be read correctly by a UNIX based workstation, for example. Consult a computer software vendor for commercially available programs that can perform this task.

B.2 Audio/Non-Audio Bit

Most, if not all, digital audio cards for computers supply the S/PDIF digital audio bitstream marked as "Audio" data. Many consumer decoders use this indicator to determine if the serial data bitstream contains linear PCM or Dolby Digital data. A consumer decoder receiving input from an audio card may fail to decode the Dolby Digital data, reproducing the data as bursts of full-level noise at the Dolby Digital frame rate (32 Hz).

In order to avoid this situation, the "Audio/Non-Audio" bit within the S/PDIF bitstream from the digital audio card must be set to indicate non-audio data. Commercial products are available that can accomplish this, such as the Lexicon LFI-10, or a simple circuit can be constructed to read the bitstream and invert the "Audio/Non-Audio" bit.

Appendix C
Dynamic Range Control (DRC)

C.1 Background

Dolby Digital uses a novel approach to applying Dynamic Range Control (DRC) to audio program material. Rather than compressing the dynamic range of the audio in an irreversible way, Dolby Digital encoders generate *compression gain* (also referred to as control) *words* that are carried in the Dolby Digital bitstream. When the bitstream is decoded, the compression gain words are applied to the audio material according to user settings. Dolby Digital decoders can be commanded to provide full, reduced, or even no dynamic range compression at all. This allows end users to adjust the amount of dynamic range compression to suit individual tastes and needs.

The compression gain words are computed based on a number of separate input parameters, including the audio program material, the program Dialog Normalization (*dialnorm*) value, and the selected Dynamic Range Compression profile (also referred to as a preset). Although these words are referred to as *gain* words, they can take on negative values and thus actually represent program attenuation.

An accurate setting of the *dialnorm* value is crucial to the proper operation of DRC. The *dialnorm* value is a Dolby Digital parameter setting 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. This level is specified on an absolute scale ranging from -1 dBFS to -31 dBFS. Dolby Digital decoders attenuate programs based on the *dialnorm* value in order to achieve uniform playback level. This feature of Dolby Digital is termed Dialog Normalization (also referred to as volume normalization). The amount of adjustment is determined by the difference between the *dialnorm* value and the relevant reference playback level (the *dialnorm* reference playback level is -31 dBFS for Line mode; +11 dB raises this level to -20 dBFS for RF mode). As a result, all DRC calculations consider the input level *relative to the dialnorm value*, and not in an absolute sense. Thus, setting the *dialnorm* value properly is a critical step in calibrating the DRC system.

There are currently six Dynamic Range Compression profiles, described in detail in Table C-1: Film Standard, Film Light, Music Standard, Music Light, Speech, or None. Five compression regions including a *null band* are defined for each profile. The null band is positioned relative to the *dialnorm* reference playback level with its upper and lower limits determined by the Dynamic Range Compression profile. Audio lying within the null band is unaffected whereas excursions above and below are subject to dynamic range compression, with softer material being amplified and louder material attenuated.

There are two forms of compression gain words in the Dolby Digital bitstream: *dynrng* words and *compr* words. The *dynrng* words are used in Line Mode and occur in the Dolby Digital bitstream once every 5.3 milliseconds (48 kHz sample rate). The *compr* words are intended primarily for use in DRC for set-top boxes that are connected to the RF (antenna) input of a television set (RF Mode). These words occur once every 32 milliseconds (48 kHz sample rate), and generally yield greater negative values due to more aggressive overload protection than the *dynrng* words.

Even though these gain words are sent as discrete values, and can change abruptly from one word to the next, the application of the gain is always applied smoothly from sample to sample in the Dolby Digital decoder.

C.2 Dynamic Range Control (DRC) Algorithm Overview

The Dolby Digital encoder generates the compression gain words from the input parameters using a sophisticated algorithm. In addition to computing the desired dynamic range adjustment, this algorithm also includes an overload protection component that ensures that decoders applying the compression gain words do not result in overload, even when downmixing multiple coded channels to fewer output channels. The high-level overview of this algorithm is shown in Figure C-1.

