ORIGINAL

ORIGINAL FILE

RECEIVED

NOV 18 1987

Federal Communications Commission
Office of the Secretary

Before the
FEDERAL COMMUNICATIONS COMMISSION
Washington, DC 20554

In the Matter of	)	
	)	
Advanced Television Systems and Their Impact on the Existing Television Broadcast Service	)	MM Docket No. 87-268

COMMENTS OF DOLBY LABORATORIES

Dolby Laboratories ("Dolby") submits these comments in response to the Commission's Notice of Inquiry ("NOI") in the above-captioned proceeding, FCC 87-246, 2 FCC Rcd 5125 (1987).

TABLE OF CONTENTS

Page

Summary of Comments 2

Dolby and its Interest in This Proceeding 2

The Broadcast Television Service Must be Allowed to Incorporate Improved Sound 3

Digital Sound Using Delta Modulation Techniques is Both Technically Excellent and Cost Effective 6

Dolby's ADM Sound Can be Added to Today's NTSC Signal Without Perceptible Degradation or Interference 10

Dolby's Delta Modulation Approach Can be Incorporated into ATV 12

Conclusion 13

0+11

Summary of Comments

The Commission must consider improved television sound as it examines systems for improved video. Television sound should be improved to a quality level comparable to digital compact discs. The Dolby digital sound approach, which employs adaptive delta modulation, produces sound that is perceived to be the quality of digital compact discs, and yet the consumer sound decoder is very low in cost. The Dolby digital approach is already in use in several countries instead of analog sound. Moreover, the Dolby sound system may be compatible with the current NTSC broadcast signal format, since we believe it can be added to the NTSC broadcast signal without perceptible degradation or interference. Most importantly, it can be incorporated into any advanced television system format by means of either frequency division multiplexing or time division multiplexing.

Dolby and its Interest in This Proceeding

Dolby Laboratories was founded in 1965 by American engineer and physicist Ray Dolby, with the aim of developing practical noise reduction systems for improving sound quality in a variety of professional and consumer environments. Today, Dolby noise reduction systems are used throughout the world for professional and consumer tape recording, and for motion picture sound. Research and development continues in both analog and digital audio.

Dolby Laboratories' corporate goal is to create processes and products which represent genuine long-term improvements in sound quality.

The advanced television technology standards that are the subject of this inquiry are a matter of great interest to us. We firmly believe that improved television sound quality is an essential element of advanced television technology. We hope that our experience and expertise in sound technology can assist the Commission in reaching decisions that best serve the public interest.

The Broadcast Television Service Must be Allowed to Incorporate Improved Sound

The current BTSC analog sound system used with televsion broadcasting[1], while theoretically capable of excellent performance, performs poorly in practice. Because of the practical problems of inter-carrier analog sound detection, there are numerous "birdie" and "buzz" products in analog television sound receivers. The resulting sound quality is inferior to most other common consumer audio formats: compact disc, compact cassette, FM broadcasting, VCR, etc.

---

[1]BTSC multichannel television sound was approved in 1984. Second Report and Order in Docket No. 21323, 49 Fed. Reg. 18100; EIA Television Systems Bulletin No. 5, "Multichannel Television Sound: BTSC Recommended Practices (July 1985).

The relatively poor quality of television sound is tolerated since the sound is rarely reproduced by anything resembling a "Hi-Fi" system. This is changing since the introduction of the BTSC stereo system. Better equipment is being used to listen to TV sound. Programmers are beginning to produce better quality sound tracks. Many films which are available for television feature Dolby Stereo soundtracks which are very high quality. Since stereo TV sound is still a novelty, the quality of the analog sound has not yet become an issue.

It is our belief that the delivery of very high quality sound with the television picture will deliver as much or more enjoyment to the consumer as the delivery of a higher resolution picture.[2] For this to be the case, it is necessary that the sound be of excellent quality both in production and reproduction, and that the sound not be degraded by transmission. It is essential that the broadcaster be allowed to transmit a very high quality sound signal in order to remain competitive with other media. The transmission link should not significantly degrade the sound.

