Dolby Digital Broadcast Implementation Guidelines

Dolby

Dolby Laboratories Inc

www.dolby.com

United States

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

United Kingdom

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

Dolby and the double-D symbol are registered trademarks of Dolby Laboratories.
©1998 Dolby Laboratories Inc. All rights reserved.

Issue 02
S98/11964/12240

Table of Contents

List of Figures v

List of Tables vi

Chapter 1 – Introduction

1.1 Audio Coding Standards for Digital Television (DTV) 2
1.2 System Concepts 2
1.3 Metadata 3
1.4 Audio Coding in the Digital Infrastructure 3
1.5 Dolby E for Distribution 4
1.6 The Future Infrastructure 4
1.7 Evolving the Infrastructure 4

Chapter 2 – Program Emission

2.1 The ATSC Audio System 7
2.2 The Consumer Receiver 10
2.2.1 Definitions and Performance 10
2.2.2 The Dolby Digital Decoder 10

Chapter 3 – Program Contribution and Distribution

3.1 5.1-Channel Audio 13
3.1.1 Plant Infrastructure 13
3.1.2 Dolby E Data Rate Reduction 13
3.1.3 Generating Metadata 17
3.1.4 Equipment Requirements 18
3.2 A Near-Term Solution 19
3.2.1 Dolby Surround in Broadcast 19
3.2.2 Dolby Surround Principles of Operation 23
3.2.3 Downmixing 5.1-Channel Audio 26

Chapter 4 – Multichannel Monitoring

4.1 Production Environment 31
4.2 Incoming Quality Control 31
4.3 Master Control 32

Chapter 5 – Professional Products

5.1 Dolby Surround 33
5.1.1 Dolby SEU4 Encoder 33
5.1.2 Dolby SDU4 Decoder 33
5.2 Dolby Digital 33
5.2.1 Professional Encoders 33
5.2.2 Professional Decoders 34

Chapter 6 – Miscellaneous Information

6.1 Technical Assistance 35
6.2 Contacting Dolby Laboratories 35
6.3 Trademark Usage 37

Appendix A – Basics of Perceptual Coding ..... A - 1

Appendix B – Metadata

B.1 Metadata ..... B - 1
B.1.1 *Dialnorm* - Dialogue Normalization ..... B - 1
B.1.2 *Dynrng* - Dynamic Range Compression ..... B - 2
B.1.3 *Compr* - Compression Control ..... B - 2
B.2 Overload Protection via dynrng and compr ..... B - 3
B.3 Likely Values for dialnorm ..... B - 3
B.4 Choosing dialnorm for a Program Item ..... B - 3
B.4.1 Speech ..... B - 3
B.4.2 Music ..... B - 4

Appendix C – Dolby Digital Encoding Parameters ..... C - 1

Appendix D – Bibliography ..... D - 1

List of Figures

2-1 Dolby Digital in the Emissions Stage 7

3-1 Original Production Process – Contribution Stage 14

3-2 Post-Production Process – Contribution Stage 15

3-3 First Stage Program Distribution – Assembly or Integration 16

3-4 Second Stage Program Distribution – Master Control 17

3-5 Original Production Process – Two-Channel Analog or Digital System 20

3-6 Post-Production Process – Contribution Stage 21

3-7 Packaging or Integration at the Affiliate – Distribution Stage 22

3-8 Two-Channel Analog Master Control – Distribution Stage 23

3-9 Dolby Surround Encoder 24

3-10 Dolby Surround Pro Logic Decoder Block Diagram 25

3-11 Dolby Surround (Lt/Rt) Compatible Downmix Signal Flow 26

3-12 Stereo (Lo/Ro) Downmix Signal Flow 27

3-13 Encode/Decode Downmix Signal Flow 29

A-1 The Effect of Masking ..... A - 1

A-2 Coding (quantization) Noise Below the Masking Curve ..... A - 2

List of Tables

3-1 Center Downmix Levels 28
3-2 Surround Downmix Levels 28

Chapter 1 – Introduction

It is the best of times for broadcast television which, like other consumer media, is moving from analog to digital program delivery. Digital broadcast technology allows the delivery of more programs, higher definition pictures, and discrete multichannel Dolby Digital audio.

Unfortunately, it is also the worst of times. To fully transition to DTV with the highest quality multichannel sound, the entire broadcast plant infrastructure may have to be upgraded and rebuilt. The exact form of the new infrastructure is unknown. Ideally, a number of small evolutionary steps will be taken to gently morph existing infrastructure into a new form that meets future requirements. Realistically, however, one or more large steps will probably be necessary.

Broadcasters are urgently seeking information about how to make the most practical transition to the new infrastructure. This document is an attempt to provide some of the needed information. Nearly anyone involved with the broadcast transition to digital television, or in the production of program material that will be broadcast over digital television, will find some of this information useful.

Dolby Laboratories, an innovative force in audio technology and expertise, has guided the movie industry from limited bandwidth monophonic sound, through a stage of matrixed surround sound, into the current age of Dolby Digital's discrete multichannel audio. Dolby contributed to the birth of the home theater industry by making Dolby Surround technology viable for domestic applications. Dolby Digital technology is currently supplying even higher quality home theater audio through the laser disc (LD) and DVD media. With the advent of the Advanced Television Systems Committee (ATSC) Digital Television (DTV) Standard, and similar standards in the cable and DBS industries, broadcasters will soon deliver high quality Dolby Digital entertainment experience to their own audiences.

The introduction of Dolby Digital has widespread ramifications to the fundamental infrastructure of the broadcast industry. Dolby Laboratories has a long history of introducing new audio technologies to the entertainment industries and providing support during and after the critical transition period. To implement the new audio capabilities offered by Dolby Digital, Dolby is committed to helping broadcasters evolve their infrastructure in a practical and cost-effective manner by:

- providing information such as that found in this manual;
- working with individual broadcasters, networks, and industry committees;
- developing ancillary technologies and standards;
- providing useful equipment and software.

Dolby's efforts to support the broadcast industry implementation of multichannel Dolby Digital are in the early phases. We don't have all the answers, but are hard at work trying to find them. This document represents the current state of knowledge and will continue to evolve as we, as well as others, research and develop additional techniques, equipment, standards, and operating practices. We welcome your input. Please point out where you think we have gone wrong, or where additional efforts are needed. Comments and inputs may be delivered by e-mail, to the following address: hdtv@dolby.com, or via facsimile to the attention of DTV at 415-863-1373. For further contact information, please refer to *Section 6.2 - Contacting Dolby Laboratories*.

1.1 Audio Coding Standards for Digital Television (DTV)

Virtually all DTV delivery in the United States will make use of the Dolby Digital audio coding system (also known as either AC-3 or Dolby D). Dolby Digital is the standard adopted by both the ATSC, in their A/52 document, *Digital Audio Compression Standard (AC-3)*, and by the International Telecommunications Union Radio Communication Bureau, in their Recommendation BS.1196. The U.S. FCC rules specify that U.S. terrestrial broadcasters use the ATSC A/53, *Digital Television Standard*, which specifies the use of ATSC A/52 (AC-3) audio. The cable TV industry also specifies the use of ATSC A/52 audio in the standards issued by the Society of Cable Telecommunications Engineers (SCTE). An increasing number of the direct broadcast satellite providers are beginning to use Dolby Digital audio. All of the ATSC Standards, as well as the informative A/54 document, *Guide to the Use of the ATSC Digital Television System*, are available from the ATSC web site at www.atsc.org.

1.2 System Concepts

The broadcast infrastructure extends from original production capture of program elements all the way to reproduction by the consumer. The signal path can be broken down into smaller, logical elements. For example, production is separate from delivery. Delivery can be broken into distribution and emission phases. Distribution is the delivery of essentially finished works from the output of the production or post-production process, to the point of emission. Emission, the final stage of delivery to the consumer, is the phase that uses the Dolby Digital coding system. The contribution phase is the delivery of program elements to the production or post-production process. Contribution, distribution, and emission are used in this document to describe various portions of the broadcast infrastructure; refer back to this section to clarify these essential definitions. Some ambiguity among these terms is inevitable due to the functional overlap between phases. For example, an emission signal (intended for delivery to the consumer) is sometimes used in distribution and then passed directly to the consumer. Certain applications use a distribution signal path for the contribution of program content.

1.3 Metadata

The Dolby Digital coding system offers several unique features that solve some of the practical problems in the delivery of high quality, versatile audio to a large and diverse audience. Some of these features are realized by using auxiliary information included inside the Dolby Digital emission bitstream to control aspects of the decoding and reproduction of the audio. This control information, called metadata (meaning data about the data), is not part of the audio signal content but affects decoding and audio reproduction. Since this audio metadata affects the consumer's experience of the program audio, it is highly desirable that the audio metadata be generated (or authored) during the post-production process where its effect on the audio can be properly evaluated. For this to be practical, the infrastructure must be capable of transmitting both the multichannel audio program and metadata through the entire distribution and emission phases. Metadata does not affect the contribution phase, which occurs prior to post-production. For further information on metadata, please refer to *Appendix A – Basics of Perceptual Coding*.

1.4 Audio Coding in the Digital Infrastructure

There is a strong trend towards the use of digital audio in many areas of the broadcast infrastructure. All new tape and disc storage media store digital audio and all new transmission links send digital audio. Although the most common representation of digital audio uses linear PCM (pulse code modulation), certain phases within the broadcast infrastructure require a reduction in the number of bits required to represent the digital audio program. The Dolby Digital coding system reduces the data rate of a 5.1-channel program from 6 Mb/s (6 channels, 20 bits, 48 kHz), down to 384 kb/s (a reduction ratio of 15 to 1). Data rate reduction is essential in emission, useful in distribution, and occasionally useful in contribution.

Dolby Digital is intended for use in the final emission link to the consumer, where operation at a reduced data rate is needed and additional coding generations are not normally necessary. Higher quality coding is essential in distribution because there may be many generations of encoding/decoding and the signal may be processed between generations. While a decoded signal may sound fine, certain types of signal processing may reveal coding artifacts that are otherwise inaudible. A distribution coder should have a much higher coding margin (the difference between the level of coding artifacts and their audibility) than an emissions coder.

Dolby Digital can be used in distribution applications and can provide distribution quality if operated at a higher data rate than is used for emission. Tests conducted at Dolby, and verified by some broadcasters, have shown that Dolby Digital is transparent after ten encode/decode cycles when operated at a data rate of 384 kb/s per two channels. In the 5.1-channel mode, a few encode/decode generations are possible by using a data rate of 640 kb/s. Some applications, such as laser disc (LD) mastering, record the Dolby Digital bitstream onto the digital audio track of a VTR. There are, however, some practical limitations to the use of Dolby Digital in the distribution phase. Dolby Digital encoders impose significant audio delays (70 - 160 ms) that require compensation. A signal in the encoded domain cannot generally be edited or switched in combination with the picture without mutes occurring in the final decoded output. The signal must instead be decoded to PCM, then edited or switched, and re-encoded.

1.5 Dolby E for Distribution

Dolby Laboratories will soon introduce a new coding technology designed specifically for distribution applications. The Dolby E-type audio coding technology will allow much of the existing two-channel linear PCM digital audio infrastructure (VTRs, routers, embedders, etc.) to carry multichannel audio. A single Dolby E bitstream will mimic a two-channel 20 (or 16) bit 48 kHz PCM audio signal and carry up to eight discrete audio channels. Existing editing and switching practices may be performed on the encoded Dolby E bitstream without causing mutes or other audible disturbances when the bitstream is decoded back to baseband PCM. Dolby E is intended to facilitate the transition from two- to 5.1-channel audio without requiring complete replacement of the existing infrastructure. The Dolby E system will also convey the audio metadata throughout the distribution system. Additional details on the Dolby E system and product availability will become available during 1998 and early 1999.

1.6 The Future Infrastructure

Dolby envisions that the future infrastructure will utilize linear PCM audio in production and post-production. Certain program elements (such as those arriving via satellite or telco links), however, may pass through a small number of generations of professional quality audio coding, such as Dolby E. Any audio processing manipulation, including the generation of the audio metadata that affects the final reproduction at the Dolby Digital decoder, should be done in post-production where the subjective effects can be properly monitored and evaluated. The remainder of the distribution and emission signal paths should, ideally, transparently deliver the audio experience designed in post-production to the home listener.