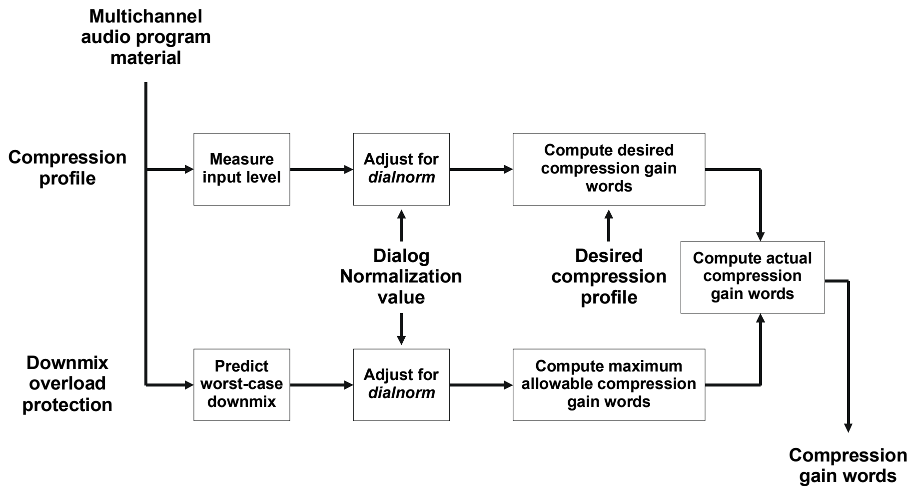


Figure C-1 Overview of Dynamic Range Control Algorithm

The multichannel audio program is first analyzed to determine the overall input level, as well as to determine the worst-case linear PCM downmix. Both of these measures are then adjusted as necessary by the Dialog Normalization (*dialnorm*) value.

Next, the *dialnorm*-adjusted input level measurement is used to compute the desired compression gain words, based on the selected Dynamic Range Compression profile. Separately, the *dialnorm*-adjusted worst-case downmix value is used to compute the maximum allowable compression gain word, i.e., the largest gain word that can be accommodated by the decoder without causing overload.

Finally, the desired compression gain word is compared to the maximum allowable gain word to determine the actual compression gain word that can be inserted into the Dolby Digital bitstream. If the desired gain word will result in decoder overload, then the actual gain word is limited to a value no larger than the maximum allowable gain word.

C.3 Compression Characteristic

The compression characteristic used by the Dynamic Range Compression profiles is made up of five regions that encompass the entire audio dynamic range. The regions are: constant boost for very soft signals, variable boost for moderately soft signals, no compression action for average signals (the null band), variable cut for moderately loud signals, and constant cut for very loud signals.

These five regions are defined using five key parameters: the maximum boost and maximum cut, the boost ratio and cut ratio, and the null band width as shown in the center portion of Figure C-2. A sixth parameter, not shown in Figure C-2, specifies where the null band is located relative to the *dialnorm* reference playback level—from this location, the location of all other regions can be determined.

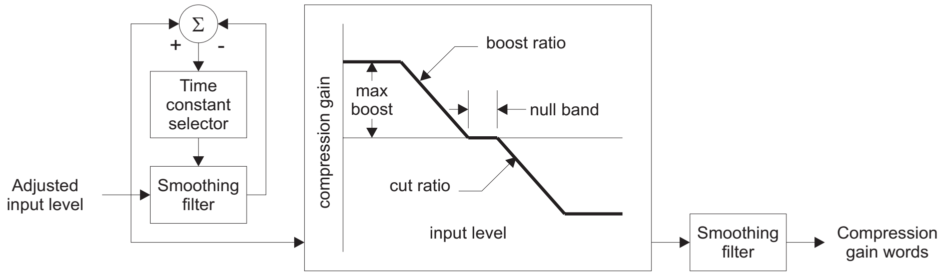


Figure C-2 Dynamic Range Compression Core

The purpose of the null band is to provide an area in which no compression action is applied to the program material. This is useful if the program has already been compressed, or is of limited dynamic range. The location of this null band is very important: It is located relative to the *dialnorm* reference playback level so that normalized signals are not further amplified or attenuated. For some of the Dynamic Range Compression profiles, the null band is positioned so its lower edge coincides with the reference level. This is done to minimize the compression action for excursions above the *dialnorm* reference playback level. Excursions below the reference level are already minimally processed due to the rather slow time constants for compression decay.

C.4 Dynamic Range Compression Profiles

All of the parameters that describe the dynamic range compression core are specified by the selected Dynamic Range Compression profile. This includes the time constant selection parameters (decay and attack thresholds, fast and slow time constants for both attack and decay) as well as the compression characteristic (max boost and cut, boost and cut ratio, null bandwidth and position).

Table C-1 provides the specific parameter values for each of the Dynamic Range Compression profiles (in Line Mode). These profiles are available in version 6.3.0 of the Dolby Digital professional encoder software. The type of profiles and the values presented are subject to revision in future software releases. For the compression characteristic fields, the range of input level is shown for each of the regions, measured in dB relative to a full-scale sine wave.