---

[2]Dolby has undertaken market research studies indicating that improved sound is desired by consumers, and than more than half of those interviewed would be willing to pay more for receivers with improved sound. Video Magazine's 1986 subscriber survey showed that improved sound was important to many subscribers in buying their next VCR. A BASF consumer research study showed that the Dolby name, which signifies improved sound quality, is important to a majority of consumers in buying pre-recorded audio cassettes. A survey by the Newspaper Advertising Bureau concludes that improved sound at movie theaters is part of the reason that people prefer the movie theater "experience" to television.

We are not alone in believing that improved sound quality is needed. At the October 8, 1987 hearing of the House Telecommunications Subcommittee, a number of witnesses testified to the need for compact-disc quality digital sound as part of a high definition television system.[3]

The Public Broadcasting Service has also testified to the need for improved television sound.

> PBS, like other broadcasters, has long recognized that a weak link in the television program distribution system is the poor quality of sound reproduction.[4]

For this reason, PBS is planning to distribute digital stereo sound to its affiliates using satellite scrambling technology.

The National Broadcasting Corporation has also recognized the benefits of a digital audio transmission system.

> Digital studio delivery to the home will give the consumer the ultimate in audio quality from the television set.[5]

---

[3]See, for example, Statement of North American Philips Corporation at p. 5.
[4]Statement of Michael E. Hobbs, Senior Vice President, Policy and Planning, PBS, before the Senate Subcommittee on Communications, July 31, 1987, at p. 3.
[5]R. Hoffner, "Trends in Television Audio--Now and in the Future", 1987 NAB Engineering Conference Proceedings, p. 201.

It would be a tragic mistake for the Commission to adopt an improved video format without considering the need for improved sound. Improved sound with digital techniques is readily feasible at relatively low cost, and should become a part of the television broadcast signal.

Digital Sound Using Delta Modulation Techniques is Both Technically Excellent and Cost Effective

The Dolby digital audio coding approach employs adaptive delta modulation ("ADM") to produce high audio quality at relatively low data rates and low cost.[6] There are other digital sound techniques available for broadcast use, but these are either more expensive, or require higher data transmission rates, or have other undesirable characteristics.

For example, digital broadcast sound using pulse-code modulation ("PCM") can produce better quality sound than any analog approach. However, PCM requires relatively high data rates (350-1000 kilobits/sec) and is notorious for objectionable sounds that result from even small numbers of errors in the transmitted data.

---

[6]C. Todd and K. J. Gundry, "A Digital Audio System for Broadcast and Pre-Recorded Media", 75th AES Convention, preprint no. 2071, March 1984; C. Todd, "Efficient Digital Audio Coding and Transmission Systems", 39th NAB Broadcast Engineering Conference, 1985. In order to assure that the Commission has a complete record, we are attaching to these comments the papers cited in footnotes 6, 7, 11, 12 and 15.

Compared with PCM, delta modulation offers improved signal-to-noise ratio at low data rates, and does not require high precision components in the design and manufacture of decoders. Moreover, isolated bit errors have only a minor audible effect.

Companding techniques can be used to reduce the channel capacity needed for transmission of a sound signal. In other words, it permits more efficient use of the spectrum. Companding can be either analog or digital. Digital companding combined with delta modulation is referred to as adaptive delta modulation.

Digital companding may be susceptible to modulation of the background noise level by the audio signal. Traditional techniques for controlling noise in adaptive delta modulation are a compromise between speed of response in handling transient signals, and tolerance of errors that produce momentary deviations in gain.

The Dolby approach improves over these traditional techniques by sending a noise reduction control signal via a separate low data rate bit stream. Use of the control signal eliminates the gain deviations that are present in the traditional approach.

Using this approach, a high quality audio channel can be transmitted in a data stream of 200-300 kilobits/sec. This is far less than the 300-1000

kbps that would be needed for a PCM-encoded sound channel. Even lower data rates could be achieved, using more complex coding methods, but this would increase the equipment costs substantially.

Not only is the Dolby ADM approach as good as or better than PCM, it is perceptibly as good as any digital sound approach can be. Other digital sound approaches, using higher data rates or different, more expensive coding schemes, can reproduce sound with a quality that is measurably somewhat better in the laboratory. However, to the listener, the Dolby ADM approach is perceptibly equivalent to these. To the human ear, it is virtually perfection.[7]

For this reason, we do not believe that there is the likelihood of significant improvement in sound technology.[8] For a single-hop transmission path to the home, it might be possible to code equivalent audio quality into a lower data rate than with ADM, but this would greatly increase costs with no perceptible increase in quality.