Distribution begins at the output of the post-production process, where the audio will be encoded into the Dolby E format. The signal can remain in the Dolby E format through tape generations, satellite links, signal routing, and content editing. When the audio signal must undergo additional processing, such as cross-fades to other programs and additional sweetening, the signal will be decoded to baseband LPCM and then re-encoded into Dolby E. While the Dolby E system provides sufficient audio quality to survive a large number of concatenated encode/decode generations, the fact that many typical signal manipulations can be performed without decoding minimizes the actual number of coding generations required. At the end of the distribution chain, the Dolby E encoded signal will be converted to the Dolby Digital (Dolby D) format for final delivery to the consumer at the low emission data rate.

1.7 Evolving the Infrastructure

Many of today's broadcast television stations using a two-channel analog audio plant can easily jump into DTV broadcasting by feeding their two-channel signal into the analog inputs of the Dolby DP567 two-channel broadcast audio encoder. The main encoded digital output connects to the video encoder/multiplexing equipment using an industry standard interface employing a 75 ohm unbalanced (BNC) connection. Since a two-channel signal can convey Dolby Surround matrix encoded multichannel audio, these stations can easily deliver a form of multichannel audio to their early DTV audience.

In order to implement discrete multichannel audio, a digital plant with at least two channels is required. While a multichannel PCM audio plant can be implemented, the Dolby E distribution coder enables a two-channel digital plant to work with discrete multichannel digital audio. Dolby recommends that broadcasters currently using an analog audio plant transition to a digital audio plant equipped with at least two AES-3 channel pairs (four PCM audio channels). This will provide compatibility with the equipment used by some broadcasters and most network centers, and with the recording capability of many digital VTRs. These two PCM stereo pairs will be able to carry two Dolby E streams for a total capacity of sixteen audio channels. This should be sufficient for many years to come.

Chapter 2 – Program Emission

2.1 The ATSC Audio System

The ATSC Digital Television Standard, document A/53, defines a system capable of transmitting high quality audio, video, and ancillary data over a single 6 MHz channel. The system provides video resolution that is five times greater than conventional NTSC and three times the number of audio channels than is possible with BTSC stereo. These higher performance audio and video signals require data rate reduction to fit into the same bandwidth provided by current RF channels. Using Dolby Digital allows up to 5.1 channels of audio to be carried in a single data stream at 384 kb/s (the rate required by ATSC standard A/53 for the Complete Main (CM) channel).

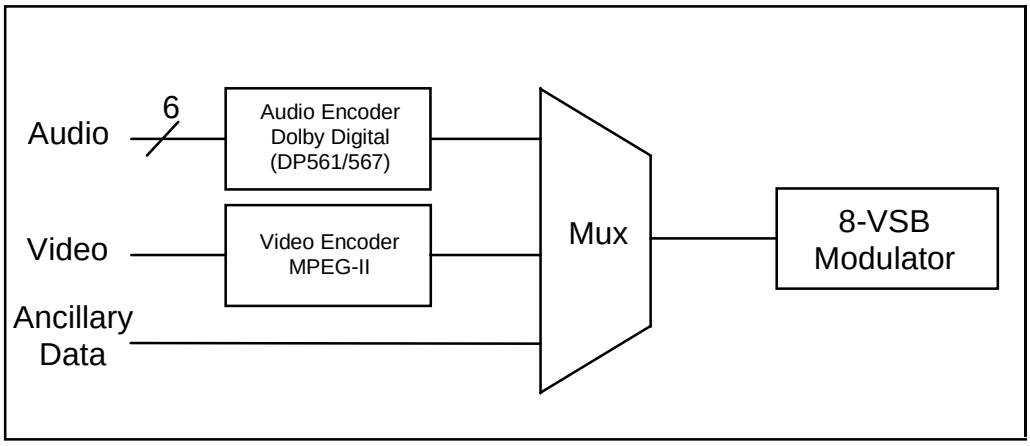

Figure 2-1 Dolby Digital in the Emissions Stage

Dolby Digital is a perceptual audio coding system based on the principles of human hearing and psychoacoustic phenomena coupled with a common bit pool to maximize data efficiency. Mathematical models of the human auditory system are used to determine which parts of the signal input to the coder are perceptible. The most audible components of the signal require the highest resolution and are encoded with the greatest number of bits. Fewer bits are allocated to the masked, or non-audible portions of the signal. In this manner, Dolby Digital allows more efficient audio encoding than linear PCM while still maintaining comparable audio quality. For further information, please refer to *Appendix A – Basics of Perceptual Coding*.

It is important to recognize that Dolby Digital is more than just a highly efficient audio coding system. It was also designed to deliver a multichannel audio program in a format that satisfies a wide range of listening requirements. For example, some listeners may have expensive home theater equipment capable of reproduction over a wide dynamic range. Most listeners, however, use systems that reproduce mono, stereo or Dolby Surround and are often located in noisy environments without isolation. Dolby Digital is an extremely flexible system that satisfies mono, stereo, and Dolby Surround

reproduction by providing compatible downmixes and compressing the dynamic range. For the first time, both the program producer and the end listener can affect how a soundtrack will ultimately be heard. The program producer can deliver a multichannel soundtrack with control parameters that adjust relative playback level and dynamic range, as well as downmix settings, for one- and two-channel compatibility. In many cases, the end user can apply, partially apply, or ignore these parameter settings in order to listen to the program in the manner most appropriate to their equipment and preferences. It is extremely important to recognize that producers and broadcasters must correctly set and monitor these parameters for the end user to derive proper functionality.

Control parameters settings are sent as separate data, called Bit Stream Information (BSI), in the header portion of the Dolby Digital frame. The BSI includes the number and type of channels (audio coding modes), dynamic range compression (*dynrng*), and dialogue (loudness) normalization (*dialnorm*). *Dynrng* and *dialnorm* are two parameters referred to as *metadata*.

Designations for the audio coding mode describe the number of front and rear channels, and whether there is an LFE (Low Frequency Effects) channel present. For example, a full Dolby Digital 5.1-channel signal is designated as 3/2L or 3/2/.1. This denotes three front channels, two rear channels and an LFE channel. A program designated as 2/0 has two front channels, no rear channels, and no LFE channel. Additional data in the BSI indicates whether a 2/0 signal is true stereo or Lt/Rt (Left total and Right total), which is decoded into four channels (left, right, center, and mono surround) by a Dolby Surround Pro Logic decoder (present in nearly all licensed Dolby Digital decoders).

To prevent loud sounds from being too loud, and to insure that low-level sounds are audible, broadcast television has relied on heavy dynamic range compression. Prior to Dolby Digital, this has been a single-ended process that permanently changed the audio signal. Once compressed, there was no way for the viewer to expand the audio back to its original state. Dolby Digital cleverly solves this problem by creating a BSI word during encoding, called *dynrng*. When used by a decoder, *dynrng* provides up to 48 dB of dynamic range compression on a block-by-block basis (every 5 ms). All Dolby Digital decoders default to using the *dynrng* values. Since it is separate information, the viewer can scale the value of *dynrng* to restore all, or part, of the original dynamic range. Using Dolby Digital, the broadcaster can compress the audio to suit most of the audience while still allowing viewers the option for reproduction with the full dynamic range.

Loudness normalization has been a long-standing problem for broadcasters. Program-to-program loudness differences have normally been fixed with audio processors designed to minimize dynamic range. Dolby Digital addresses this in a different manner. Another BSI word generated during encoding, *dialnorm*, provides the decoder with information to adjust playback level. This control word represents the difference between full-scale output and the LeqA value of the dialogue. LeqA, or loudness equivalent, is the long-term RMS value of the A-weighted audio level. The output of the decoder is variable in 1 dB increments from -31 (unity) to 0 (a 31 dB attenuation in output level). If television programs are assigned the proper *dialnorm* values, the decoder will adjust the output level appropriately, producing consistent loudness independent of the source. For further information on *dynrng* and *dialnorm*, please refer to *Appendix B – Metadata*.

The following are the key features of Dolby Digital. The encoder must be presented with full dynamic range program material, as well as the proper metadata, for these features to function properly.

- Delivers from one (mono) up to six discrete audio channels
- Highly efficient audio coding: approximately 15 times less data when delivering 5.1 channels compared to six channels of 20-bit linear PCM
- Supports output data rates from 32 kb/s to 640 kb/s
- Full-bandwidth frequency response in all five main channels
- LFE (Low Frequency Effects) channel for extended bass (3 Hz – 120 Hz)
- Encoder accepts word lengths of up to 24 bits
- AC-3 algorithm provides 24-bit dynamic range
- Uniform, level-matched playback for all sources
- Calibrated playback volume via “floating” reference level determined by the program producer
- Provides producer-set/user-controlled dynamic range compression
- Allows multichannel programs to be downmixed to stereo or mono
- Two-channel downmix can be Dolby Surround encoded
- Lo/Ro downmix can be optimized by the producer during post-production
- Future improvements in encoding algorithms can benefit all listeners, even those with existing decoders

2.2 The Consumer Receiver

2.2.1 Definitions and Performance

All licensed Dolby Digital products can, in some way, decode a Dolby Digital bitstream into audio. Some devices, such as DTV receivers and DVD players, are referred to as source, or set-top products since they decode only the delivery format for which they have been designed. Devices that have a Dolby Digital input but do not necessarily incorporate a Dolby Digital source are classified as decoder products.

All source or decoder products are able to perform dynamic range compression and dialogue normalization by applying *dynrng* and *dialnorm* values to the decoded audio. All products must be able to downmix the multichannel audio into various compatible formats. Source products must provide conventional two-channel audio outputs for compatibility with current consumer A/V systems. Source products, therefore, normally have an internal two-channel Dolby Digital decoder that downmixes 5.1-channel programs into a Dolby Surround (Lt/Rt) compatible two-channel output suitable for external Dolby Surround Pro Logic decoding. Source products must also provide access to the Dolby Digital bitstream via a DDIF (Dolby Digital Interface, standardized as IEC 61937) output, which allows full-featured decoding by an external Dolby Digital decoder. Multichannel Dolby Digital decoders typically contain a Dolby Surround Pro Logic decoder as well to maintain compatibility with the large volume of existing program material. This allows the same multichannel decoder to derive appropriate outputs from both 5.1-channel and Dolby Surround encoded programs. An exception to this rule is a product called a Dolby Digital “adapter.” The purpose of an adapter is to add Dolby Digital decoding to an existing Dolby Surround Pro Logic system. In this case, the Dolby Surround Pro Logic decoder within the Dolby Digital decoder is redundant. It is usually included, however, owing to straightforward digital implementation.

2.2.2 The Dolby Digital Decoder

Dolby Digital decoders can operate in two *modes*, called *Line* and *RF*, designed into the Dolby Digital decoder IC to provide simplified implementations of dynamic range compression, dialogue normalization and downmixing based on the application. Line Mode is used in most two-channel source (set-top) decoders, two- and 5.1-channel televisions, and outboard Dolby Digital adapters. RF Mode has been optimized for products that create a downmixed version of the audio that will be used for subsequent channel 3-type RF modulation or for listening to the internal speakers of a television.

Line and RF Modes always use *dialnorm*, the dialogue normalization setting, to maintain a consistent dialogue level between programs and DTV channels. dialogue is reproduced at a constant level of -31 dBFS (decibels below full-scale) in Line Mode and at -20 dBFS in RF Mode. In RF Mode, overall program level is boosted 11 dB within the decoder and the peaks are limited (using a metadata parameter called *compr*) to prevent overload of the digital-to-analog converters. In products containing RF modulators or demodulators for conventional analog audio, the analog RF Mode output is boosted a further 6 dB relative to those RF sources. By limiting headroom to 20 dB above the

average dialogue level, over-modulation of the television receiver is prevented while providing an acceptable average RF modulation level close to that of a quality television broadcast. For further information on *dialnorm*, *dynrng*, and *compr*, please refer to *Appendix B – Metadata.*

Dolby Digital decoders are capable of delivering a wide range of outputs with little or no intervention from the listener. The decoder can be commanded by parameters set during encoding to reconfigure to the appropriate mode. For example, a properly encoded Dolby Digital bitstream can automatically activate Dolby Surround Pro Logic decoding when the television station switches from 5.1-channel to Dolby Surround two-channel program material.