Table C-1 Dynamic Range Compression Profile Parameter Values

Profile Name	Film Standard	Film Light	Music Standard	Music Light	Speech
time constant (tc) selection					
attack threshold	15 dB	15 dB	15 dB	15 dB	10 dB
decay threshold	20 dB	20 dB	20 dB	20 dB	10 dB
fast attack tc	10 ms	10 ms	10 ms	10 ms	10 ms
slow attack tc	100 ms	100 ms	100 ms	100 ms	100 ms
slow decay tc	3 s	3 s	10 s	3 s	1 s
fast decay tc	1 s	1 s	1 s	1 s	200 ms
hold off period	10 blocks (53 ms)	10 blocks (53 ms)	10 blocks (53 ms)	10 blocks (53 ms)	10 blocks (53 ms)
Compression characteristic					
max boost (abs range)	6 dB (-43 dBFS)	6 dB (-53 dBFS)	12 dB (-55 dBFS)	12 dB (-65 dBFS)	15 dB (-50 dBFS)
boost ratio (abs range)	2:1 (-43 to -31)	2:1 (-53 to -41)	2:1 (-55 to -31)	2:1 (-65 to -41)	5:1 (-50 to -31)
null band width (abs range)	10 dB (-31 to -21)	20 dB (-41 to -21)	10 dB (-31 to -21)	20 dB (-41 to -21)	10 dB (-31 to -21)
cut ratio (abs range)	20:1 (-21 to +4)	20:1 (-21 to +4)	20:1 (-21 to +4)	2:1 (-21 to +9)	20:1 (-21 to +4)
max cut (abs range)	24 dB (+4 dBFS)	24 dB (+4 dBFS)	24 dB (+4 dBFS)	15 dB (+9 dBFS)	24 dB (+4 dBFS)

See notes to table on next page.

- The *dialnorm* reference playback level of -31 dBFS (Line mode) provides ample headroom to accommodate program material with a very wide dynamic range (signal peaks being possibly 20 dB or greater than the average level). In many cases, there is sufficient headroom for downmixed signals as well. The absolute levels shown in the compression characteristic fields are derived from the *dialnorm* reference playback level and are therefore lower than typical for traditional compressors or limiters.
- Some absolute ranges extend higher than 0 dBFS. Since a full-scale sine wave cannot exceed 0 dBFS, these ranges should be interpreted as extrapolated extensions of the allowable range. As a result, it may not be possible in practice to achieve the maximum cut compression gain words.
- If the "None" profile is selected, the dynamic range compression core generates desired compression gain words that are set to 0 dB gain, i.e., no boost or cut. Selecting the "None" profile however, does not disable the downmix overload protection function and it is possible that the actual compression gain words in the Dolby Digital bitstream will be less than 0 dB.

Appendix D
Dolby Digital Time-Domain Filters

There are five time-domain filters available in Dolby Digital professional encoders. They are the 90-Degree Phase Shift filter, the Digital Deemphasis filter (not supported in the Dolby Model DP561 encoder), the DC Highpass Filter, the Channel Bandwidth Lowpass Filter (not supported in the Dolby Model DP561 encoder), and the LFE Lowpass Filter.

D.1 90-Degree Phase-Shift Filter

Purpose

The 90-Degree Phase Shift filter provides a means for an encoding engineer to create a multichannel Dolby Digital bitstream that can be downmixed to a Dolby Surround-compatible Lt/Rt output. Without this filter, point-source elements panned from Surround to Center in the multichannel mix would seem to pan from Surround to Left and then to Center when downmixed to Lt/Rt and reproduced using a Dolby Surround Pro Logic decoder.

This filter should generally be used whenever encoding a multichannel signal unless it is known that the 5.1-channel source does not contain point-source element pans. For example, if the source was recorded using five discrete microphones placed in the corners of an auditorium, there is no panning between channels and the filter could be safely disabled. If in doubt, use a DP562 to downmix the 5.1-channel program to Lt/Rt, Dolby Surround Pro Logic decode the Lt/Rt signals, and then set the filter to the setting that sounds best.

Description

The 90-degree phase-shift is created using a very long FIR filter. Since this filter introduces a significant time delay, the other four channels are delayed using a PCM delay line so that all six channels are kept in sample alignment. This filter has exactly 90-degree phase shift at all frequencies. The magnitude response is flat across most of the spectrum, rolling off at the lower edge of the audio band (-3 dB below 30 Hz).

