Standards and Practices for Authoring Dolby® Digital and Dolby E Bitstreams

Issue 3

Dolby Laboratories, Inc.

Corporate Headquarters

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

European Headquarters

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

Table of Contents

1 Scope and Purpose

Dolby Laboratories has developed products and technologies that enable audio professionals to create their own Dolby Digital data streams, allowing greater control and precision. This document is intended to encourage the use of these tools by mixers, engineers, producers, and directors to master their own audio.

Authoring Dolby Digital bitstreams is a mastering process, and great care should be taken to properly select the metadata parameters that control the consumer's decoder or set-top box and affect the listening experience. It is therefore important to keep this process as close to the creation of the content as possible so that the intent of the artist is conveyed accurately to the consumer. The use of Dolby tools and technologies in the authoring process provides the artist and/or producer unprecedented control over how their work is experienced by the consumer.

2 Definitions

2.1 Dolby Digital

Dolby Digital (AC-3) is intended for the transmission of audio into the home through digital television broadcast (either high or standard definition), DVD, or other media. Dolby Digital can carry anywhere from a single channel of audio up to a full 5.1-channel program, including metadata. In both digital television and DVD, it is commonly used for the transmission of stereo as well as full 5.1 discrete audio programs. Dolby Digital is designed for maximum fidelity and space efficiency, and should only pass through one encode/decode cycle.

2.2 Dolby E

Dolby E is specifically intended for the distribution of multichannel audio within professional production and distribution environments. Any time prior to delivery to the consumer, Dolby E is the preferred method for distribution of multichannel/multi-program audio with video. Dolby E can carry up to eight discrete audio channels configured into any number of individual program configurations (including metadata for each) within an existing two-channel digital audio infrastructure. Unlike Dolby Digital, Dolby E can handle many encode/decode generations, and is synchronous with the video frame rate. Like Dolby Digital, Dolby E carries metadata for each individual audio program encoded within the data stream. The use of Dolby E allows the resulting audio data stream to be decoded, modified, and re-encoded with no audible degradation. As the Dolby E stream is synchronous to the video frame rate, it can be routed, switched, and edited in a professional broadcast environment.

2.3   Metadata

Metadata is additional control information that is carried along with the encoded audio program and provides essential information about the audio to a Dolby Digital decoder. Metadata provides many important functions including dynamic range control for less-than-ideal listening environments, level matching between programs, downmixing information for the reproduction of multichannel audio through fewer speaker channels, and other information. Metadata makes Dolby Digital a complete delivery system for audio, rather than just an audio compression system.

2.4   Integrated Receiver Decoders (IRD) and Set-Top Boxes (STB)

When receiving digital television broadcasts, whether via terrestrial, cable, or satellite transmission, either an integrated receiver or a set-top box is needed to separate the audio and video components from the carrier signal. Every STB has a built-in Dolby Digital decoder that supplies an analog stereo downmix of the program audio (either Lt/Rt or Lo/Ro, see Section 3.6 for more information). Some STBs may also offer a mono signal (derived from the Lo/Ro signal) modulated over an RF/antenna output. In addition to these outputs, a digital output is provided for connection to an external decoder.

2.5   Dolby Digital Surround EX

Dolby Digital Surround EX™ is an extension to the Dolby Digital 5.1 format. It was introduced to the movie-going public with *Star Wars:* Episode One—The Phantom Menace. Originally developed for theatres, this format has migrated into consumer products and media. Surround EX is primarily used for DVD soundtracks, although it may be incorporated into the broadcast chain at some point in the future.

The format itself is different from that found in a conventional 5.1 home theater environment. A back surround channel is added, creating a center surround channel between the left and right surround speaker channels. This additional "center" or back surround channel is achieved through a matrix encode of the three discrete surround channels (Ls, Bs, Rs) during the audio mastering process. This creates a stereo-compatible Lst/Rst (Left Surround Total, Right Surround Total) for the surround channels. In this way, a DVD released in the Surround EX format is compatible with all existing home theater configurations.

The metadata stream contains a specific parameter that can be flagged to indicate that the Dolby Digital audio stream is encoded in Surround EX. This metadata parameter is informational only, and simply allows those consumer decoders that are capable of decoding in the Surround EX format to switch automatically into this mode.

3   Perceptual Coding vs. Metadata

It is generally not necessary to monitor the Dolby Digital coding process from encode to decode simply to preview the effects of the coding algorithm. However, it is very important to monitor the effects of *metadata* on the source audio either through the encode/decode process or through the use of a DP570 Multichannel Audio Tool. The simple reason is that the Dolby Digital encoding process is specifically designed to be transparent, while metadata is intended to optimize the program audio for playback in a variety of home listening environments.

3.1   The Sound of Perceptual Coding as Implemented in Dolby Digital

Dolby Digital is a data reduction technology that achieves data compression through the removal of redundant audio information. The audio reproduced through a Dolby Digital decoder is not identical to the original source audio. During encoding, the Dolby Digital algorithm selects the portions of the audio that would not normally be heard by the human ear and removes them through a process known as *Perceptual Coding.* Perceptual coding uses the natural properties of the human ear to ignore signals masked by adjacent frequencies and the differences in level found within full bandwidth program audio. More information of the specifics of perceptual coding can be found in several publications on the Dolby website, www.dolby.com.

The Dolby Digital encoded/decoded signal sounds perceptually the same as the original audio signals. The masking properties of the human ear allow Dolby Digital to achieve a better than 15:1 compression ratio from original source digital audio with little or no perceived difference.

The ratio varies based on applicable sampling rates, Dolby Digital data rates, and bit resolution. Dolby Digital preserves the resolution of the source digital audio.

Note: Sample rate conversion in the DP569 Dolby Digital Multichannel Encoder, when enabled, limits the encoded signal to a maximum 20-bit resolution. However, the DP569 is capable of 24-bit resolution when the sample rate converter is disabled.

The Dolby Digital coding system is designed to preserve the fidelity of the original source audio, and nearly ten years of critical use by audio professionals in DVD, film, multimedia, and broadcast has proven it to be a reliable, accurate, and transparent coding system.

3.2   The Sound of Metadata

Unlike the nuts and bolts of the perceptual coding process, which is designed to be audibly identical to the source master, metadata is specifically intended to optimize

the sound of the decoded program based upon a consumer's listening conditions. It is therefore vitally important to monitor the effects of metadata while mastering a Dolby Digital data stream.