Since speaker systems and configurations vary widely among listeners, Dolby Digital decoders provide a feature called bass management (redirection), which allows low frequencies to be sent to the speakers most capable of reproducing them. Although all five main channels in a Dolby Digital 5.1-channel system are capable of full-bandwidth reproduction, some speakers in a home system may not have adequate low frequency response. For example, a system without a subwoofer may reproduce bass much better in the left, center, and right speakers than in the surround speakers. In this case, the bass frequencies should be directed only to the front speakers and not to the surrounds. Another system may have a subwoofer but lack other speakers that adequately reproduce bass frequencies. In this case, all of the bass should be directed to the subwoofer. For further information on bass management, please refer to the *Dolby Digital Professional Encoding Manual* available on the Dolby web site at www.dolby.com/tech/.

Chapter 3 – Program Contribution and Distribution

3.1 5.1-Channel Audio

3.1.1 Plant Infrastructure

Current television plant audio distribution facilities are generally limited to routing two, and sometimes four channels of analog or digital (two AES/EBU pairs) audio. Video tape recorders (VTRs) are either two-channel analog or four-channel digital (two AES/EBU pairs) devices. Telco and satellite links have data rate, or bandwidth restrictions that limit their baseband capacity to less than the six channels required by programs carrying 5.1-channel sound.

Broadcast plants also lack the facilities required to handle metadata control information. Metadata is an essential part of the Digital Television system and must be included in the final transmission.

3.1.2 Dolby E Data Rate Reduction

Dolby Digital, with its highly efficient data-rate reduction, is used between the Transmitter and home Receiver to conserve bandwidth. Within the Contribution and Distribution portion of the DTV system, AES/EBU signal paths are available that can carry data at 2.304 Mb/s if all 24 bits of each channel in the pair are used (assuming a sampling rate of 48 kHz). Most professional digital VTRs are limited to 20-bit audio, yielding a data rate of 2 Mb/s. These rates are high enough to carry between six and eight data rate reduced digital audio channels and their accompanying metadata with enough margin to allow the audio to be decoded, processed and re-encoded eight to ten times. Even with a 16-bit wide path, multiple channels (with some restrictions) and a metadata path can be provided.

A significant advantage of using the Dolby E distribution coder to carry the multichannel audio within an existing AES/EBU infrastructure is that the data rate in the plant and Telco and satellite links can be identical. This avoids the generation losses caused by re-encoding simply to change data rates to accommodate different transport channels.

The following Figures show how Dolby E coding equipment can be used to provide a path for six to eight audio channels and metadata through a television plant equipped with at least one layer of AES/EBU routing.

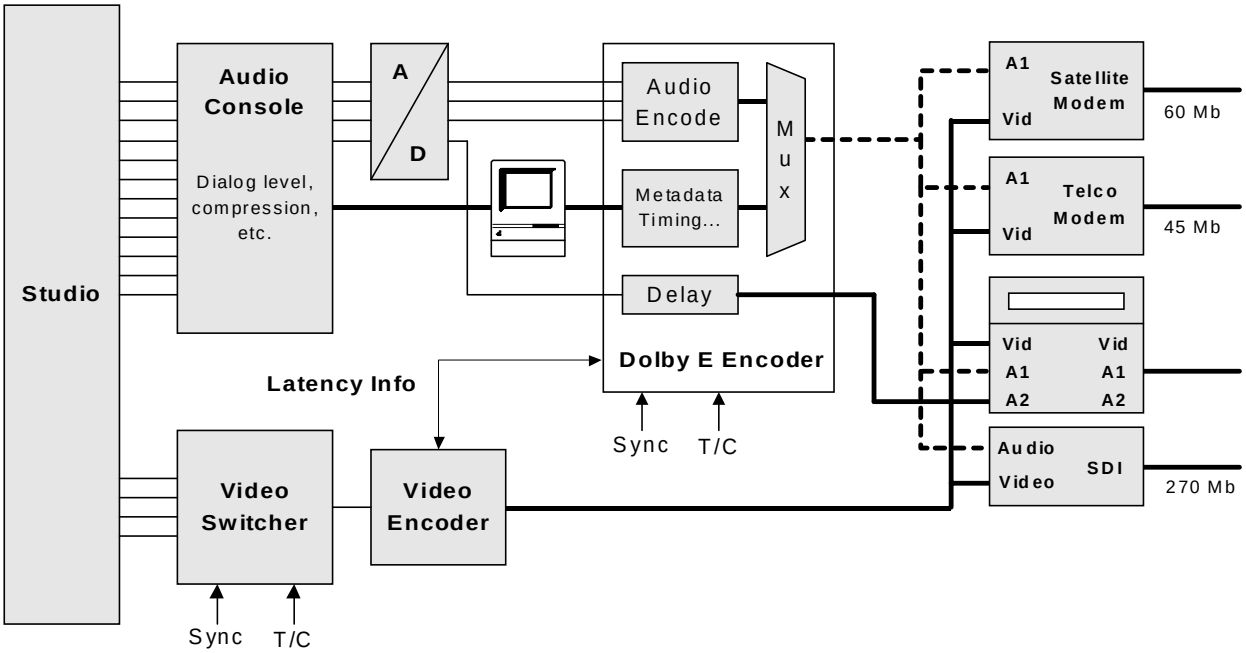

Figure 3-1 Original Production Process – Contribution Stage

Even in the studio, the path is full of bottlenecks for a six-channel signal. Current tape machines, routing switchers and embedders cannot handle all the channels or metadata required by the program. Limited bandwidth produces a similar problem for HD video. The solution to both problems is to use data reduction techniques capable of withstanding several generations of encoding/decoding, and a multiplex to carry the metadata.

Audio and video coders impart significant delays to their signals. The “Latency” connection between the coders allows them to compensate for the delays so that the audio and video signals are synchronous as they go to the modems, tape machine or embedder.

The terminal, shown in the metadata path, is used to insert or modify the information needed by the Dolby Digital emission encoder located at the end of the contribution and distribution systems.

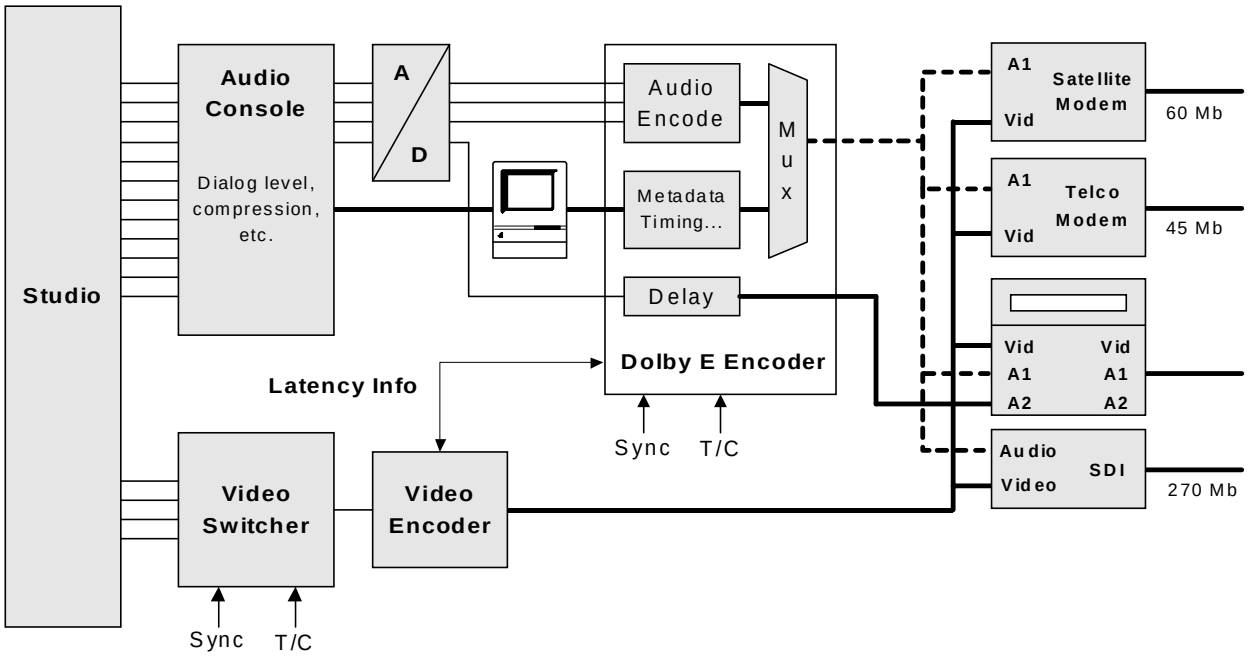

Figure 3-2 Post-Production Process – Contribution Stage

Post-production must be done in the linear (decoded) domain. Audio and video may need to be kept synchronized (hence the latency connection) unless the work is done independently, in which case both signals have to be accurately time-stamped to place the audio back in sync with the video.

Since metadata cannot be carried through the router or audio console, it must be bridged around the production tools. The data bridge can be very simple or, in complex plants, might be an additional layer on the router. The path to the output encoder includes a terminal that allows the data to be modified to conform to the changes made during post-production.

The output of the post-production process consists of completed programs or portions of programs. For the purposes of this document, it is assumed that the Network Origination Center will distribute finished program segments to the affiliate stations where they will be assembled into complete programs or sequences of programs and commercials.

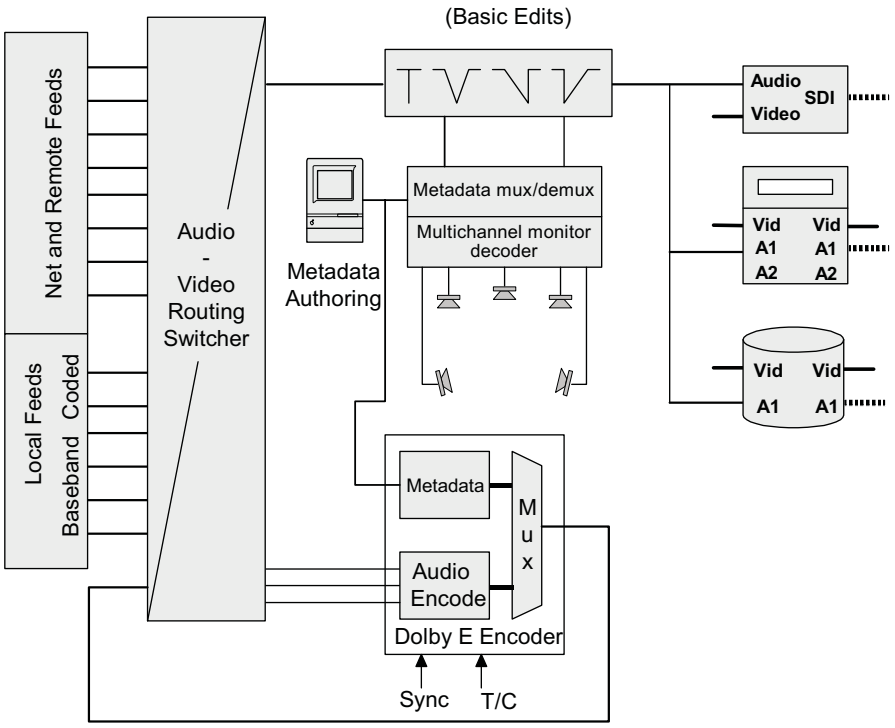

Figure 3-3 First Stage Program Distribution – Assembly or Integration

The first step in the distribution system brings programs, and other feeds, to the Affiliate where they are assembled into completed sequences. It is assumed that almost all the feeds to the affiliate are encoded using light data reduction techniques. The (local) signals that are not encoded can be routed through a utility coder that is “wrapped around” the routing switcher. Unencoded or linear signals may be included in the program packaging process as long as there are enough channels in portions of the plant to handle them. The four basic transitions (Take, V Fade, Fade Take, and Take Fade) are probably sufficient to handle the bulk of the program packaging requirements. If voice-overs or crossfades are necessary, the signals must be decoded, processed and re-encoded.