D.2 Digital Deemphasis Filter

Purpose

The Digital Deemphasis filter is used to de-emphasize any 50/15 µs pre-emphasized linear PCM signals that may be presented to the inputs of the Dolby Digital encoder. Pre-emphasis is a technique that was once commonly used to reduce the harshness of A/D and D/A converters. In a system using pre-emphasis, the analog input signal is passed through a high-frequency-boost shelf filter and then A/D converted. At the output of the playback system, the D/A converter reconstructs the analog output signal and passes it through a high-frequency-cut shelf filter. Any high-frequency noise introduced in the A/D or D/A would be attenuated by the shelf filter cut amount (10 dB). Pre-emphasized linear PCM signals are identified using a special channel status subside bit in the S/PDIF or AES/EBU digital audio bitstream. Depending on the state of this bit, the D/A converter knows whether or not to pass the analog signal through the high-frequency-cut shelf filter.

Dolby Digital bitstreams do not carry this pre-emphasis bit, and thus all Dolby Digital bitstreams are assumed to be encoded without pre-emphasis. If pre-emphasized linear PCM samples are to be encoded with a Dolby Digital encoder, they must first be de-emphasized otherwise the output of the decoder sounds unnaturally bright. This can be done by using the Digital Deemphasis filter that is built into the Dolby Digital encoder. The preferred method for this is to have the filter automatically switch in whenever the pre-emphasis bit is detected in the S/PDIF or AES-EBU channel status subcode fields.

Description

The Digital Deemphasis filter is a first-order high-frequency-cut shelf filter designed to match the required analog filter at all sample rates. The Dolby Model DP561 encoder does not support this filter.

D.3 DC Highpass Filter

Purpose

The DC Highpass Filter is used to block DC from being Dolby Digital encoded. This is important, as a DC offset requires some amount of data rate to encode even though it is not audible, thereby wasting bits. Another benefit of using this filter is that the meter values do not get stuck at the DC offset level during very quiet passages (DC offset can easily be greater than -60 dBFS). The DC Highpass Filter should always be enabled unless the encoding engineer is absolutely sure that there is no DC in the input audio. To check for DC, disable the filter and view the meters during very quiet passages to see if they remain fixed above -60 dBFS.

Description

The DC Highpass Filter is a first-order high-pass filter with unity gain across the entire audible spectrum, and a rolloff at very low frequencies (-3 dB below 1 Hz).

D.4 Channel Bandwidth Lowpass Filter

Purpose

The Channel Bandwidth Lowpass Filter is used to roll off the high frequency content in the input signal at a frequency just below that specified by the Dolby Digital audio bandwidth boundary. Using this filter ensures that the audio signal is completely contained within the Dolby Digital audio bandwidth. By providing a smooth transition at the upper bandwidth edge, this filter helps to minimize artifacts that may arise if the input signal contains significant high-frequency energy. In general, this filter should be enabled unless the encoding engineer is confident that the input signal does not contain appreciable high-frequency energy above the Dolby Digital audio bandwidth.

Description

The Channel Bandwidth Lowpass Filter is a sixth-order low-pass filter with unity gain over most of the audio spectrum and a steep rolloff just below the Dolby Digital audio bandwidth boundary. Note that the Dolby Digital audio bandwidth is variable, and depends on the selected Data Rate, Audio Coding Mode (also referred to as Channel Mode), and input audio sample rate. The Dolby Model DP561 encoder does not support this filter.

D.5 LFE Lowpass Filter

Purpose

The LFE Lowpass Filter is used to band-limit the LFE channel input signal in the Dolby Digital encoder. If this filter is not enabled, wideband signal content fed to the LFE channel input of the encoder will produce significant audible artifacts at the output of the decoder. This filter should always be enabled unless it is known that the LFE channel input signal has already been properly band-limited.

Description

The LFE Lowpass Filter is an eighth-order low-pass filter with unity gain at low frequencies and a steep rolloff at the upper edge of the LFE bandwidth (-3 dB above 120 Hz).

Appendix E
Mix and Mastering Data Sheets

Dolby Mix Data Sheet

Date ____/____/____

Project ____________________ **Project No.** ____________________

Client ____________________ **Producer** ____________________

Studio ____________________ **Engineer** ____________________

Sampling Frequency	☐ 32 kHz	☐ 44.1 kHz	☐ 48 kHz	
Bit Resolution	☐ 16-bit	☐ 18-bit	☐ 20-bit	☐ 24-bit
Time Code Format	☐ 25 fps	☐ 29.97 NDF		
Tape Format	☐ ADAT	☐ DA-88	☐ ½" Digital	☐ Data Cartridge (JAZ)
Surround Level SPL Calibration	☐ Equal to Front	☐ -3 dB to Front		