Metadata provides the tools necessary for audio programs to be reproduced accurately and artistically in many different listening situations from full-blown home theaters to in-flight entertainment, regardless of the number of speaker channels, quality of playback equipment, or relative ambient noise level. While an engineer or content producer takes great care in providing the highest quality audio possible within their program, they have no control over the vast array of consumer electronics or listening environments that will attempt to reproduce the original soundtrack. Metadata provides the engineer or content producer greater control over how their work is reproduced and enjoyed in almost every conceivable listening environment.

3.3   The Three Ds: Dialogue Normalization, Dynamic Range Control, and Downmixing

Metadata provides a number of key parameters that are specifically designed to control the sound of the program delivered to the consumer, depending upon selections made at the consumer's decoder that reflect their unique listening environment. These metadata parameters are known generally as Dialogue Normalization, Dynamic Range Control, and Downmixing.

Note: Rather than a single metadata parameter, the function of *Downmixing* within the consumer's Dolby Digital decoder is controlled by several specific metadata parameters, and, as with the other two Ds, care must be taken in monitoring and selecting these metadata parameters.

The engineer is ultimately responsible for optimizing the multichannel mix for best reproduction in the optimal listening environment so care should be taken to ensure that less optimal listening environments consisting of fewer speaker channels or high ambient noise levels are supported. For example, enjoyment of a DVD, game console, or digital television program should not be limited to only those consumers with full-blown home theater systems. Dolby Digital and metadata together provide a method to achieve the highest quality audio reproduction without compromising the integrity of the original encoded audio, regardless of the number of speaker channels, relative ambient noise levels, or quality of equipment in a playback system.

3.4   Dialogue Normalization

The Dialogue Normalization (also known as Dialogue Level or Dialnorm) parameter within the Dolby Digital stream provides a relative value to the home decoder or set-top box that adjusts the audio to a predetermined replay loudness level. This value aids in level matching between program content and media types (i.e., DVD, DTV,

DBS, etc.). Setting the Dialogue Normalization parameter is crucial to the proper operation of home decoders and provides three main functions:

1. In multimedia postproduction, the level-matching of dialnorm allows different content (i.e., titles, short movies, game replays, etc.) to be interspersed, maintaining the same comfortable listening level in the home. During the digital broadcast of a television program, dialnorm allows commercials, news breaks, and the like to be interspersed within the program so that the viewer doesn't have to constantly reach for the volume control during breaks.

2. Consumer requirements for home enjoyment of multichannel audio vary widely. Not everyone who owns a DVD player or receives a digital television broadcast listens in an ideal home theater environment; for example, many laptop computers offer DVD players and DTV receivers. The specifications, price, and capabilities of speakers, amplifiers, DVD players, etc., range from the basic to near-professional. Some care must be taken so that consumers of multichannel audio programs can be reasonably assured that the mixer's intent will be carried through their chosen playback system to their ears. In addition to level-matching between different program content and media, dialnorm ensures that the full dynamic range of the program can be reproduced without clipping the digital-to-analog (D/A) converters. As this level shift occurs well before the consumer's volume control knob, the decision on how loud to listen to a particular program is still up to the consumer.

3. A properly set dialnorm value provides the *null band* within the dynamic range profile (see Section 3.5) where the audio level is neither raised nor lowered. Without a properly set dialnorm parameter, reduced dynamic range listening modes may not have the intended effect of allowing softer portions of the signal to be audible while simultaneously reducing the decibel level of explosions, gunshots, or similar effects that may disturb others in the home.

Typically, consumers of a DVD movie or digital television program set their system's playback volume to a comfortable level centered around the level of dialogue within the program. This comfortable listening level depends on the consumer's taste and the intelligibility of the dialogue component. The parameter is called "Dialogue Normalization" because it uses the knowledge of the dialogue level within a program to make sure the consumer's comfortable listening level remains consistent between programs. This concept is derived from a standard practice in film mixing.

That is not to say that dialogue is required to set the Dialogue Normalization parameter. In music-only DVDs or television broadcasts, the dialnorm parameter can be thought of as the average volume level of the program.

Dialogue Normalization, in simple terms, is exactly the same as turning the volume down a bit on a consumer's home stereo. However, simply adjusting the volume on a home stereo provides none of the other advantages of dialnorm: providing a reference for reduced dynamic range listening conditions, accurate and musical dynamic range compression, and clipping protection prior to the D/A circuitry.

Dialogue Normalization in a Dolby Digital decoder cannot be defeated. Dialogue Normalization alone does not assert any compression or expansion on the program material but simply adjusts the audio to a standardized level (See ATSC A52 www.atsc.org/Standards/stan_rps.html).

That being said, Dialogue Normalization works in partnership with two other optionally assertable (or defeatable) parameters within the Dolby Digital bitstream, collectively called the Dynamic Range Control Profiles. The Dialogue Normalization parameter determines the area within the program audio where the Dynamic Range Profiles are inactive, setting a null band between the soft and loud portions of the program where no audio processing occurs and defining the upper and lower limits of the Dynamic Range Profiles.

Again, while the Dialogue Normalization parameter is required and not defeatable, the Dynamic Range Profiles are optional and can be turned off in a properly implemented Dolby Digital decoder. This means the choice to listen to a program at a reduced dynamic range (so as not to disturb the neighbors), or to the same program in all its full-volume, earth-shaking glory, is entirely up to the consumer.

3.5   Dynamic Range Control

Dynamic Range Control (also known as Dynamic Range Compression) within the Dolby Digital data stream consists of two profiles: Line Mode and RF Mode. These two profiles do not change the content of the encoded audio within the bitstream, but are used by the Dolby Digital decoder to adjust the extremes of the program material within the listening environment to account for those instances where it is preferable or necessary to listen to the program at a reduced dynamic range.

Line mode provides a moderate amount of compression when compared with RF mode, and also allows the user to adjust the low-level boost and high-level cut parameters within a home decoder when not downmixing. This adjustment or scaling of the boost and cut areas allows the consumer to customize the audio reproduction for their specific listening environment. To avoid clipping, the scaling feature is not available in certain downmixing situations. RF mode is designed for peak-limiting situations where the decoded program is intended for delivery through an RF input on a television, such as through the antenna output of a set-top box. The RF Mode Profile is also used for a common feature on consumer decoders known as "Midnight Mode," which provides enough dynamic range compression to ensure that an action movie or game won't wake up others in the home.