A six-channel decoder and monitoring system is provided for quality assurance operations and to listen to (or preview) the program material as it will be presented in the various home settings. Metadata authoring is included to allow the operator to change or correct various parameters, such as *dialnorm* values, to ensure consistent program-to-program performance.

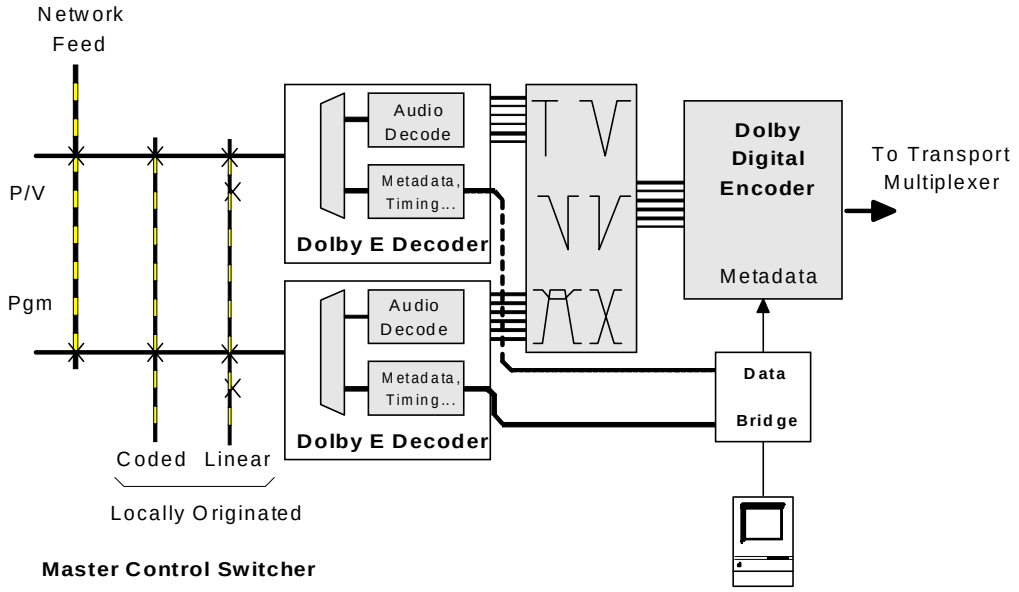

Figure 3-4 Second Stage Program Distribution – Master Control

The Master Control switcher is usually the last component in the Distribution chain. M/C can be followed by a Dolby E encoder if the signal must be routed to additional locations as well as to the transmitter.

A Master Control switcher routes multichannel audio signals that have undergone light data reduction and decodes the signal prior to performing a crossfade or voice-over. The metadata is demuxed and bridged around the multichannel audio processor. The terminal is included to allow optional addition to, or modification of metadata before it is sent to the Dolby Digital emission encoder. Multichannel audio monitoring facilities are assumed but not shown in the Figures. Since this is normally the last point in the Distribution chain, multichannel audio monitoring facilities should allow previewing the effects of the home receiver's decoder to allow metadata modification, if necessary.

3.1.3 Generating Metadata

Figure 3-1 and Figure 3-4 show a generic facility for authoring metadata. Much of the information required by the Dolby Digital emission encoder is static (changes only occur with program format changes) while the *dynrng* metadata changes in real time.

Several different methods to author the information are needed. Most of the static information can be set in production, post-production, or when packaging the signal. The *dynrng* information is produced in post-production or in the encoder by monitoring the audio and producing the data required by the receiver to control the dynamic range of the decoded signal. Since the Master Control area is the last point to monitor or modify metadata, some type of facility is necessary there.

None of the metadata authoring tasks have been fully defined yet, so the authoring tools and methods of using them cannot be described at this time. Consultations with broadcast operations specialists will be necessary to fully define the tools.

3.1.4 Equipment Requirements

Any digital audio system that carries non-linearly coded (typically rate-reduced) audio data cannot be permitted to modify the information in any way. Operations, such as gain changes, modify the data and destroy the encoding scheme, rendering it useless. Any operation which may corrupt the coded information should be avoided.

It is important that bit 1of byte 0 of the Channel Status data (the audio/non-audio flag) of the AES/EBU signal carrying the Dolby E coded audio be set to “1” to indicate that it is a data signal and not a linear or baseband audio signal. The audio/non-audio flag must also be set at the output of any equipment (such as a DVTR) carrying the Dolby E signal.

Gain Controls

Input (sensitivity) and output (level) controls must have unity gain positions or bypass functions that ensure the data recorded and reproduced by, or passing through the system is identical to the input data.

It should be possible to select the unity gain function manually or automatically. In the automatic mode, the unity gain function should be selected when the audio/non-audio bits of the input AES-3 signal or the signal being reproduced are set to “1.”

Word Length

A system must not change the word length of a non-audio signal by truncation. If bits 0, 1, and 2 of byte 2 of the channel status are set to “0” (the default state), then any non-audio data must be assumed to be 20 bits long. If the word length of the non-audio signal is not 20 bits, the channel status information (byte 2, bits 0–5) must be set to indicate the intended word length and must either be carried through the system or set appropriately at the system output.

Splicing and Switching

Recording systems, switchers, editing systems, and similar devices must be able to make end-to-end splices in data streams. The system switching points and the structure of the non-linearly coded information must be compatible so that switching operations do not destroy the integrity of the data.

Signal Processing

Processes that can modify the data, such as crossfades, fades to or from silence, and rounding or dithering, must not be attempted when handling non-linearly coded data.

Sample Rate or Data Rate

The data rate of a non-audio signal cannot be changed by the interpolation process, which changes the sampling rate of digital audio signals, without destroying that data. If it is necessary to transfer the data between two systems operating at different data (or sampling) rates, the non-linearly coded information must be “repacked” in the payload portion of the second channel without modifying the data. This can be accomplished by changing the quantity of stuffing bits. Supervisory information describing the non-linear signal may also have to be modified to reflect the changes made.

3.2 A Near-Term Solution

3.2.1 Dolby Surround in Broadcast

Approximately 60-70% of current television audio facilities are equipped for two-channel analog signals. Since Dolby E requires at least one AES/EBU digital audio pair, it cannot be used in plants with only analog equipment.

A Dolby Surround encoder, on the other hand, matrix encodes left, center, right, and mono Surround signals into a pair of Lt/Rt (Left total and Right total) analog channels which can be used in all stations. Although Dolby Surround lacks some of the operational features of the Dolby Digital system, such as independent surrounds and the low frequency effects channel, it is a common release format for virtually all multichannel productions. Over 31 million Dolby Surround decoders are currently in use by consumers, so this provides a practical method for distributing multichannel programs to the listener before Digital TV receivers are in place.

A near-term solution has to provide for the inability to control remotely located Dolby Digital emission encoders. Since there are no signal paths for metadata from the network origination center, the encoders must remain in a suitable default broadcast mode until they can be controlled. This default broadcast mode must provide satisfactory service to as many listening situations and as many program formats as possible.

If a Dolby Digital encoder is in the 2/0 mode and the Dolby Surround Mode parameter is correctly set, the receiver can then use this setting to decode an Lt/Rt signal into four channels. Lt/Rt signals are decoded by a receiver such that the in-phase information is reproduced by the center channel, firmly anchoring important information, such as dialogue, to the screen, and the out-of-phase information correctly appears in the mono surround channel. Lt/Rt signals are both mono and stereo compatible (although surround channel information is not present in a mono downmix).

True six-channel signals must be mixed in post-production to four channels (left, right center, and mono surround) and then Dolby Surround encoded to function correctly in this scenario. This signal is reproduced by the various receiver types described in *Section 2.2 - The Consumer Receiver*. The following diagrams show how this short-term solution might be implemented.

In a two-channel plant, the only viable multichannel option is to use Dolby Surround encoded audio (Lt/Rt). It is assumed that the video will undergo limited date rate reduction using a video coder in the plant and on the inter-facility links. In this case, the baseband, or linear audio should be delayed to match the video encoding latency. A connection between the video coder and the audio delay unit ensures that the audio and video stays in sync from the input to the next stage in the process.

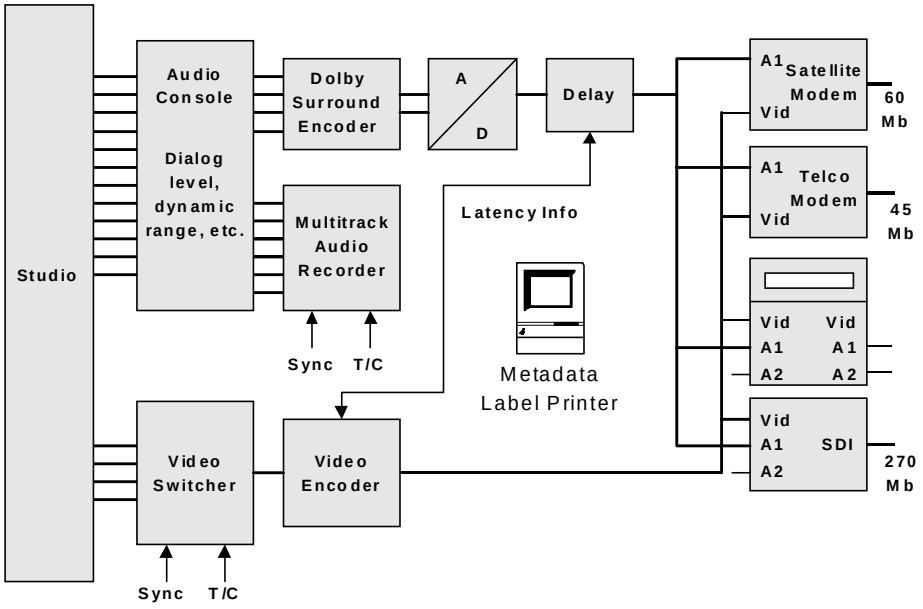

Figure 3-5 Original Production Process – Two-Channel Analog or Digital System

Figure 3-5 shows a single AES/EBU pair from the ADC through the delay to the DVTR and SDI embedder. The audio payload data rate on a single AES pair (about 2 Mb/s) permits multiplexing with the video for inclusion on a satellite or Telco circuit. This allows distribution of the two-channel Dolby Surround encoded signal throughout the entire system.

If it is necessary to handle more than four channels, the audio must be done double system. The extra handling and synchronization problems may be worth the effort for special projects. Since there is no path for metadata in this system, the terminal can be used for other purposes such as printing tape box labels.