CHI	CH2	CH3	CH4	CH5	CH6	CH7	CH8

Time Code	Program
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	

Notes:

Dolby Mastering Information

Date ____/____/____

Project ______________ **Project No.** ______________

Client ______________ **Producer** ______________

Studio ______________ **Engineer** ______________

Sampling Frequency ☐ 32 kHz ☐ 44.1 kHz ☐ 48 kHz

Bit Resolution ☐ 16-bit ☐ 18-bit ☐ 20-bit ☐ 24-bit

Time Code Format ☐ 25 fps ☐ 29.97 NDF

Tape Format ☐ ADAT ☐ DA-88 ☐ ½" Digital ☐ Data Cartridge (JAZ)

Dolby Digital Encoding Information	Bitstream Information	Processing
Audio Service Configuration LFE Filter ☐ ON ☐ OFF Dial Norm Setting: ________ Mix Level: ________ Data Rate: ________	☐ Default ☐ Copyright Bit Center Mix Level: ________ Surround Mix Level: ________	☐ Digital Deemphasis ☐ DC Highpass Filter ☐ Bandwidth Lowpass ☐ LFE Lowpass Filter
Surround Channel Processing ☐ 90-Degree Phase-Shift ☐ 3 dB Attenuation	**Dynamic Range Compression** ☐ None ☐ Speech Film: ☐ Std. ☐ Light Music: ☐ Std. ☐ Light	

Mode	CHI	CH2	CH3	CH4	CH5	CH6	CH7	CH8

Time Code	Program
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	
: : ;	

Appendix F
Glossary

Note*: The items with the in the definition are Dolby terms.*

AC-3	Audio Coding algorithm number 3. Refer to Dolby Digital.
ADPCM (Adaptive Differential Pulse Code Modulation)	A pulse code modulation (PCM) system typically operating at a high sampling rate whereby coding is based on a prior knowledge of the signal to be processed (i.e., greater than, equal to, or less than the previous sample). The system is adaptive in that digital bits of code signify different sizes of signal change depending on the magnitude of the signal. Refer to PCM.
AES (Audio Engineering Society)	The official association of technical personnel, scientists, engineers, and executives in the audio field.

AES/EBU (Audio Engineering Society/European Broadcasting Union) interface	The serial transmission format standardized for professional digital audio signals (AES3-1992 AES Recommended Practice for Digital Audio Engineering – Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data). A specification using time division multiplex for data, and balanced line drivers to transmit two channels of digital audio data on a single twisted-pair cable using 3-pin (XLR) connectors. Peak-to-peak values are between 3 and 10V with driver and cable impedance specified as 110 ohms.
ATSC (Advanced Television Systems Committee)	The Advanced Television Systems Committee was formed to establish voluntary technical standards for advanced television systems, including high definition digital television (HDTV).
Bitstream (also bit stream)	A binary signal without regard to grouping according to character. A continuous series of bits transmitted on a line.
Codec	A device for converting signals from analog to coded digital and then back again for use in digital transmission schemes. Most codecs employ proprietary coding algorithms for data compression.
DAB (Digital Audio Broadcasting)	NRSC (National Radio Systems Committee) term for the next generation of digital radio broadcast.
Data rate	The speed at which digital information is transmitted, typically expressed in hertz (Hz), bits/second (b/s), or bytes/sec (B/s).

Data reduction (also referred to as data compression)
The process of reducing the number of recorded or transmitted digital data samples through the exclusion of redundant or unessential samples.

DBS
Direct Broadcast by Satellite.

De-emphasis
A change in frequency response characteristic complementary to that introduced by pre-emphasis. Refer to Pre-emphasis.

Dialog Normalization
Refer to Chapter 4.

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. 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.

Dolby Digital (formerly AC-3)
A perceptual audio coding system based upon transform coding techniques and psycho-acoustic principles. Frequency-domain processing takes full advantage of noise masking by confining quantization noise to narrow spectral regions where it will be masked by the audio signal. Designed as an emissions (delivery) system, Dolby Digital provides flexible coding of up to 5.1 audio channels at a variety of data rates. In addition, Dolby Digital bitstreams carry informational data about the associated audio Refer to Metadata.