In contrast to compression as used on musical instruments in a recording studio to make them sound *punchy* and *fat*, dynamic range compression limits the softest and the loudest portions of an audio program to maintain a comfortable and intelligible listening level. For example, when enjoying a movie (on DVD, video, or broadcast TV) at lower volumes, the softer portions of the program (whispers, softer dialogue, etc.) are more difficult to hear, requiring greater volume, thereby making the louder portions (explosions, onscreen arguments, gunshots, etc.) too loud for comfortable

listening. Additionally, when enjoying a television broadcast or DVD in an environment with a high level of ambient noise, quieter portions of the program are drowned out by the ambient noise. When these "Dynamic Range Profiles" are asserted within the decoder, the decoder raises the level of the softer portions of the program and lowers the level of the louder portions, allowing the user to enjoy the movie without having to reach for the volume control constantly.
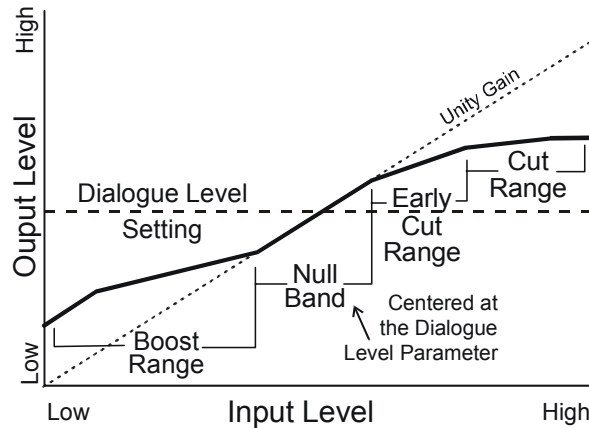

Figure 1 Dynamic Range Control within a Dolby Digital Decoder

There are two dynamic range metadata profiles to set within a Dolby Digital data stream. Each of these two profiles has six presets available to choose from, including None.

Music Light (No early cut range)
Max Boost: 12 dB (below –65 dB)
Boost Range: –65 dB to –41 dB (2:1 ratio)
Null Band Width: 20 dB (–41 dB to –21 dB)
Cut Range: –21 dB to +9 dB (2:1 ratio)

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

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

Film Standard
Max Boost: 6 dB (below –43 dB)
Boost Range: –43 dB to –31 dB (2:1 ratio)
Null Band Width: 5 dB (–31 dB to –26 dB)
Early Cut Range: –26 dB to –16 dB (2:1 ratio)
Cut Range: –16 dB to +4 dB (20:1 ratio)

Speech
Max Boost: 15 dB (below –50 dB)
Boost Range: –50 dB to –31 dB (5:1 ratio)
Null Band Width: 5 dB (–31 dB to –26 dB)
Early Cut Range: –26 dB to –16 dB (2:1 ratio)
Cut Range: –16 dB to +4 dB (20:1 ratio)

The details of the dynamic range characteristics for each preset are shown for reference only and are not adjustable, and, as mentioned previously, the dialogue normalization value determines the placement of the null band. Upon decoding, however, a customer has the option of scaling the amount of boost or cut, depending on the feature set available in their home decoder. Once again, the ability to scale only applies to Line mode.

Because of the relationship between dialogue normalization and dynamic range control, it is necessary to select the appropriate dialnorm value prior to previewing dynamic range profiles. As the amount of dynamic range compression used is ultimately selected by the consumer for their own specific listening needs, it is important to preview the dynamic range control profiles through a variety of compression settings before selecting one to include in the metadata stream.

3.6   Downmixing

Simply put, downmixing allows consumers to enjoy a 5.1-channel television broadcast, DVD, or game console without requiring a complete home theater setup. More technically, downmixing is a function of Dolby Digital that allows a multichannel program to be fully reproduced over fewer speaker channels than for which the program is optimally intended.

As with stereo mixing, where the mix is monitored in mono on occasion to maintain compatibility, multichannel audio mixing requires the engineer to reference the mix to fewer speaker channels to ensure compatibility for downmixing situations. In this way, Dolby Digital, using the metadata parameters that control downmixing, is an "Equal Opportunity Technology" in that every consumer who receives the Dolby Digital data stream can enjoy the best audio reproduction possible, irrespective of the number of channels in their playback system.

It is important to understand the output signals found on each piece of equipment that can receive a Dolby Digital program in the home (see Table 1). Set-top boxes (STB), such as those used for the reception of terrestrial, cable, or satellite digital television typically offer an analog mono signal modulated on the RF/Antenna output, a line level analog stereo signal, and an optical or coaxial digital output. DVD players offer an analog stereo signal as well as a digital output, and possibly six-channel analog outputs as well. Portable DVD players offer an analog stereo signal, headphone, and digital outputs. DVD players found in computers and game consoles offer a digital output and possibly six-channel analog outputs, as well as analog stereo and headphone outputs. 5.1-channel amplifiers, decoders, and receivers have six-channel analog outputs and possibly six speaker-level outputs.

In all cases, the analog stereo output is a downmixed version of the Dolby Digital data stream while the digital output carries the Dolby Digital data stream to a downstream decoder or integrated amplifier with Dolby Digital capability.

The analog stereo output of these units can be one of two different stereo downmixes. One is a stereo-compatible Dolby Surround downmix of the multichannel source program that is suitable for Dolby Surround Pro Logic decoding. This downmix is called left-total/right-total or Lt/Rt. The other type of downmix is a simple stereo representation (called left-only/right-only, or Lo/Ro) suitable for playback on a stereo hi-fi or via headphones, and from which a mono signal is derived for use on the RF/Antenna output from a set-top box. The difference between the two downmixes is how the surround channels are handled. The Lt/Rt downmix sums the surround

channels and adds them in phase to the left channel and out of phase to the right channel. This allows a Dolby Surround Pro Logic decoder to reconstruct the L/C/R/S channels for a Pro Logic home theater. The Lo/Ro downmix adds the left and right surround channels discretely to the left and right speaker channels. This preserves the stereo separation for stereo-only monitoring and provides a mono-compatible signal. The LFE channel is not included in either of these downmixes.

On most equipment, the consumer can, through the product's user interface, choose the appropriate downmix for their playback system. Certain metadata parameters allow the engineer to select how the stereo downmix is constructed and which downmix is preferred, although the Lt/Rt downmix is usually the default.

As previously mentioned, user adjustment of Dynamic Range Control (DRC) is limited while downmixing; the use of the stereo analog outputs of DVD players, set-top boxes, and game consoles are no different. Typically, the consumer is not able to adjust the cut or boost parameters when using the stereo outputs and the only DRC available is the selection of a "Midnight Mode" or the equivalent.