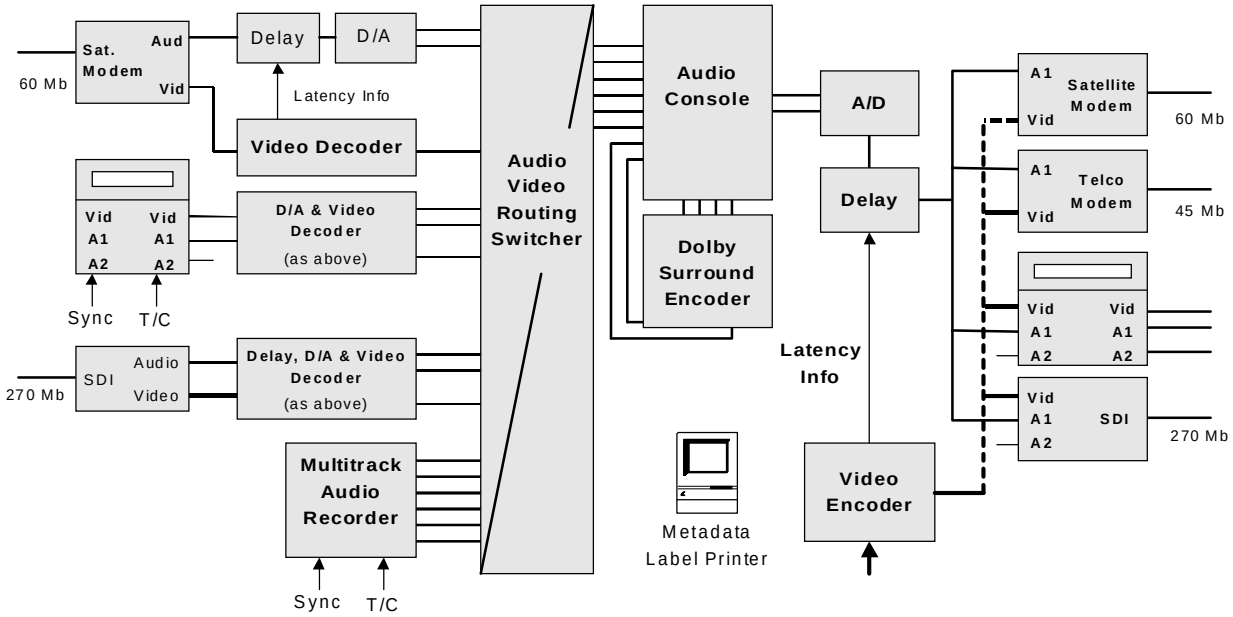

Figure 3-6 Post-Production Process – Contribution Stage

The contributed material is recovered from the inter-facility links, or played back from a DVTR. If necessary, each of the outputs is run through a delay to compensate for the video decoding process. Synchronization of any multitrack material is maintained by chase-locking the audio to the video or house time code.

A utility Dolby Surround encoder is “wrapped around” the audio console to deal with any material not already in Lt/Rt format. The console output again passes through a delay to compensate for the encoding delay of the accompanying video material, assuming that it exists. There is no metadata signal path, so a terminal is provided to track the information and print new tape box labels as needed.

Audio and video may require synchronization (hence the latency connection) unless the work is done double system, in which case both signals have to be accurately time-stamped so the audio can be laid back in sync with the video.

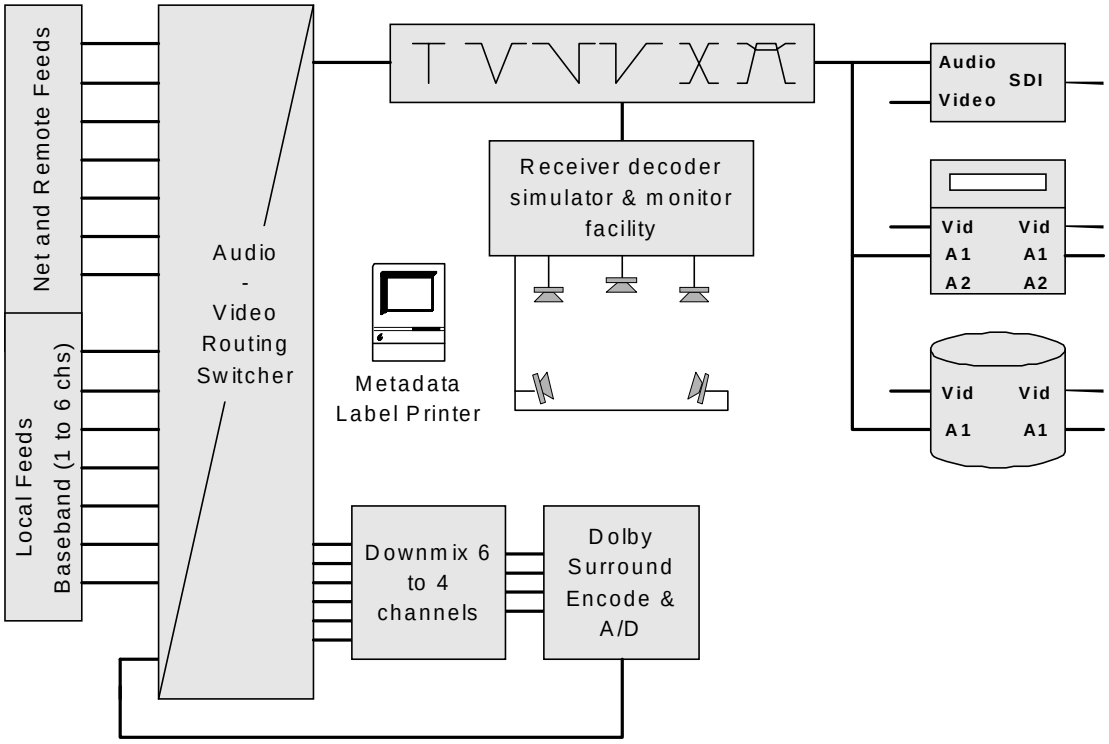

Figure 3-7 Packaging or Integration at the Affiliate – Distribution Stage

Network, remote, and local feeds are combined into the final "to air" sequence. Most of the incoming feeds are Dolby Surround encoded. Other (local) signals may be in mono, stereo, discrete four-channel, or 5.1-channel format.

The Dolby Surround material can be treated as if it were any two-channel material. Fades, cuts, crossfades, and voice-overs do not affect the Dolby Surround encoding, as long as the proper inter-channel balance and phase are maintained. A Dolby Surround Pro Logic decoder in the home receiver gracefully decodes the main signal formats currently in use.

If a Dolby Digital encoder is in the 2/0 mode and the Dolby Surround Mode parameter is correctly set, the receiver can then use this setting to properly decode an Lt/Rt signal into four channels. Lt/Rt signals are decoded by a receiver such that the in-phase information is reproduced by the center channel, firmly anchoring important information such as dialogue to the screen, and the out-of-phase information correctly appears in the mono surround channel. Lt/Rt signals are both mono and stereo compatible (although surround channel information is not present in a mono downmix).

True 5.1-channel signals must be mixed in post-production to four channels (left, right, center, and mono surround) and then Dolby Surround encoded to function correctly in this scenario. This signal is reproduced by the various receiver types described in *Section 2.2 - The Consumer Receiver*.

The monitoring facilities should simulate the decoder in the home receiver to allow the operator to check the metadata values carried through the rest of the system. A terminal is provided to input data to the automation system if that is the path used to get metadata to the Dolby Digital emission encoder. Otherwise, the information must be carried to the next stage manually.

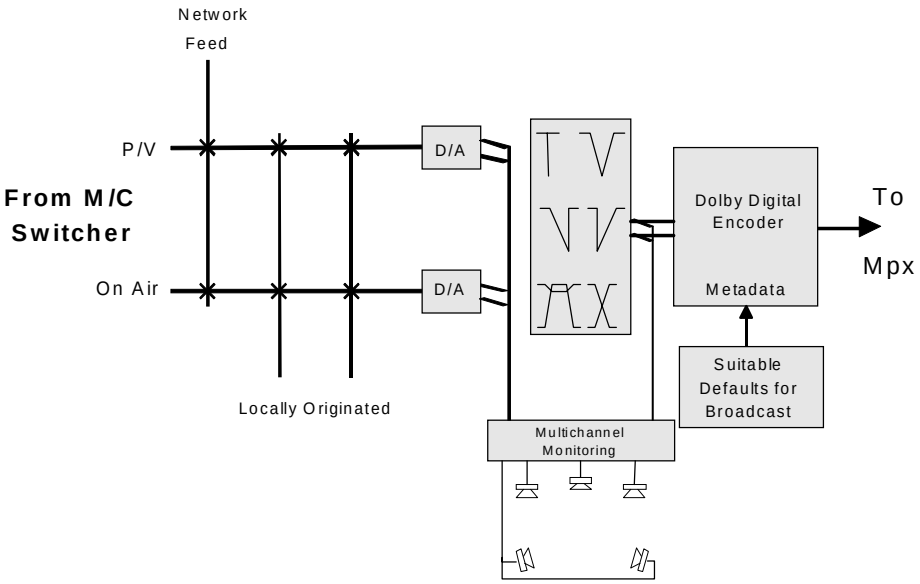

Figure 3-8 Two-Channel Analog Master Control – Distribution Stage

In the short-term, the Dolby Digital emission encoders must transmit a set of suitable default broadcast metadata parameters because there is no path for this information between the affiliate and the Network Origination Center.

The biggest change needed for Master Control is the addition of a multichannel audio monitoring system that allows the operator to listen to an emulation of the Dolby Digital decoder used in domestic receivers. It also provides a last opportunity to change the metadata, if necessary. Otherwise, the Master Control area can remain very much as it is now.

3.2.2 Dolby Surround Principles of Operation

Introduction

Dolby Surround is a two-step, encode/decode process involving both recording and playback. When a Dolby Surround soundtrack is produced, four channels of audio information (left, right, center, and mono surround) are encoded onto two audio tracks using a Dolby Surround encoder. These two tracks are then carried on stereo program sources, such as videotapes and TV broadcasts, into the home where they can be processed by a Dolby Surround decoder to recreate the surround sound experience.

The Dolby MP (Motion Picture) Matrix

The encoder accepts four separate input signals (L, C, R, S), and creates two final outputs, left-total and right-total (Lt/Rt). Refer to the conceptual drawing below.

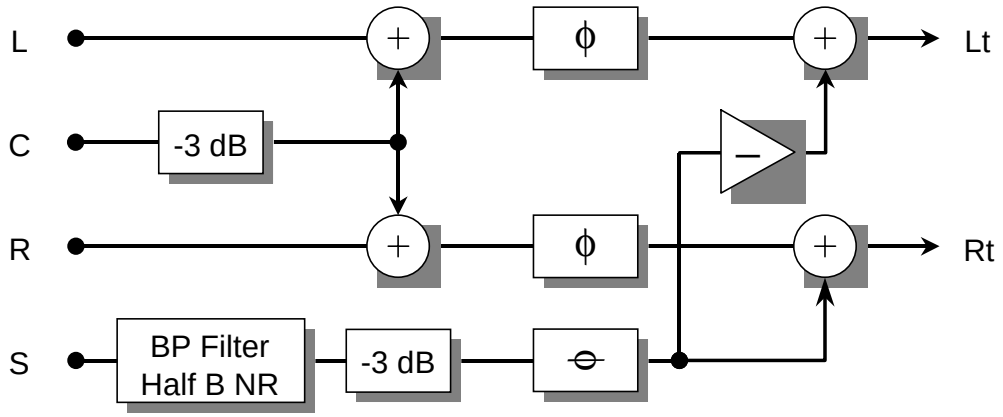

Figure 3-9 Dolby Surround Encoder

The encoder takes the left and right inputs and passes them to the Lt/Rt outputs without modification except for an all-pass network (ϕ) in each channel. The center channel input is divided equally to left and right signal paths with 3 dB attenuation to maintain constant acoustic power in the mix. The surround channel is attenuated 3 dB, band-pass filtered with a 100 Hz high-pass and 7 kHz low-pass filter, encoded with a modified form of Dolby B-type noise reduction, and then all-pass filtered (ϕ). The differential between the front channel all-pass (ϕ) and surround channel all-pass (θ) is a constant 90° over frequency. The surround channel is summed into the Lt/Rt outputs with polarity inversion performed prior to addition to the Lt output.

Clearly there is no loss of separation between the left and right signals; they remain completely independent. It is not as obvious that there is also no theoretical loss of separation between the center and surround signals. Since the surround signal is recovered by taking the difference between Lt and Rt, the identical center channel components in Lt and Rt exactly cancel each other in the surround output. Likewise, since the center channel is derived from the sum of Lt and Rt, the equal and opposite surround channel components cancel each other in the center output.

For this cancellation technique to maintain high separation between center and surround signals, the amplitude and phase characteristics of the two transmission channels must be as close as possible.

For further information on Dolby Surround, please refer to the *Dolby Pro Logic Surround Decoder, Principles of Operation* document available on the Dolby Web site at www.dolby.com/tech/. For information on Dolby Digital decoder downmixing, please refer to *Section 3.2.3 - Downmixing 5.1-Channel Audio*.

The Dolby Surround Pro Logic Decoder

Dolby Surround Pro Logic operates by continuously monitoring the incoming signals to determine the degree and direction of the dominant signal in the soundfield. The circuit then dynamically cancels adjacent-channel crosstalk with a vector-cancellation steering system. As the leakage signal is eliminated, the steered signal is increased proportionally to maintain constant acoustic power of all signals; no modulation side-effects are created.