While most advanced video approaches are still in the development stage, availability is not a problem for the Dolby audio coding approach.[9] The

---

[7]This is seen quite clearly in Figure 12 of Annex V, CCIR Rep. 632-3 ("Subjective Test Results on Adaptive Delta Modulation for High Quality Sound Coding"). The data points plotted in columns A and B are the Dolby ADM-encoded signals and the original source signals. They are perceived to be identical. It is only when data errors are introduced into the Dolby ADM signal that the listeners perceive a difference.

[8]NOI, para. 84.

[9]NOI, para. 40.

technology is relatively mature, and has been in use for several years throughout the world. It is an integral part of the B-MAC system used for private satellite networks in the U.S. and other countries.[10] It has been adopted for use in Australia as part of the satellite distribution service of the Australian Broadcasting Network.[11] It has been tested for broadcast use in Great Britain, Sweden and Finland. Unlike the U.S., these countries rejected analog stereo sound for broadcast television and instead adopted digital PCM sound. It has been tested in several other countries, and tests of compatibility with NTSC (discussed below) have been performed in the United States.[12]

In addition, the Dolby audio coding technique is extremely cost-effective.[13] The decoding circuitry is inherently simple and low cost (a stereo decoder integrated circuit is already available), and little precision is required in any of the components. We estimate that the digital decoding circuitry, in mass production, would add only about $10 to the manufacturing cost of TV sets. This is comparable to the cost of analog stereo circuitry, and much less

---

[10]See K. Lucas, "B-MAC and HDTV - Does it Fit?", Proceedings of the Third International Colloquium on Advanced Television Systems: HDTV '87, at p. 4.3.2

[11]See Annex V, CCIR Rep. 632-3 ("Subjective Test Results on Adaptive Delta Modulation for High Quality Sound Coding") and CCIR Rep. 953-1 ("Digital Coding for the Emission of High-Quality Sound Signals in Satellite Broadcasting"), Recommendations and Reports of the CCIR, Volumes X and XI, part 2 (16th Plenary, Dobrovnik, 1986)

[12]C. Todd, "A Compatible Digital Audio Format for Broadcast and Cable Television", CE-33 IEEE Transactions on Consumer Electronics, p. 297 (1987).

[13]NOI, para. 40.

than other digital audio approaches. This low cost is achieved by incorporating the complexity in the encoding equipment, of which only a few units are needed.

Dolby's ADM Can be Added to Today's NTSC Signal Without Perceptible Degradation or Interference

We believe the Dolby ADM sound signal can be added to the current NTSC video channel without causing perceptible degradation to the video or increasing interference to other broadcast licensees.[14] This means that it could be implemented today, on an experimental basis, while this proceeding evolves toward a final decision on standards.

Based on work done in Sweden and other countries, we propose that the ADM sound signal can be carried as a 512 kpbs signal (which includes some extra housekeeping and auxiliary data) that is modulated as a differential QPSK carrier, 4.85 MHz above the visual carrier.[15] The carrier level can be -20 dB relative to the visual carrier.[16]

---

[14]See the Todd paper cited in footnote 12, above.

[15]While Sweden uses a 7 MHz channel spacing as compared with 6 MHz in this country, the analog sound carrier is spaced 1.5 MHz below the visual carrier of the upper adjacent TV channel. Since this adjacent channel spacing is identical to the adjacent channel spacing in the U.S., we believe that the testing done in Sweden applies equally to the U.S. See A. Nyberg, "Digital Multi-Channel Sound for Television", 1987 IEEE Int'l Conf. on Consumer Electronics at p. 92.

[16]The 512 kbps signal could as well be used by the broadcaster for any purpose other than sound, if desired.

We have tested a number of carrier frequencies and a number of levels relative to the visual carrier. At a carrier frequency of 4.85 MHz, the worst-case degradation[17] to the audio SAP channel is about 9 dB. This adds noise to the SAP channel that is audible under ideal laboratory conditions, but probably not noticeable under practical conditions. At this carrier frequency, there is less than 1 dB noise increase in the BTSC stereo sound. Placing the digital sound carrier at 4.90 MHz would virtually eliminate the effect on the SAP channel.[18]

A digital sound approach can also be implemented in a cable TV environment, with adjacent channel operations on all channels. Based on tests done in Europe, the digital QPSK signal does not appear to cause interference into the adjacent visual carrier, because its noise-like spectrum does not create any patterning on the screen. Interference from the adjacent channel visual carrier into the digital sound does not seem to cause data errors, but it does reduce the margin against error.