Dolby Surround	A passive system that matrix encodes four channels of audio into a standard two-channel format (Lt/Rt). When the signal is decoded using a Dolby Surround Pro Logic decoder, the left, center and right signals are recovered for playback over three front speakers and the surround signal is distributed over the rear speakers. Refer to Lt/Rt and Dolby Surround Pro Logic.
Dolby Surround Pro Logic (DSPL)	An active decoding process designed to enhance the sound localization of Dolby Surround encoded programs through the use of high-separation techniques. Dolby Surround Pro Logic decoders continuously monitor the encoded audio program and evaluate the inherent soundfield dominance, applying enhancement in the same direction and in proportion to that dominance. Refer to Dolby Surround and Lt/Rt.
Downmix	A process wherein multiple channels are summed to a lesser number of channels.
DRC (Dynamic Range Control)	A feature of Dolby Digital that allows the end user to retain or modify the dynamic range of a Dolby Digital encoded program upon playback. The amount of control is dictated by encoder parameter settings and decoder user options.
DSS	Direct Satellite System.
DTV (Digital Television)	A term used for all types of digital television including High Definition Television (HDTV) and Standard Definition Television (SDTV).

DVB	Digital Video Broadcasting.
DVD (Digital Versatile Disc)	A type of optical disc. There are two possible sizes: 12 cm (standard) and 8 cm. There are five types of 12 cm discs: DVD-5, single-sided, single-layer 4.7 billion byte capacity; DVD-9, single-sided, dual-layer, 8.54 billion byte capacity; DVD-10, dual-sided, single-layer, 9.4 billion byte capacity; DVD-14, dual-sided, single/dual-layer, 13.24 billion byte capacity; DVD-18, dual-sided, dual-layer, 17.08 billion byte capacity. There are three formats of read-only discs: DVD-ROM (games and computer use), DVD-Video (movies), and DVD-Audio (music-only). There are also write-once and rewritable disc formats.
Dynamic Range	The ratio, expressed in decibels (dB), of the maximum to the minimum signal level, whether alone or in relation to a device or system.
Dynamic Range Compression	Level adjustment applied to an audio signal in order to limit the difference, or range of the loudest to the softest sounds.

EBU (European Broadcasting Union)	Created in 1950 and headquartered in Geneva, Switzerland, the EBU is the world's largest professional association of national broadcasters. The EBU assists its members in all areas of broadcasting, briefing them on developments in the audiovisual sector, providing advice, and defending their interests via international bodies. The Union has active members in European and Mediterranean countries and associate members in countries elsewhere in Africa, the Americas, and Asia.
HDTV (High Definition Television)	The standard for digital television in North America that includes a definition for picture quality of at least two million pixels (compared to 336,000 pixels for NTSC). The audio standard is Dolby Digital.
Headroom	The difference (in dB) between the nominal level (average) and the maximum operating level (just prior to “unacceptable” distortion) in any system or device. Because it is a pure ratio, there is no unit or reference-level qualifier associated with headroom—simply "dB"; headroom expressed in dB accurately refers to both voltage and power.
IEC (International Electrotechnical Commission)	A European organization (headquartered in Geneva, Switzerland) involved in international standardization within the electrical and electronics fields.

ITU (International Telecommunications Union)	Headquartered in Geneva, Switzerland, this international organization is involved in coordinating global telecom networks and services. Both governmental and private sector representatives are included.
JPEG (Joint Photographic Experts Group)	This is a group of experts nominated by national standards bodies and major companies to work to produce standards for continuous tone image coding. The 'joint' refers to its status as a committee working on both ISO and ITU-T standards. The 'official' title of the committee is ISO/IEC JTC1 SC29 Working Group 1, and is responsible for both JPEG and JBIG standards.
LAeq	An Leq measurement using A weighting. Refer to Leq and Weighting.
Leq	Leq represents the continuous noise level, equivalent in loudness and energy, to the fluctuating sound signal under consideration. Refer to LAeq.
LFE (Low-Frequency Effects)	The optional LFE channel (also referred to as the “boom” channel) carries a separate, limited, frequency bandwidth signal that complements the main channels. It delivers bass energy specifically created for subwoofer effects or low-frequency information derived from the other channels. The LFE channel is the “.1” in 5.1-channel audio.

Line Mode

A Dolby Digital decoder operational mode. The *dialnorm* reference playback level is -31 dBFS and *dynrng* words are used in dynamic range compression. Refer to Dynamic Range Compression.

Lo/Ro (Left only, Right only)

A type of two-channel downmix for multichannel audio programs. Lo/Ro downmixes are intended for applications where surround playback is neither desired nor required.

LPCM (Linear Pulse Code Modulation)

A pulse code modulation (PCM) system in which the signal is converted directly to a PCM word without companding, or other processing. Refer to PCM.

LS/RS (Left Surround, Right Surround)

The actual channels or speakers delivering discrete surround program material.