Some metadata parameters assist in achieving an appropriate downmix, helping to ensure that the intention of the engineer/content producer translates correctly across these environments. Specifically, metadata provides control over how certain channels are "folded" into the resulting downmix. As with DRC, downmixing is ultimately the choice of the consumer and dependent upon their unique listening environment.

While the engineer is tasked with optimizing the multichannel mix for reproduction in an ideal monitoring environment, it is important to reference the mix in downmixing conditions to ensure cross-platform compatibility and the proper selection of the downmixing metadata parameters. The many different consumer listening modes can be heard through front-panel selections on either the DP562 Reference Decoder or the DP570 Multichannel Audio Tool.

Table 1 Example of Consumer Products and Output Options

	Digital I/O	5.1-Channel Analog Outputs	Two-Channel Analog Outputs	RF Remodulated Output	Notes
5.1-Channel Amplifier	X	X			The regular A/V amp
5.1-Channel Decoder		X			
High-End DVD Player	X	X	X		
DVD Player	X		X		
PC	X	(X)	X		Includes games consoles
High-End Set-Top Box	X	X	X	X	Often HDTV
Set-Top Box	X		X	X	Usually SDTV
IDTV	X		X		TV set with an integrated digital TV tuner
High-End TV	X	X			Large screen TV with 5.1 speakers

4 DVD Authoring System Overview

4.1 DP569/DP562 with Dolby Recorder

This the most common setup for Dolby Digital encoding and decoding in use today. It provides a Dolby Digital bitstream that can be stored on a PC computer and used within any DVD authoring package as well as for reference decoding. It has also been used in mixing situations, since configuring an encoder back-to-back with a decoder was the only way to monitor the effects of metadata on program audio, until the introduction of the DP570 Multichannel Audio Tool.
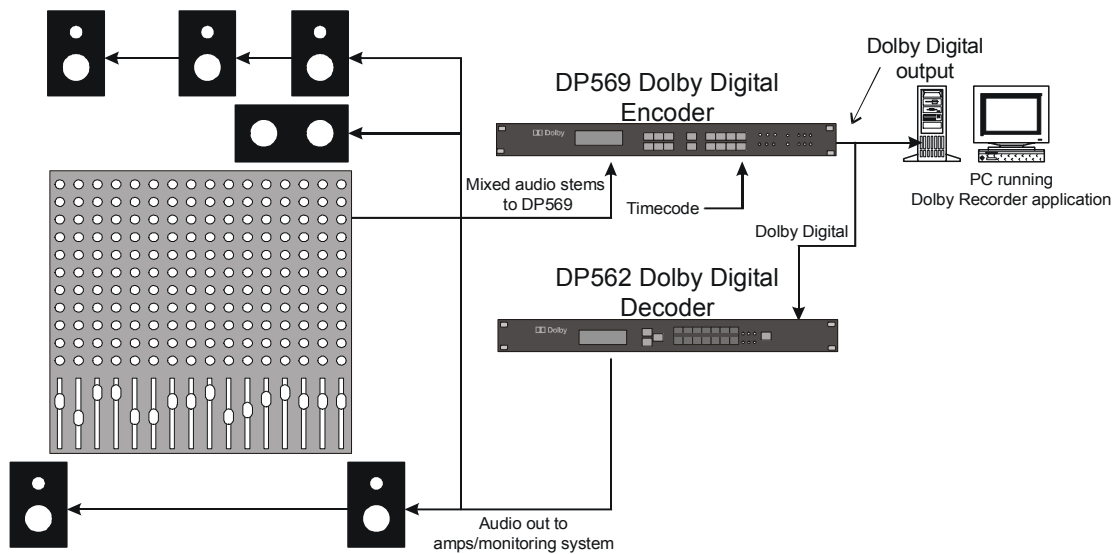
Figure 2 DP569/DP562 with Dolby Recorder

This method, while effective, suffers from two main drawbacks.

1. The resulting audio from the decoder is out of sync with the picture. The coding delay inherent in the encoder, coupled with the latency of the decoder, creates a minimum overall sync error of about 219 ms. The coding delay within Dolby Digital encoders is adjustable from a minimum of 187 ms up to a maximum of 450 ms, while Dolby Digital decoders, adhering to SMPTE 337M, add a latency of 32 ms on the decode side. While the resulting audio gives the engineer an accurate representation of the effects of the selected metadata parameters, the delays involved create a less-than-ideal monitoring situation.

2. This configuration provides no method to use common console functions such as Solo/PFL and Speaker Dim/Mute during mastering. As the monitoring system is connected to the output of the DP562 in this configuration, any solo or dim/mute functions existing on the console cannot be used.

These three issues notwithstanding, using an encoder and decoder back-to-back provides an accurate method of selecting and previewing metadata parameters for the creation of a Dolby Digital data stream, as well as creating the actual stream itself.

4.2   DP570/DP569 with Dolby Recorder

The preferred method of selecting, previewing, and authoring metadata for inclusion within a Dolby Digital data stream is through the use of the DP570 Multichannel Audio Tool paired with the DP569 Encoder. This configuration uses the DP570 for both the monitoring and authoring functions of metadata, and frees the DP569 to simply encode the audio and create the resulting Dolby Digital data stream with the selected metadata parameters.
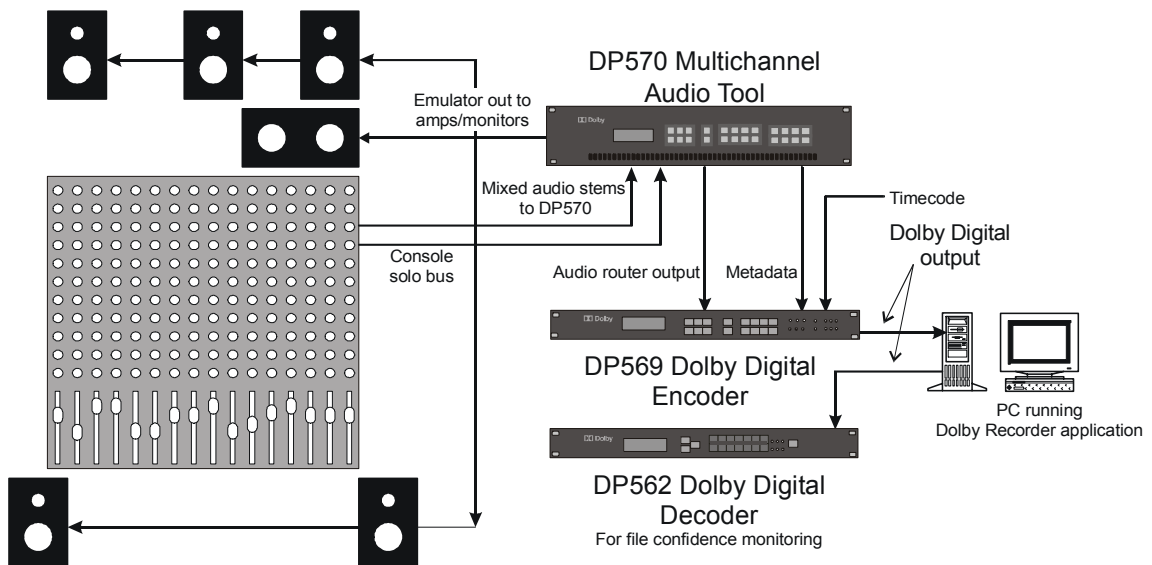
Figure 3 DP570/DP569 with Dolby Recorder