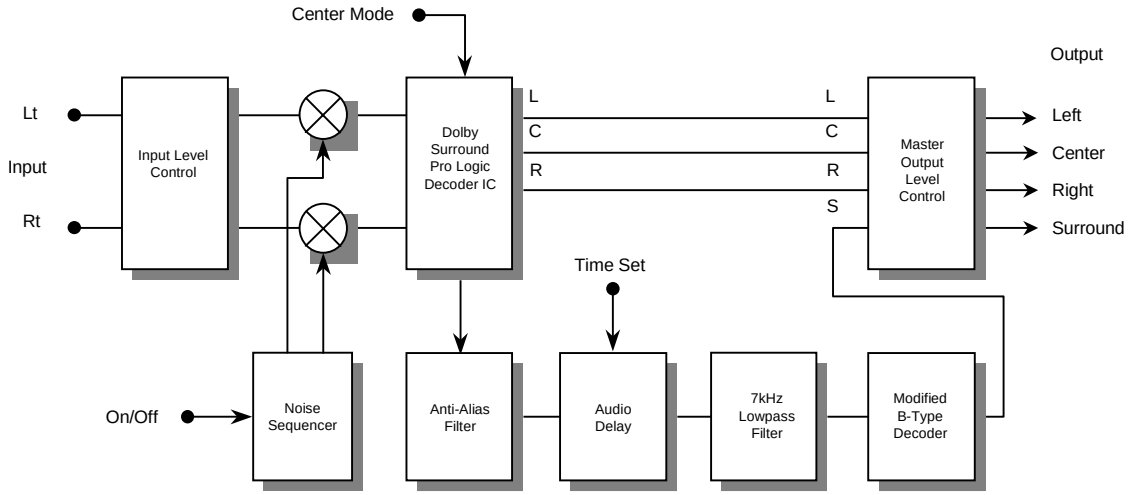

Figure 3-10 Dolby Surround Pro Logic Decoder Block Diagram

- After the matrix decoding is complete, time delay, low-pass filtering, and mild noise reduction are applied in the remainder of the surround decoding chain.
- Time delay ensures that any front channel sounds that happen to leak from the surround speakers arrive at the listener just after the front channel sounds. This helps prevent the leakage from pulling the sound image away from the screen. Such leakage may be caused by phase or amplitude differences present in the input signals caused by azimuth misalignment, frequency response errors, or simple balance errors in the stereo program.
- The 7 kHz low-pass filter is used for several reasons. Most importantly, for a given azimuth error between the two audio channels, the leakage signal magnitude increases with frequency, making separation at high frequencies much more difficult to achieve. Dialogue sibilance, for example, can distract from the surrounds without use of the filter. Reducing the high-frequency content also makes the surround speakers sound further away and more difficult to localize, which benefits those seated closest to the surround speakers.

Modified Dolby B-type NR is used to reduce noise as well as front-channel signal leakage. The amount of processing was reduced from 10 to 5 dB to prevent the encoded surround signal from significantly altering the nature of the left and right channel signals, and to avoid any side-effects caused by processor mis-tracking.

All Dolby Surround Pro Logic decoders include basic features to optimize a variety of speaker configurations. To allow the use of a smaller center speaker without sacrificing bass quality, a Center Mode switch offers a Normal setting that redirects the bass from the center speaker to the main left/right outputs. If no center speaker is available, the Phantom Mode splits the entire center signal equally to the left/right outputs. The Wide Mode provides a wideband center signal when a full-range speaker is used.

All decoders include an internal calibration signal generator and level adjustment controls to properly balance all the channels. A master volume control adjusts all channels without disturbing the balance among the output channels.

3.2.3 Downmixing 5.1-Channel Audio

Multichannel broadcast audio must be able to meet the needs of consumers with any type of playback system, including Dolby Surround Pro Logic decoders, as well as stereo and mono. Dolby Digital receivers, set-top boxes, and DVD players can downmix five-, four- or three-channel program material to a two-channel mix compatible with Dolby Surround, stereo, or mono. This mix will be available from the two-channel output of the receiver. A consumer who does not have a separate Dolby Digital decoder can obtain a Dolby Surround, stereo, or mono compatible downmix for an existing system. Downmix compatibility should be checked when preparing multichannel program material for broadcast. Dolby Digital provides two downmix types: Dolby Surround (Lt/Rt) compatible downmix and stereo (Lo/Ro) downmix. Dolby Digital decoders can offer the choice of whether to downmix to Dolby Surround (Lt/Rt) compatible or stereo (Lo/Ro).

Dolby Surround Compatible Downmix

A Dolby Surround (Lt/Rt) compatible downmix is accomplished in two stages, encoding and decoding. Refer to the figure below.

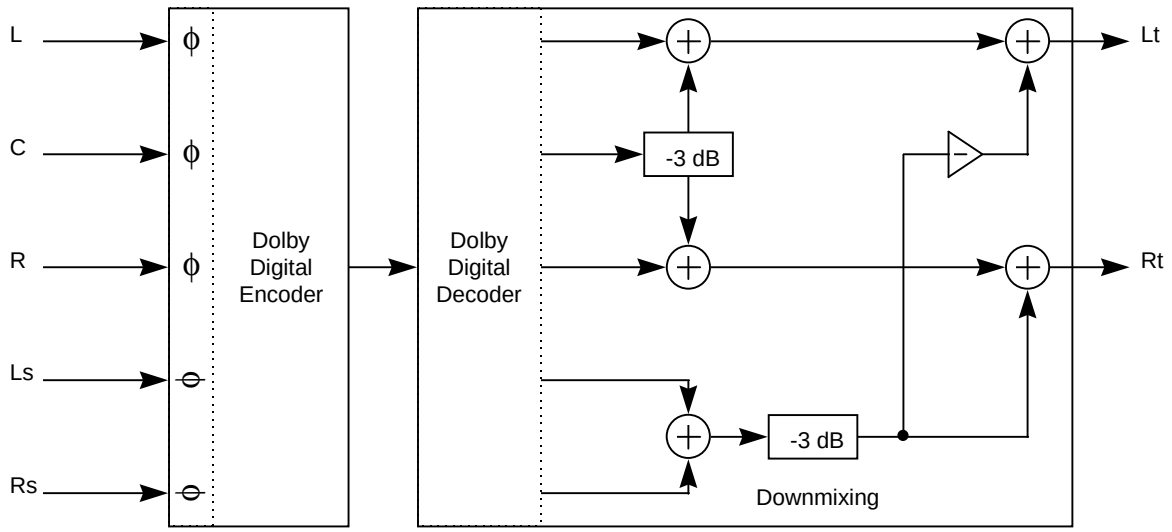

Figure 3-11 Dolby Surround (Lt/Rt) Compatible Downmix Signal Flow

During the encoding process, a 90° phase shift is applied equally to both surround channels, which helps maintain smooth front to back panning and reduces cross-talk between the center and surround channels. Applying the 90° phase shift during encoding removes the need for the receiver to perform this function in DSP.

In the decode process, the receiver:

- attenuates the center channel 3 dB and directs it equally to the left and right outputs;
- combines both surround channels, attenuates them 3 dB, directs the polarity-inverted signal to the left output and the unchanged signal to the right output.

Stereo Downmix

A stereo downmix (Lo/Ro) is compatible with a standard two-channel stereo listening environment. Refer to the figure below.

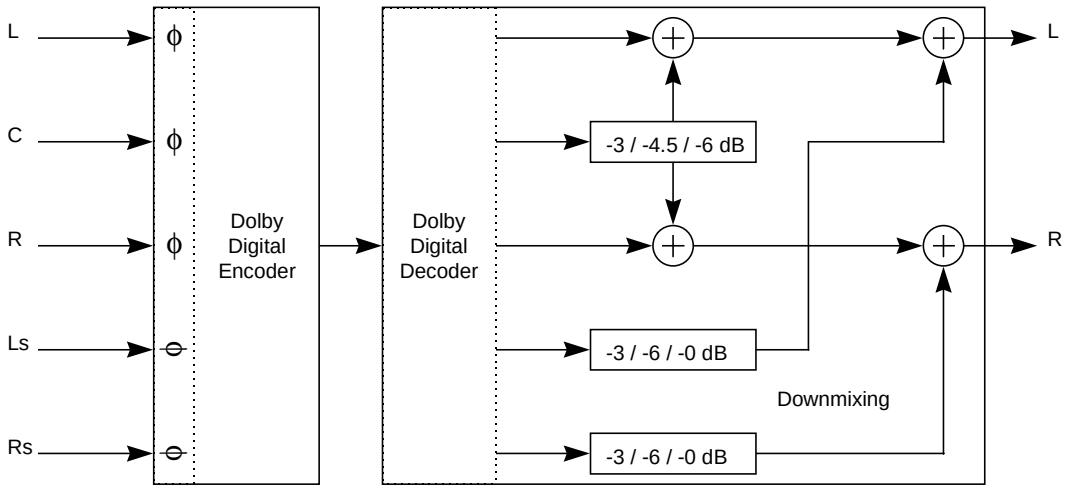

Figure 3-12 Stereo (Lo/Ro) Downmix Signal Flow

In the decode process, the receiver directs the center channel signal equally to the left and right outputs while directing the left surround channel to the left output and the right surround channel to the right output. A choice of attenuation levels is available for the center and surround channels. This is important because center and surround levels may seem excessive with a standard 3 dB attenuation. During encoding, the amount of attenuation applied to the center and surround channels may be selected in the bitstream information section of the encoder.

Table 3-1 Center Downmix Levels

Center Downmix Level	Mix Level
1 (00) (default)	0.707 (-3.0 dB)
2 (01)	0.596 (-4.5 dB)
3 (10)	0.500 (-6.0 dB)

Table 3-2 Surround Downmix Levels

Surround Downmix Level	Mix Level
1 (00) (default)	0.707 (-3.0 dB)
2 (01)	0.500 (-6.0 dB)
3 (10)	0 (-∞ dB)

LFE Channel

The LFE, or .1 channel, is not included in any downmix. It is therefore necessary to audition the 5.1-channel material in downmix mode to assure that the LFE information is not essential; it may already exist in the left, center, or right main channels. If the LFE channel contains important, discrete audio information that is not present in any of the main channels, then some, or all of the LFE information should be redirected to the left, center and/or right main channels.

Monitoring Through an Encoder/Decoder

Monitoring the program material through a Dolby Digital encoder and decoder is the best way to properly check downmixing for compatibility. Refer to Figure 3-13.

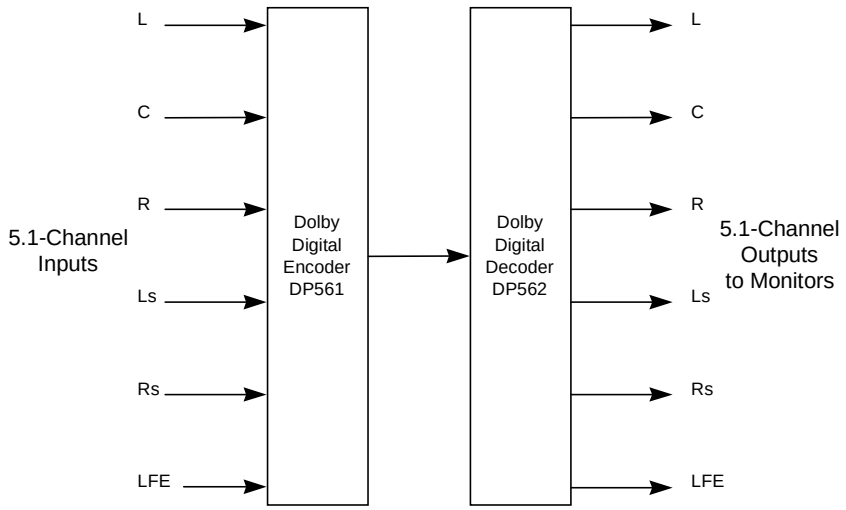

Figure 3-13 Encode/Decode Downmix Signal Flow

The DP561 Dolby Digital encoder allows adjustment to parameters, such as the Center and Surround downmix levels, to be monitored in a downmix with the DP562 Dolby Digital professional reference decoder. The Dolby DP562 decoder also provides easy checking of all downmix configurations including Dolby Surround Pro Logic decoding of a Dolby Surround (Lt/Rt) compatible downmix from five-, four-, and three-channel program material.