Lt/Rt (Left total, Right total)

Two-channel delivery format for Dolby Surround. Four channels of audio, Left, Center, Right, and Surround (LCRS) are matrix encoded for two-channel delivery (Lt/Rt). Lt/Rt encoded programs are decoded using Dolby Surround and Dolby Surround Pro Logic decoders. Refer to Dolby Surround and Dolby Surround Pro Logic.

Metadata ("data about the data")

The descriptive and supporting data that is connected to the program or the program elements. It is intended to both aid the direct use of program content and support the retrieval of content as needed during the post-production process.

MPEG (Moving Picture Experts Group)	A working group within SMPTE who set, among other things, specifications for compression schemes for audio and video transmission. A term commonly used when referring to their associated data compression technologies (MPEG).
NTSC (National Television System Committee)	Named after a committee that worked with the FCC in formulating standards for the current United States analog color television system. Now describes the American system of color telecasting consisting of 525 lines transmitted at 29.97interlaced frames per second. It is a composite of red, green, and blue signals for color and includes a FM frequency for audio and an MTS signal for stereo.
PAL (Phase Alternating Line)	A European television standard that uses 625 lines of resolution (100 more than NTSC) transmitted at 25 interlaced frames (50 fields) per second.
PCM (Pulse Code Modulation)	Pulsed modulation in which the analog signal is sampled periodically and each sample is quantized and transmitted as a digital binary code.
Peak value	The maximum numerical value reached whether the signal is positive or negative.
Perceptual Coding	Refer to Appendix A, The Dolby Digital Algorithm - Theory of Operations.

Pink noise

A type of noise whose amplitude is inversely proportional to frequency over a specified range. Pink noise is characterized by a flat amplitude response per octave band of frequency (or any constant percentage bandwidth), i.e., it has equal energy, or constant power, per octave. Pink noise can be created by passing white noise through a filter having a 3 dB/octave slope.

Pre-emphasis

An intentional change made in the frequency response of a recording system to improve the signal-to-noise (S/N) ratio or to reduce distortion. Typically, a high-frequency boost is used during recording, followed by complementary de-emphasis (a high-frequency cut) during playback.

RF Mode

A Dolby Digital decoder operational mode intended primarily for cable set-top boxes that are connected to the RF (antenna) input of a television set. The *dialnorm* reference playback level is -20 dBFS and *compr* words are used in dynamic range compression. Refer to Dynamic Range Compression.

RMS (Root Mean Square)

The value assigned to an alternating current or voltage that results in the same power dissipation in a given resistance as dc current or voltage of the same numerical value. Calculated as 0.707 of peak amplitude of a sine wave at a given frequency.

SDTV

Standard Definition Television.

SECAM ("Séquentiel Couleur avec Memoire" sequential color with memory)
A color television system with 625 lines per frame and 50 fields per second developed by France and the former USSR and used in some countries that do not use either NTSC or PAL.

S/N
Refer to SNR.

SNR (Signal to Noise Ratio)
The ratio of the magnitude of the signal to that of the noise, generally expressed in decibels. An audio measurement of the residual noise of a unit, stated as the ratio of signal level (or power) to noise level (or power), normally expressed in decibels. The "signal" reference level must be stated. Typically this is either the expected nominal operating level, +4 dBu for professional audio, or the maximum output level, usually around +20 dBu. The noise is measured using a true RMS type voltmeter over a specified bandwidth, and sometimes using weighting filters. All of these criteria must be stated for an SNR specification to have meaning.

SMPTE (Society of Motion Picture and Television Engineers)

An international technical society devoted to advancing the theory and application of motion-imaging technology including film, television, video, computer imaging, and telecommunications. Members of the Society are engineers, executives, technical directors, cameramen, editors, consultants, and specialists in film processing, film and television production and post-production, and practitioners from almost every other discipline in the motion-imaging industry. The Society was founded in 1916, as the Society of Motion Picture Engineers. The T was added in 1950 to embrace the emerging television industry. The SMPTE is recognized around the globe as a leader in the development of standards and authoritative, consensus-based recommended practices (RPs), and engineering guidelines (EGs). The Society serves all branches motion imaging including film, video, and multimedia.

S/PDIF (Sony/Philips Digital Interface Format)

A consumer version of the AES/EBU digital audio interconnection standard. The format uses a 75-ohm coaxial cable with RCA connectors and has nominal peak-to-peak values of 0.5V. The frame ordering differs slightly than that of AES/EBU, specifically in the channel status information. Refer to AES/EBU interface.

S/PDIF is equivalent to IEC 61937 when used for data, as in Dolby Digital.