This configuration provides several key features.

1. The DP570 Multichannel Audio Tool provides the ability to monitor the effects of metadata in real time, without the delays found in an encode/decode configuration. This allows audio and picture to remain synchronous while selecting and previewing various metadata parameter settings.

2. The software remote and graphical user interface (GUI) provide a simple procedure for the measurement and selection of the Dialogue Normalization (dialnorm) parameter. As dialnorm is the single most important parameter in the metadata stream, this procedure helps eliminate the guesswork in setting this important value.

3. The metadata selections made are automatically forwarded to the DP569 encoder, limiting user input to a single, user-friendly device in the selection and preview of metadata.

4. The solo bus within the console is still active, and the DP570 provides dim and mute functions for audio production. Additionally, the DP570 provides a General Purpose Input/Output (GPI/O) connector for external control of key features within a console-driven hardware remote control. This allows many controls within the DP570 to be located close to the engineer and even integrated within the console itself.

4.3 Surround EX Encoding with DP570/DP569/EX-EU4 and Dolby Recorder

This configuration, a modification of the previous one, prepares, creates, and monitors a Dolby Digital bitstream encoded in Surround EX™ for DVD.
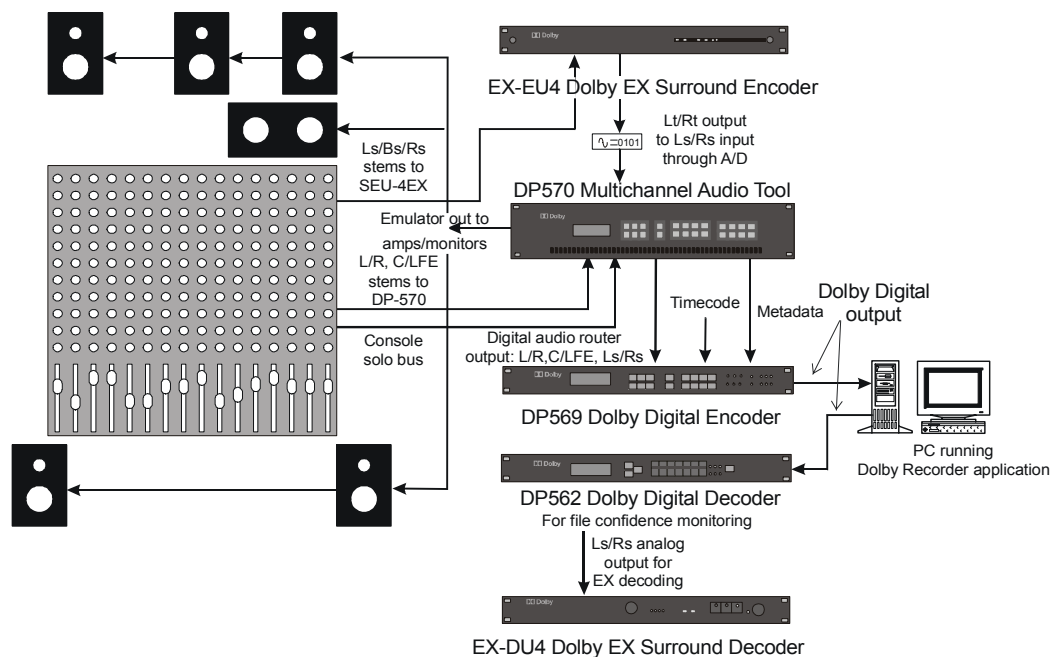
Figure 4 Surround EX Encoding with DP570/DP569/EX-EU4 and Dolby Recorder

As in the previous example, the console master functions are still active and the DP570 provides dim and mute functions for audio production.

The mixed "6.1" stems from the console are separated, and the surround channels (Ls, Bs, Rs) are sent to the EX-EU4 while the front channels (L, R, C, LFE) are sent directly to the DP570 (appropriate A/D or D/A conversion is implied). After encoding the three surround channels in real time as analog audio, the Lt/Rt output of the EX-EU4 is sent to the Ls/Rs inputs on the DP570 after A/D conversion. When the EX button is pressed on the DP570, the DP570 performs a Pro Logic decode on the Ls/Rs inputs and sends the decoded signals to the appropriate speaker channels for production monitoring.

Standard coding delays are present within the DP569 Dolby Digital Encoder, and standard latencies are present during confidence checking through the DP562 decoder.

In this example, the analog outputs of the DP562 are used to relay the Ls/Rs channels to the EX-DU4 Surround EX Decoder.