Mixing Considerations

By understanding the downmixing process, the following problems that can arise when mixing new 5.1-channel material or auditioning existing multichannel program material can be easily detected and avoided.

- The LFE is not included in any downmix. It is therefore necessary to audition the 5.1-channel material in downmix mode to assure that the LFE information is not essential; it may already exist in the left, center, or right main channels. If the LFE channel contains important, discrete audio information that is not present in any of the main channels, then some, or all of the LFE information should be redirected to the left, center, and/or right main channels.
- A downmixed and decoded mono surround channel is derived from the discrete stereo surround channels by summing them together, attenuating them 3 dB, and applying a 100 Hz to 7 kHz band-pass filter.

  The downmix assumes stereo surrounds contain discrete information. In many cases, however, the stereo surround channels actually contain mono information. In these cases, the downmixed mono surround may appear to be several dB louder than the discrete stereo surrounds. You may wish to attenuate the stereo surround level for a better balance between the discrete stereo and downmixed mono surround.

If the discrete stereo surrounds contained significant information above 7 kHz, the mono surround may appear to be several dB quieter. Equalizing information in the 5 – 7 kHz region of the stereo surround channels may help the downmixed mono surround better match the level of the discrete stereo surrounds.

- Audio in the left and right main channels that contains incoherent phase information, such as stereo reverb effects or material recorded with stereo microphones, may shift to the surround channel in a surround decoded downmix, while remaining discrete left and right in the 5.1- or two-channel mix. Although this may result in a pleasing effect in many cases, panning the left and right channels of the particular effect slightly towards the center can help eliminate the problem.
- Information in the left and right main channels that is mono and equal in level collapses to the center channel in a surround decoded downmix. If this occurs, it is best to treat the actual sound effect itself and not the entire left or right channel. A small delay (8 ms) may be added to either the left or right sound effect, or musical instrument, to shift the phase slightly and cause it to remain hard left and right. If this technique is applied, mono compatibility should be checked.
- Surround channel information is not present in a mono downmix. If the surround channel is deemed important to the overall program material, then some of the surround signal should be redirected to the center channel. Panning the effect to the center and surround channel can be a good compromise.

Chapter 4 – Multichannel Monitoring

4.1 Production Environment

Many factors contribute to the effective design of a multichannel production facility. Room dimensions, ergonomics, and equipment all play an important role in creating a comfortable and efficient environment for audio production. Speaker choice and placement are critical for accurate multichannel monitoring. Employing the Dolby Model DP562 professional reference decoder allows for both Dolby Digital and Dolby Surround Pro Logic decoding. Once the installation is complete, playback levels and center/surround delay settings must be calibrated. These and many other factors directly affect the quality and consistency of the resultant audio.

For further information on audio production environments, please refer to the *Dolby Digital Professional Encoding Manual* and *Dolby Surround Mixing Manual* available on the Dolby web site at www.dolby.com/tech/.

4.2 Incoming Quality Control

The facility should have a quiet audio QC room to assess the quality of incoming audio sources. Noisy electronic equipment, such as tape machines, should be placed in an area isolated from the listening space. Control and monitoring requirements should include muting for individual playback channels, overall gain control, metering, and an audio phase oscilloscope. Five identical full-range monitor speakers and a complementary subwoofer should be installed, each with its own amplifier.

In the short term, a playback system suitable for Dolby Surround is sufficient. However, acquiring the equipment necessary for 5.1-channel Dolby Digital monitoring, which will be required in the near future, is a more efficient use of financial resources because the DP562 Dolby Digital professional reference decoder includes a Dolby Surround Pro Logic decoder. In particular, a 5.1-channel Phase Scope is a wise purchase.

The Dolby DP562 decoder provides all current monitoring modes with provisions for future configurations. The unit has a bass management system that allows low frequency material to be played back through several combinations of speakers and subwoofers. For further information, please refer to the *Dolby Digital Professional Encoding Manual* available on the Dolby Web site at www.dolby.com/tech/.

4.3 Master Control

A 5.1-channel audio system should be installed in Master Control for proper audio monitoring. This allows the operator to verify operation of the system and the on-air audio. As with the QC room, acquisition of Dolby Digital equipment for the current implementation is a wise investment.

The DP562 Dolby Digital professional reference decoder is required to decode Dolby Digital and Dolby Surround signals, while the SDU4 Dolby Surround decoder only decodes Dolby Surround. For further information on Dolby Surround monitoring in Master Control, please refer to the *Technical Requirements for Master Control* document available on the Dolby web site at www.dolby.com/tech/. For further information on Dolby Digital monitoring, refer to the *Dolby Digital Professional Encoding Manual* on the Dolby web site.

Chapter 5 – Professional Products

5.1 Dolby Surround

This section describes Dolby Surround professional products manufactured by Dolby Laboratories. For further information on mixing in Dolby Surround please refer to the *Dolby Surround Mixing Manual* available on the Dolby web site at www.dolby.com/tech/.

5.1.1 Dolby SEU4 Encoder

The SEU4 Dolby Surround encoder receives four analog audio input signals (left, right, center, and mono surround) and matrix encodes them into two analog audio output signals, Lt and Rt, for transmission and recording.

5.1.2 Dolby SDU4 Decoder

The SDU4 Dolby Surround decoder receives two matrix-encoded analog audio input signals, Lt/Rt, and decodes them into four analog audio outputs (left, right, center, and mono surround) using Dolby Surround Pro Logic decoding technology. The unit also provides switchable stereo and mono monitoring modes to evaluate mix compatibility. A master fader controls all four monitor output channels, allowing variations in overall listening level while maintaining channel-to-channel playback balance.

5.2 Dolby Digital

This section contains information on Dolby Digital professional products. For further information please refer to the *Dolby Digital Professional Encoding Manual* available on the Dolby web site at www.dolby.com/tech/.

5.2.1 Professional Encoders

Dolby Digital professional encoders are available as either hardware or software-based devices. Choosing an encoder depends on the application requirements. Hardware encoders operate in real-time, encoding the PCM audio input into a Dolby Digital stream with little process delay (milliseconds). A real-time encoder is ideally suited for live broadcast applications that require minimal delay for encoding and transmission. Hardware encoders also permit real-time output monitoring with a Dolby Digital decoder. Software encoders, although usually not real-time, are ideal for batch processing and encoding a series of PCM files or streams in a single session. It is likely that software encoders will soon approach the real-time performance of hardware encoders as the processing power of PCs increases.

Dolby Laboratories Encoders

Dolby Laboratories offers stand-alone hardware audio encoders, such as the Dolby DP567 (two-channel) and the Dolby DP561 (5.1-channel), that are ideally suited for broadcast applications that require real-time Dolby Digital encoding. For further information on Dolby manufactured encoders, please contact Dolby Laboratories as described in *Section 6.2 - Contacting Dolby Laboratories*.

Dolby Laboratories Licensed Encoders

Hardware encoders are available from Dolby licensees as PCI cards and integrated audio/video encoders. Software encoders have been developed by Dolby licensees for platforms such as SGI, Apple Macintosh, and Windows-based PCs.

5.2.2 Professional Decoders

Consumer decoder products have many exciting and useful features but often lack the functionality required by a professional production environment. For example, real-time monitoring of Dolby Digital stream parameters is critical when encoding and monitoring audio material. Another essential feature of a professional decoder is the capability to emulate any type of decoder on the market. Since many consumer decoders have some of their decoding modes (dynamic range compression, down-mixing, etc.) preset at the factory, it is important for the encoding engineer to emulate that type of decoder in the production studio. Professional decoders typically feature a rack-mount style chassis and professional input/output connections (XLR, AES/EBU).

Dolby DP562 Professional Reference Decoder

The DP562 Dolby Digital professional reference decoder monitors Dolby Digital bitstream parameters in real-time while logging errors or faults for quality assurance. The Dolby DP562 can emulate any Dolby Digital consumer decoder in addition to decoding Dolby Surround material using a Dolby Surround Pro Logic decoder. For further information on the Dolby DP562, please contact Dolby Laboratories as described in *Section 6.2 - Contacting Dolby Laboratories*.

Dolby Laboratories Licensed Professional Decoders

Two types of licensed professional decoders are currently available: Broadcast baseband decoders, which are suited for fixed-mode broadcast monitoring, and confidence decoders, which are limited in features but are integrated with Dolby Digital encoders.

Chapter 6 – Miscellaneous Information

6.1 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 may also be obtained by sending e-mail to info@dolby.com with a description of the desired documents and a complete mailing address.

Dolby also has a staff of engineers who can assist with soundtrack production, encoding, implementation, and trademark use questions. Dolby engineers are also available to provide on-site assistance with room configuration and calibration, soundtrack 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 given in the following section.

6.2 Contacting Dolby Laboratories

You may contact Dolby from anywhere in the world by e-mail using the following addresses:

Address	Use
hdtv@dolby.com	Questions on HDTV audio production, encoding, and implementation issues
tsa@dolby.com	To apply for a Dolby trademark agreement (TSA)
info@dolby.com	General information and inquiries

6.2.1 General information and inquiries

In addition to its headquarters in San Francisco, Dolby has offices around the world equipped to provide information on soundtrack production and encoding. Use the following list to contact the nearest Dolby office for direct assistance.

A wide variety of technical and trademark information can be found on the Dolby web site at www.dolby.com.

San Francisco

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

Los Angeles

Dolby Laboratories
3375 Barham Boulevard
Los Angeles, CA 90068-1446
Telephone 213-845-1880
Facsimile 213-845-1890

New York

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

United Kingdom

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

Shanghai Liaison Office

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
e-mail dolbysho@public2.sta.net.cn

Japan Liaison 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

6.3 Trademark Usage

Dolby Laboratories encourages the use of Dolby trademarks to identify soundtracks that have been encoded using Dolby technologies. This is an effective way to inform listeners of the soundtrack format, and the use of standard logos promotes easy recognition in the marketplace. Like any trademark, however, Dolby logos may not be used without permission. Dolby Laboratories therefore provides a royalty-free Trademark and Standardization Agreement (TSA) for companies who wish to use Dolby trademarks. This agreement must be signed by the company that owns the program material being produced. 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, however, that these facilities refer their clients to us for trademark licensing information.

If you would like to use the Dolby Digital or Dolby Surround logos you can apply for a Dolby Trademark and Standardization Agreement (TSA) by sending e-mail to tsa@dolby.com or by contacting Dolby Laboratories at any of the locations given in *Section 6.2 - Contacting Dolby Laboratories*. When sending written requests please indicate the type of trademark license you would like (Dolby Digital and/or Dolby Surround) and include your name, your company name, mailing address, and the type of media that your soundtracks will be distributed on (such as DTV broadcast, DVD, CD, 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 coming months 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.

Appendix A – Basics of Perceptual Coding

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, which can mask other frequencies. Masking is the decreased audibility of one sound due to the dominant presence of another sound in a nearby frequency. Refer to Figure A-1. The degree of masking and resulting audibility depends on the amplitude of the dominant signal and its relationship to other frequency components of the waveform.

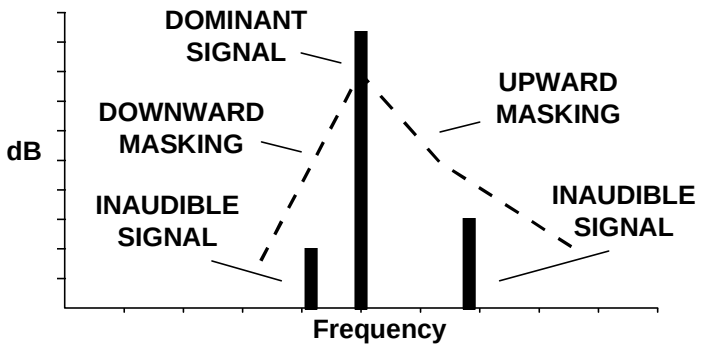

Figure A-1 The Effect of Masking

Dolby Digital takes full advantage of auditory masking by imitating the operation of the human ear. Dolby Digital encoders divide the audio spectrum into narrow frequency bands and groups them with respect to the critical bands. Using an auditory masking model, bits are allocated for each channel from a common bit pool. Coding (quantization) noise is minimized by constraining it 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-2.