SPL (Sound Pressure Level)	The SPL of a sound is equal to twenty times the logarithm (base 10) of the ratio of the root-mean-square (RMS) sound pressure to the reference sound pressure. As a point of reference, 0 dB-SPL equals the threshold of hearing, while 140 dB-SPL produces irreparable hearing damage.
Weighting	In a sound level meter, this is a filter that creates a response that corresponds to the ear's varying sensitivity at different loudness levels. *A weighting* corresponds to the sensitivity of the ear at lower listening levels. The filter design *weights* or is more sensitive in certain frequency bands than others. The goal is to obtain measurements that correlate well with the subjective perception of noise.
ANSI A-weighting	The A-curve is a wide bandpass filter centered at 2.5 kHz, with ~20 dB attenuation at 100 Hz, and ~10 dB attenuation at 20 kHz. Therefore, it tends to heavily roll off the low end, with a more modest effect on high frequencies. It is essentially the inverse of the 30-phon (or 30 dB-SPL) equal-loudness curve of a Fletcher-Munson.
ANSI B-weighting	The B-weighting curve is used for intermediate level sounds and has the same upper corner as the C-weighting, but the lower amplitude corner is 120 Hz.
ANSI C-weighting	The C-curve is basically "flat," with -3 dB corners of 31.5 Hz and 8 kHz, respectively.

CCIR 468-weighting	This filter was designed to maximize its response to the types of impulsive noise often coupled into audio cables as they pass through telephone switching facilities. The CCIR 468-curve peaks at 6.3 kHz, where it has 12 dB of gain (relative to 1 kHz). From here, it gently rolls off low frequencies at a 6 dB/octave rate, but it quickly attenuates high frequencies at ~30 dB/octave (it is down -22.5 dB at 20 kHz, relative to +12 dB at 6.3 kHz).
CCIR ARM-weighting or CCIR 2 kHz-weighting	This curve derives from the CCIR 468-curve above. Dolby Laboratories proposed using an average-response meter with the CCIR 468-curve instead of the costly true quasi-peak meters used by the Europeans in specifying their equipment. They further proposed shifting the 0 dB reference point from 1 kHz to 2 kHz (in essence, sliding the curve down 6 dB). This became known as the CCIR ARM (average response meter), as well as the CCIR 2 kHz-weighting curve.
White Noise	A random signal having the same energy level at all frequencies (in contrast to pink noise which has constant power per octave band of frequency).

Appendix G
Bibliography

Davidson, Grant A., Fielder, Louis D., and Link, Brian D., “Parametric Bit Allocation in a Perceptual Audio Coder,” AES 97th Convention, November 1994, Preprint 3921, available on the Dolby web site at www.dolby.com/tech/.

Davis, Mark F., “The AC-3 Multichannel Coder,” AES 95th Convention, October, 1993, available on the Dolby web site at www.dolby.com/tech.

Davis, Mark F., “The Big Squeeze, The Theory and Practice of Dolby Digital,“ *Audio*, Vol. 81, No. 7, July 1997.

Davis, Mark F. and Todd, Craig C., “AC-3 Operation, Bitstream Syntax, and Features,” AES 97th Convention, November, 1994, Preprint 3910.

Dolby Laboratories, *Dolby Surround Mixing Manual*, available on the Dolby web site at www.dolby.com/tech/.

Dressler, Roger, "Dolby Pro Logic Surround Decoder, Principles of Operation,” available on the Dolby web site at www.dolby.com/tech.

Gundry, Kenneth, “Keeping the Audio Unchanged,” *Surround Professional*, Vol. 1, No. 2, January 1999.

Todd, Craig C., "Distribution and Emission of Multichannel Audio Programs, and Implications for Post-Production," 139th SMPTE Technical Conference, November 1997, Paper 139-15; also “Film and Video Origination in the Era of DTV Broadcasting and Distribution,” SMPTE 1997; also *SMPTE Journal*, October 1998.

Todd, Craig C., “Loudness Uniformity and Dynamic Range Control for Digital Multichannel Audio Broadcasting,” IBC, 1995, pp. 145-150.

Todd, Craig C., Davidson, Grant A., Davis, Mark F., Fielder, Louis D., Link, Brian D., and Vernon, Steve, “AC-3: Flexible Perceptual Coding for Audio Transmission and Storage,” AES 96th Convention, February-March 1994, available on the Dolby web site at www.dolby.com/tech/.

Vernon, Steve, “Design and Implementation of AC-3 Coders,” *IEEE Transactions on Consumer Electronics*, Vol. 41, No. 3, August 1995.

“The Digital Audio Compression Standard-ATSC A/52,” available from the Dolby web site at www.dolby.com/tech/ or the ATSC web site at www.atsc.org.