5 DTV Authoring Overview

5.1 Metadata in Master Control

Monitoring the effects of metadata on program content in a master control situation can be simply achieved through the use of the DP570.
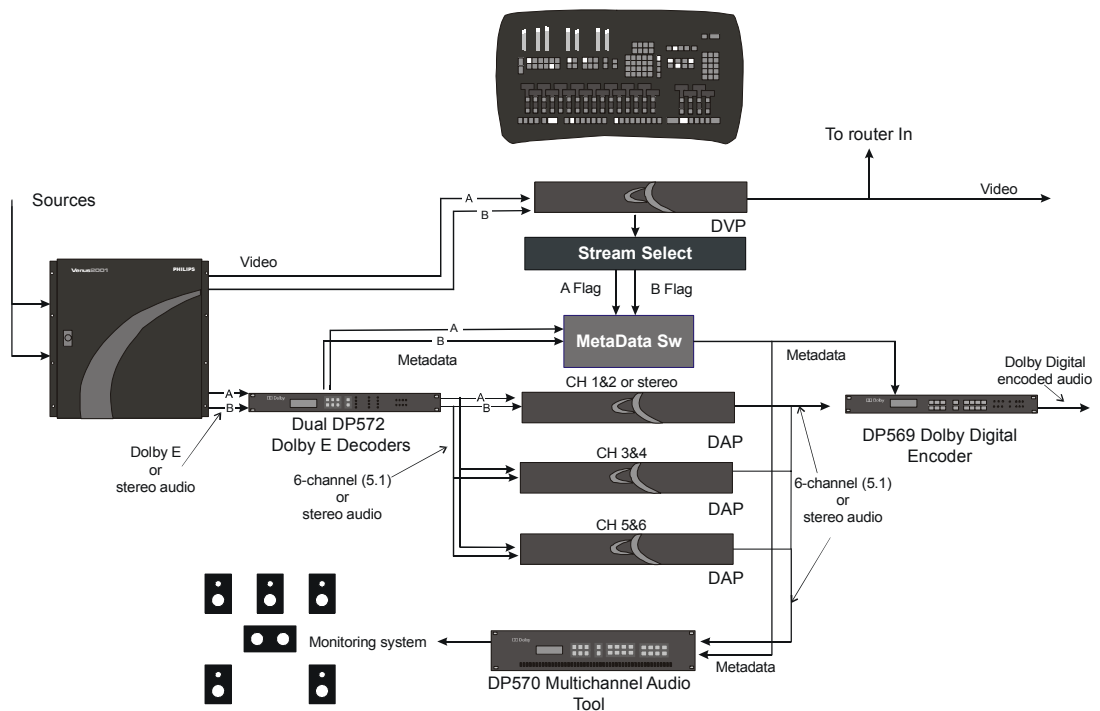
Figure 2 Multichannel Monitoring in Master Control for Digital Television

This block diagram shows a simplified master control switch between two sources. Prior to entering master control, the incoming Dolby E streams are decoded back to baseband PCM audio. The metadata for each incoming Dolby E stream is sent through a serial metadata switcher whose output feeds a downstream DP569 Dolby Digital Encoder for transmission, as well as a DP570 Multichannel Audio Tool in the monitor chain. The DP570 controls the monitoring environment and receives the appropriate audio feeds from either of the two sources in this example.

5.2   DP570/DP571 for Digital Television Distribution

The preferred method of selecting, previewing, and authoring metadata for inclusion within a Dolby E data stream is through the use of the DP570 Multichannel Audio Tool paired with the DP571 Dolby E Encoder. This configuration uses the DP570 for both the monitoring and authoring functions of metadata, and frees the DP571 to simply encode the audio and create the resulting Dolby E data stream with the selected metadata parameters.
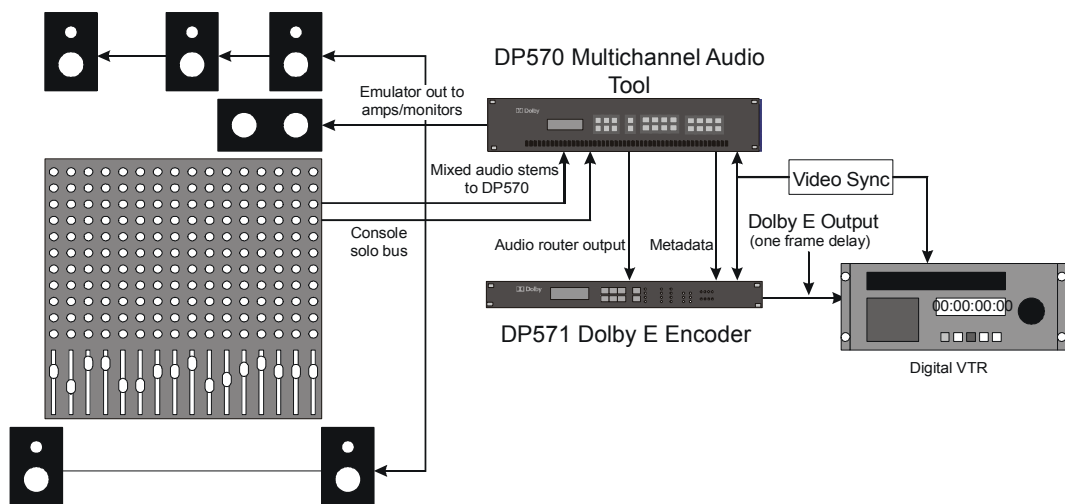
Figure 3 DP570/DP571 for Digital Television Distribution

The output of the DP571 is delayed one video frame from reference. Upon decode of the Dolby E stream, the DP572 Dolby E Decoder adds another frame of delay from the reference.

This configuration provides several key features.

1. The DP570 Multichannel Audio Tool provides the ability to monitor the effects of metadata in real time, without the latencies (and expense) inherent in using a Dolby Digital encoder and decoder in a back-to-back configuration. This allows audio and picture to remain synchronous while selecting and previewing various metadata parameter settings.

2. The software remote and GUI provide a simple and automatic procedure for the measurement and selection of the dialnorm parameter. As dialnorm is the single most important parameter in the metadata stream, this procedure eliminates the guesswork in setting this important value.

3. The metadata selections made are automatically forwarded to the DP571 encoder, limiting user input to a single, user-friendly device.

The solo bus within the console is still active and the DP570 provides dim and mute functions for audio production. Additionally, the DP570 provides a GPI/O connector for external control of key features within a console-driven hardware remote control. This allows many controls within the DP570 to be located close to the engineer and even integrated within the console itself.

6   Frequently Asked Questions

Aren't Dolby Digital and "5.1" the same thing?

No. A stream encoded in Dolby Digital can carry any number of channels, from a minimum of one mono signal to a maximum of six channels in a 5.1-channel home theater configuration. In digital television broadcast, Dolby Digital is commonly used in a two-channel stereo configuration. The flexibility of Dolby Digital as an encoding technology to carry anywhere from mono to multichannel surround programs makes it the ideal tool to carry audio in digital media.

Aren't Surround EX and "6.1" (or even "7.1") the same thing?

No. Dolby Digital Surround EX™ was first introduced to the public with the release of *Star Wars:* Episode One—The Phantom Menace. Since then, many more films have been released in theatres in the Surround EX format. Now, Surround EX is appearing on DVDs and consumer electronic equipment for use in the home.