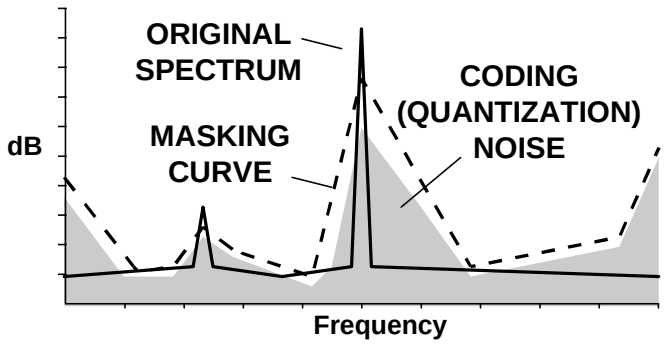

Figure A-2 Coding (quantization) Noise Below the Masking Curve

Dolby Digital is an extremely efficient coding system intended for use in the final emission link to the consumer, where operation at a reduced data rate is needed and additional coding generations are not normally necessary. Higher quality coding is essential in distribution because there may be many generations of encoding/decoding and the signal may be processed between generations. While a decoded signal may sound fine, certain types of signal processing may reveal coding artifacts that are otherwise inaudible. For further information, please refer to *Section 1.4 - Audio Coding in the Digital Infrastructure*. For further information on Dolby Digital please refer to *Appendix D – Bibliography* at the end of this document and the technical papers available on the Dolby web site at www.dolby.com/tech/.

Appendix B – Metadata

B.1 Metadata

Metadata is non-audio information that accompanies the audio in the Dolby Digital bitstream, and is used by the decoder to control the level and/or dynamics of the reproduced audio. The actual Dolby Digital audio information carries the signals fed into the encoder with no changes in level or dynamics; such changes may be introduced only at the moment of decoding and, in general, at the option of the listener. For further information, please refer to the *Dolby Digital Professional Encoding Manual* and the ATSC document A/52, *Digital Audio Compression Standard (AC-3)*, available on the Dolby web site at www.dolby.com/tech/.

Audio recording and transmission media are normally fully modulated to make best use of the available dynamic range. In other words, the only fundamental reference is the maximum level of the medium; lower references are merely pre-determined amounts below "full-scale." As a result, signals with a small peak-to-average ratio and peaks near full-scale have a high average and sound very loud, while signals with a wider dynamic range and a larger peak-to-average ratio, have a lower average and are comparatively quiet much of the time. Different items normalized to the same peak level demand different volume settings by the listener.

Broadcasts typically apply substantial compression both to restrict the material's dynamic range and, more significantly, to reduce the peak-to-average ratio to normalize the volume. Unfortunately, all listeners receive this processed signal. Dolby Digital is the first audio medium to address this variation in subjective volume without requiring a reduction in dynamic range. The audio bitstreams in Dolby Digital are fully modulated in accordance with conventional practice but are not compressed. The audio reproduced from the decoder is normalized to a typical average level and can, at the discretion of the listener, either remain uncompressed or undergo a variable amount of compression.

There are three types of metadata to control the reproduced average volume and dynamics: *dialnorm*, *dynrng* and *compr*.

B.1.1 *Dialnorm* - Dialogue Normalization

Dialnorm is a control word conveyed once per audio block (5 ms) that adjusts the volume setting within the decoder. *Dialnorm* is permanently enabled in all consumer decoders and ensures that program items containing speech at different levels (with respect to digital full-scale) emerge at a more constant acoustic level. *Dialnorm* can also be used to ensure that music with a wide dynamic range (and usually a lower average level) does not emerge subjectively much quieter than music with a narrow dynamic range.

It is not intended that *dialnorm* be changed within a program segment but rather that it balance the relative reproduced levels of different segments without resetting the volume control. In the broadcast environment, it is preferable to generate *dialnorm* during post-production, convey it in parallel with the audio (via the metadata distribution path) to the emissions encoder, and incorporate it into the Dolby Digital bitstream. In some applications, however, it might be necessary to set *dialnorm* manually at the emissions encoder.

Since *dialnorm* is a measure of the average loudness of a program item and is used as a reference for compression thresholds, *dynrng* and *compr* cannot be generated without an appropriate setting for *dialnorm*.

B.1.2 *Dynrng* - Dynamic Range Compression

Dynrng is a control word occurring once per audio block that can reduce the dynamic range of the reproduced audio when the decoder is in the line out mode. It is active within program segments and may often be generated to produce a result similar to the audio dynamics processors used in current analog broadcasting facilities. As with *dialnorm*, it is ideally generated during post-production, providing the sound engineer with an opportunity to audition and modify the effect. In this case, the data must be conveyed to the emissions encoder in parallel with the audio.

At the point of *dynrng* generation (post-production or at the emissions encoder), the user can choose between several compression “profiles” intended for different types of program material (movies, wide-range music, predominantly speech, etc.). In all cases, levels within a region near the *dialnorm* setting, referred to as the “null zone”, are unaffected while excursions above or below the region are reduced. The null zone represents the average amplitude level for a particular program type. Proper compression, therefore, depends on setting *dialnorm* correctly.

Many consumer decoders allow the listener to employ *dynrng* fully, partially or not at all. When fully applied, the output signal has restricted dynamics similar to those of analog broadcasting. When *dynrng* is disabled, the output dynamics are not modified and remain identical with those presented to the emissions encoder. Partially applying *dynrng* results in an intermediate modification to the dynamic range.

B.1.3 *Compr* - Compression Control

Compr is a control word that occurs once per audio frame (32 ms) that can modify the dynamic range of the reproduced audio when the decoder is in RF Mode. RF Mode provides one- or two-channel outputs intended primarily to drive an RF modulator that feeds a television set or video cassette recorder. *Compr* is active within program segments and is usually generated to produce a dynamic range similar to that of current analog broadcasting. *Compr* is ideally generated during post-production and requires a distinct path through Contribution and Distribution to the emissions encoder.

The user can choose between several compression “profiles” for *compr*, often identical to those offered for *dynrng*. In consumer decoders operating in RF Mode, *compr* is generally enabled.

B.2 Overload Protection via *dynrng* and *compr*

Dolby Digital is a multichannel medium designed to permit reproduction in many formats, most requiring downmixing in the decoder. Satisfactory loudness matching between programs with different numbers of source channels, and between decoders using different output Modes, demands the use of mixing coefficients that occasionally cause decoder overload. The Dolby Digital encoder assesses this possibility by measuring the peak level of all downmixes available in decoders. When it detects a potential overload, the "artistic" compression settings of *dynrng* and *compr* are overridden in favor of protection limiting (gain reduction) to prevent decoder overload. By correctly setting *dialnorm*, protection limiting should occur very rarely for Line Mode (gain reduction via *dynrng*). Protection limiting may be more common in RF Mode (gain reduction via *compr*), just as conventional analog broadcasts demand frequent limiting to prevent transmitter overload.

B.3 Likely Values for *dialnorm*

The dialogue output level from a Dolby Digital decoder after the application of *dialnorm* is standardized at -31 dBFS Leq. Leq, equivalent loudness, is a measure of subjective loudness expressed here in dB relative to digital full-scale. In general, listeners are likely to prefer that the average level of other program types be similar (within a few dB) to that of speech. A program source with a dialogue Leq 31 dB below full-scale requires no adjustment of the reproduced volume; *dialnorm* would be set to -31. This might occur in a movie soundtrack with unusually wide dynamic range.

For fully modulated speech (i.e., reaching digital full-scale), such as a news bulletin, the equivalent loudness level might be about -15 dBFS. In this case, *dialnorm* should be set to -15, providing a 16 dB attenuation in the decoder and yielding the standard output level for movie dialogue of -31 dBFS Leq.

Typical broadcast source material, often with a limited dynamic range and approaching digital full-scale, generally requires *dialnorm* values between -10 and -20. Only programming with a wide dynamic range, such as movie soundtracks and symphonic music, requires *dialnorm* values between -20 and -31.

B.4 Choosing *dialnorm* for a Program Item

B.4.1 Speech

For items containing speech, there are several methods for determining the required value for *dialnorm*. Since the goal is to arrive at a subjective match between different sources, most methods involve comparative listening; there is no exact answer, only what sounds about right!

- Measure the level of passages which consist primarily of speech using an equivalent loudness meter (Leq, a long-term rms of the A-weighted audio) and express the answer in dB with respect to full-scale. The result is the correct *dialnorm* setting.
- Play a reference recording of speech that has previously been normalized to the standard level with respect to full-scale and adjust the *dialnorm* until the new item subjectively matches the reference without changing the monitoring gains.

- Set up a chain including a Dolby Digital encoder and a decoder operating in RF Mode feeding a television set. Start with a *dialnorm* setting of -10 and adjust negatively (i.e., -11, -12, etc.) until your channel's speech level matches the dialogue level of conventionally processed broadcast channels. It is essential to approach from the correct direction to avoid confusion caused by gain reductions due to overload protection limiting. This method is appropriate for a permanent channel transmitting speech programming that is reasonably consistent at the source.

B.4.2 Music

When music is the most significant program content, it is necessary to estimate how a listener chooses to hear music when the volume control was set for a preferred speech level. While there is no correct value, it is clearly possible to set *dialnorm* to prevent the listener from rushing for the volume control. The second method above will work. More objective methods will be developed in the future.

Appendix C – Dolby Digital Encoding Parameters

A summary of Dolby Digital professional encoding parameters follows. For detailed information, refer to the *Dolby Digital Professional Encoding Manual* available on the Dolby web site at www.dolby.com/tech/.

Parameter Default Values

Although Dolby recommends default values for licensed encoders, parameter settings usually require adjustment for specific content and applications, such as broadcast.

In addition to the Data Rate and Audio Coding Mode, the user should pay particular attention to Dialogue Normalization, Dynamic Range Compression, Surround Channel 3 dB Attenuation, and 90° Phase-Shift.

Audio Service Configuration

The parameters in this group specify the fundamental aspects of the Dolby Digital encoded program. They include the Data Rate, Sampling Frequency, number of coded channels, Audio Coding Mode, the type of audio service, Bitstream Mode, and the Dialogue Normalization value.

Bitstream Information

The parameters in this group relate directly to the Bitstream Information (BSI) fields. They include Center Downmix Level, Surround Downmix Level, Dolby Surround Mode, Language Code, Audio Production Information Exists, Mixing Level, Room Type, Copyright Bit, and Original Bitstream.

Preprocessing Options

These parameters offer options to precondition audio input signals before they are encoded. They include Digital De-Emphasis, DC Filter, Channel Bandwidth Lowpass Filter, LFE Lowpass Filter, Surround Channel 90° Phase-Shift, Surround Channel 3 dB Attenuation, Dynamic Range Compression Preset, and RF Pre-Emphasis Filter.

Automatic Parameters

These parameters should be adjusted automatically by the encoder. They include Audio Bandwidth and Coupling.

Time Code Control

SMPTE time code values are not a required part of an elementary bitstream. Typically, MPEG system elements are used for audio/video synchronization during decoding (presentation time stamp, PTS, for example). Currently, there is no known use for the syntactic time code elements, which are optional in the BSI portion of the elementary Dolby Digital stream. Instead, the time stamp values associated with each Dolby Digital synchronization frame can be outside of the elementary stream.

Appendix D – Bibliography

Dolby Laboratories, “Dolby Digital Professional Encoding Manual,” available on the Dolby web site at www.dolby.com/tech/

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

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

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

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

Grant A. Davidson, Louis D. Fielder, Brian D. Link, “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/

Craig C. Todd, Grant A. Davidson, Mark F. Davis, Louis D. Fielder, Brian D. Link, Steve Vernon, “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/

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

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

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

“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.