The Surround EX format is different from conventional 5.1 in that an extra back surround channel is added between the left and right surround channels. To maintain compatibility, this back surround channel is matrix-encoded with the left/right surround channels to create a stereo-compatible surround signal. Consumers without Surround EX capability receive stereo surround channels, while those with Surround EX decoding capability hear three surround channels. The Surround EX format is only valid on source signals with at least two surround channels.

How do I set up the speakers on my mix stage or production studio for Dolby Digital or Dolby E authoring?

Since both Dolby Digital and Dolby E are discrete encoding processes (i.e., speaker channel separation is maintained throughout the process and no channels are "matrixed"), what is heard on the mix stage is what the consumer will hear at home, provided the listening environment is properly calibrated. While there is no specific standard for sound pressure level (SPL) during mastering for DVD or digital television broadcast, it is important that each speaker is adjusted to deliver the same SPL at the mixing position. This can be achieved by measuring the SPL with a meter at the mixing position and making appropriate adjustments while playing pink noise through each speaker individually. Some mixing stages are set to 85 dB SPL at each speaker, while others may be set somewhat lower. The important thing is to make sure that each speaker is set to the same level at the mixing position.

The subwoofer speaker should be calibrated with 10 dB of in-band gain over the center channel from 25 to 120 Hz when measured with a real-time frequency analyzer. This would equate to about 91 dBC when measured with a properly calibrated SPL meter.

Additionally, it is important that each of the main channels be equidistant from the mixing position. If this is not possible, the DP570 Multichannel Audio Tool has built-in delays that can be added to each channel as needed. Please refer to the User's Manual for more comprehensive information.

I am recording and mixing a music-only program and I don't want the consumer to assert any compression on the audio. What should I do?

While it is often preferred to monitor and select the dynamic range control parameters for the DVD release or broadcast of a feature film, despite the user-selectable nature of these parameters they still may not be desired for music-only programs. In these instances, selecting "None" as the Dynamic Range profile can effectively negate the dynamic range control parameters. While RF modes will still be active for certain specific purposes (downmix clipping protection) and the proper setting of the dialnorm parameter is still critical, setting the profile to "None" prevents the consumer from asserting any user-selectable dynamic range control.

I'm mixing a film for digital television broadcast. Should I include a Dolby Digital stream of my final mix along with the discrete tracks?

It depends. Dolby Digital is designed as a one-time decodable technology. If the program content is to be simply streamed out through a broadcast encoder, including a Dolby Digital stream on the audio tracks of your digital videotape is acceptable. This ensures that the metadata parameters you select will be transported to the consumer's home decoder, and your content will be heard as you intended.

However, if the program content will be decoded and further processed, edited, or sweetened, this would be a distribution process and therefore require a different encoding technology. The accepted coding technology for the distribution of multichannel audio for digital television broadcast is Dolby E. If Dolby E equipment is not available, the best choice for the distribution of multichannel audio for DTV is the original source audio recorded in baseband PCM format on digital multitrack media.

Can I use a 384/448 kbps data rate for stereo?

Yes. All Dolby Digital decoders support up to at least 448 kbps (DVD and DVB max spec) data rate irrespective of the number of channels encoded within the Dolby Digital data stream. The final data rate used is a production choice depending on many factors, including number of audio tracks, video data rates, and available space for the audio on a DVD or within the transmission stream, among others. While 192 kbps is a common rate for the production of 2/0 (stereo) Dolby Digital data streams, there is nothing preventing the use of higher data rates, if desired.

How can I set the DP562 decoder to emulate the most common dynamic range control settings on consumer decoders?

The front panel of the DP562 has four buttons that deal with the dynamic range control settings: **None**, **Custom**, **Line**, and **RF**. **None** is a professional set-up mode that defeats both dynamic range control and dialnorm, and is only used for installation and signal tests. Once a specific dynamic range profile has been selected, simply pressing **RF** emulates the selection of "late night mode" in a consumer decoder, while **Line** mode emulates a more moderate dynamic range control effect that a consumer may select. **Line** mode within the DP562 decoder also allows limited scalability when not downmixing.

Custom is commonly used to turn dynamic range control off while still asserting the dialnorm parameter, which emulates the consumer choosing to listen in a full dynamic range mode. Some high-end consumer decoders offer a scalable boost and cut, and Custom mode can be tailored to reflect this as well.

Note: These modes are also included in the DP570 Multichannel Audio Tool.

I'm using a consumer decoder to monitor my Dolby Digital stream. How do I set it up so that I can use it for mastering my Dolby Digital stream?

While it is preferable to use a reference decoder such as the DP562 for mastering purposes, it is possible, albeit unwieldy, to use a consumer decoder.

Consumer decoders do not offer easy access to the many downmix and dynamic range options that may be available to the consumer at home. In addition, not all consumer decoders offer every feature that is available in Dolby Digital technology. A reference decoder like the DP562 has all the necessary controls for mastering on the front panel to facilitate changing downmixing conditions and dynamic range profiles quickly.

If budget constraints require the use of a consumer decoder, it is suggested that the engineer refer to the operation manual of the device for specific information.

Why is it called a Low-Frequency Effects (LFE) channel? Isn't it just the subwoofer?

Actually, LFE and subwoofer are two very different things.

Typically, the term "subwoofer" refers to a speaker that reproduces very low-frequency information that the main channel speakers (however many there are) are incapable of reproducing. Low-frequency sounds that are normally found on the main audio channels are directed to the subwoofer speaker for added punch in the low range. In this fashion, a subwoofer acts as a complement to extend the range of the main speakers, which may find it difficult, if not impossible, to reproduce these low frequencies.

The LFE channel in contrast, is specifically produced with low-frequency information exclusive to this channel. Content producers have specifically created this LFE channel to include extra sound-effects information for emphasis in effects, such as explosions, crashes, gunshots, and such.

In a consumer playback environment, often low-frequency information from the main channels is redirected to the subwoofer speaker and added to the existing LFE channel, if any. If there is no LFE channel present in the Dolby Digital stream, the subwoofer speaker contains the low frequencies redirected from the main speakers. The capability to redirect low-frequency information is called Bass Management, and is present in all Dolby Digital decoders. The LFE channel can also be redirected to the main speaker channels through the use of Bass Management, for those playback environments lacking a subwoofer speaker.

When producing a music-only multichannel program, the entire musical contents of the program should be placed in the main channels, just as they are in stereo music recordings. The LFE channel should only be used when the bass levels are so high as to require a substantial decrease in overall program volume to accommodate them, such as might occur with the cannon shots in Tchaikovsky's "1812 Overture."

Just as with any full-bandwidth audio signal delivered to consumers via DVD, CD, or DTV, it is the job of the playback system to get the most out of the signal. Consumer Dolby Digital systems that use a combination of smaller speakers and subwoofers incorporate bass management to ensure that the speakers most able to reproduce the bass are used to the best advantage. There is no need to tailor the frequency content of the program in anticipation of the many different playback systems it may encounter in the home.

Remember: just because the "point-one" channel is there does not mean it must be used. It is perfectly acceptable to create a five-channel Dolby Digital stream (encoding in 3/2 mode, as opposed to 3/2L mode) without an LFE channel present.

Why do I need the Dolby Recorder?

The Dolby Recorder program is a Microsoft Windows-based application that allows a Dolby Digital data stream to be recorded onto a computer hard drive. It creates a file with the .ac3 file extension that can then be ported over to a DVD authoring system and married to the encoded video content.

The DP569 Encoder outputs a Dolby Digital data stream carried within the envelope of an AES/EBU digital audio pair, although the actual area taken up by the Dolby Digital data is only a fraction of the AES/EBU space. The rest of the unused area is filled out with zeros. The AES/EBU envelope allows the Dolby Digital data to be routed and stored using much of the same equipment that passes AES/EBU digital audio. However, for a DVD authoring system to use this data stream, the excess zeros must be stripped off. In simplest terms, this is what the Dolby Recorder application does.

How much space does an *.ac3 file consume on my hard drive?

It is simple to calculate the needed disk space for a *.ac3 file. First, determine the number of seconds in the program. In this example, a 90-minute program and a Dolby Digital data rate of 448 kbps is used:

5,400 sec. × 448,000 bits per sec. [Dolby Digital data rate] = 2.4192 gigabits

2.4192 gigabits / 8 [for byte conversion] = 302.4 megabytes

I'm not using Dolby Recorder. Can I save my Dolby Digital data stream to a Digital Audio Tape (DAT)?

While both Dolby Digital and Dolby E data streams can be recorded and routed much like regular PCM digital audio, the encoded signal is *data* and specific precautions must be taken to ensure that it survives any archival process. Any reclocking or resampling of the data stream destroys it and renders it irrecoverable. However, both Dolby Digital and Dolby E data fit nicely on the digital audio tracks of many digital video recorders. Additionally, Dolby Digital can be recorded onto the audio tracks of many digital multitracks. Most DAT machines do not have the capability of recording data rather than audio, although some professional units offer this feature. Generally, a DAT is not a good medium for the archiving of Dolby Digital data streams. Whenever data is recorded onto a DAT, the possibility of undecoded data being played out of the unit at maximum volume (0 dB full scale), and possibly damaging speakers and ears, exists if the DAT machine mistakenly interprets the data as audio or in any way corrupts the data stream.

Can I record Dolby E onto a DAT?

For Dolby E to maintain coherence with video frame boundaries, and thus be able to be routed and switched much like video signals, it is usually necessary for the archival medium to be referenced to video.

However, with the introduction of the DP583 Dolby Frame Synchronizer, content producers now have the ability to archive Dolby E data streams to non-video based media. The DP583 provides the ability to re-clock the Dolby E stream to a house video reference, as well as reclocking Dolby Digital and baseband PCM audio to a house sample clock reference. As with other "audio only" mediums, care must be taken so that the DAT used to archive the Dolby E data stream is not mistakenly played back in an audio system.

With the delays involved in using Dolby E, how should I compensate to maintain lip-sync on playback of a DTV program?

Dolby E delays the audio a single video frame based on the video reference for each encode or decode cycle. These delays are necessary to process the audio into the data stream and vice versa. There are two methods to compensate for these delays:

1. Record the Dolby E stream so audio and picture are in sync on tape. This requires advancing the source tracks one frame before encoding to compensate for the one frame of delay in the Dolby E encoding process. This method maintains audio/picture sync on the recording media to facilitate editing. Upon decode of the Dolby E data, the video will have to be delayed one frame to compensate for the one frame of delay in the decoding process.

2. Record the Dolby E data so the decoded Dolby E data stream will be in sync with video upon decoding of the data stream.

When using Dolby E with standard definition video, this requires advancing the audio two frames prior to encoding. This compensates for the single-frame delay during the encode cycle and the additional single-frame delay on decode. This method is useful to ensure that there will be no additional video delays or other processing necessary upon decoding of the data stream, since the audio will be in sync once decoded. This method is not recommended if the source program will be edited prior to decoding of the Dolby E data stream.

In the case of Dolby E with high-definition video, HD video machines have built-in means of compensating for the one-frame Dolby E encoding delay. Upon playback of the HD tape, the video processing built into the unit compensates for the one-frame Dolby E decoding delay, thereby maintaining sync with the picture during editing and upon decoding of the Dolby E data stream.

How do I compensate for encoding delays and decoding latencies when authoring a Dolby Digital data stream for DVD?

A Dolby Digital data stream can carry time stamp information within each Dolby Digital synchronization frame. These time stamps originate from the master timecode referenced to the video content. When a DVD is authored, the authoring system looks at the time stamps encoded within the Dolby Digital data stream and matches the timecode numbers between the audio and video components.

When mastering a Dolby Digital data stream for DVD, be sure to include the appropriate timecode signal to maintain sync with picture within the authoring system.

Does Dolby E replace Dolby Digital?

No. Dolby E is an encoding technology used only for professional distribution of audio, such as from television networks to affiliates through satellite communication or via hard media like digital videocassettes. Dolby E is specifically designed to allow numerous generations of encode/decode cycles that are sometimes necessary for the production and distribution of audio in a broadcast environment. Dolby Digital is designed to deliver digital audio to the home through digital television broadcast, DVD, or other media. As it is optimized for high quality at low data rates, Dolby Digital encoding should be performed after all final production decisions have been made and the next decode step would be in the consumer's home.

Can I buy a consumer version Dolby E decoder for my home theater?

No. Dolby E is a professional technology and is not licensed for consumer use.

How do I get from Dolby E to Dolby Digital?

To create a Dolby Digital data stream from a Dolby E stream, it is necessary to decode the source Dolby E stream to baseband PCM digital audio and then encode the selected digital audio program using a Dolby Digital encoder.

Where can I learn more about Dolby technologies?

See www.dolby.com for more details about our technologies and products. You can also find technical information about AC-3 on the Advanced Television System Committee's website, www.atsc.org.