---

[17]The worst case occurs with a high RF signal level (20 mVRMS), a very good receiver, and no video modulation.

[18]However, it should be noted that the upper edge of the TV channel is at 4.75 MHz above the visual carrier. Consequently, this carrier at 4.85 MHz is slightly outside the allocated channel and there is the possibility that it might result in some interference to non-TV operations adjacent to channels 4, 6, 13 and 69, and a carrier at 4.90 MHz could result in somewhat more interference. It is for this reason that we suggest implementation on an experimental basis.

Similarly, sync suppression scrambling, with sync information carried by pulse amplitude modulation of the 4.5 MHz sound carrier, does not seem to produce data errors but does reduce the margin for error. We believe that some additional study is needed on this point, however.

Dolby's Delta Modulation Approach Can be Incorporated into ATV

Advanced television systems can incorporate improved sound in a number of ways. For example, the sound can be added as a subcarrier above the visual carrier (as with NTSC), or it can be incorporated into the horizontal blanking interval of the video (as with B-MAC).

Based on the testing referenced above and our experience with the ADM approach, we can say that the Dolby ADM sound signal can be incorporated into virtually any advanced television or high definition television signal format that might be proposed. The Dolby ADM sound system can be either frequency-multiplexed or time-multiplexed into the TV signal.

There does not appear to be any need for the Commission to provide any special protection for ADM sound as was done for BTSC sound. The ADM decoder looks for a very special bit pattern for synchronization, before it unmutes the TV stereo sound. In contrast, the BTSC decoder looks merely for a pilot subcarrier, and could unmute improperly without the protection granted by the Commission.[19]

---

[19] 47 C.F.R. 73.682(c)

Conclusion

As the Commission proceeds to consider the technical standards for the next generation of TV broadcasting, it must keep in mind the need to improve sound as well as video. The marketplace success of digital compact discs has shown that the public appreciates improved sound quality in consumer electronics. It would be a tragic mistake for the industry if the Commission were to adopt an improved video format for television, but retain the inferior analog sound standard. This would be particularly ironic when improved digital sound techniques are already being implemented throughout the world while improved video techniques are still being studied in the laboratories.

Respectfully submitted,

David P. Robinson
Vice President, Engineering
Dolby Laboratories
100 Potrero Street
San Francisco, CA 94103
415-558-0200

Jeffrey Krauss
Consultant
15200 Shady Grove Road, Suite 450
Rockville, MD 20850

Date: November 18, 1987

PROCEEDINGS

39TH ANNUAL BROADCAST ENGINEERING CONFERENCE

NATIONAL ASSOCIATION OF BROADCASTERS
LAS VEGAS, NEVADA
1985

EFFICIENT DIGITAL AUDIO CODING & TRANSMISSION SYSTEMS

Craig C. Todd

Dolby Laboratories, Inc.

San Francisco and London

0. ABSTRACT

A digital audio coding method is described which is particularly suited to the requirements of the broadcast industry. High audio quality is achieved at relatively low bit rates (200 k bits/sec to 350 k bits/sec). Decoder circuitry is simple so receiver cost is minimized (important in point to multipoint transmission). The system is robust, and with increasing error rates, degrades gracefully and gradually. A simple error concealment scheme further reduces the audibility of errors.

1. INTRODUCTION

Digital techniques have a number of advantages for the distribution of audio signals. A primary one is that the channel coding can be considered independently of the audio coding so a great deal of flexibility exists to optimize these two items independently, which is not the case with practical analog techniques. Digital audio can be encrypted with no effect on audio quality, which is not possible with analog techniques. Given sufficient channel capacity and funding, arbitrarily high quality audio may be transmitted with conventional straightforward A-D, D-A conversion techniques and error correction systems. This paper sets forth an audio coding method which has a number of advantages over conventional techniques. These include: cost of circuitry at the recieving end of the transmission; required channel capacity to achieve perceptibly excellent audio quality; and performance under impaired reception conditions.

Note: Some of the material in this paper is derived from a paper presented at the 75th convention of the Audio Engineering Society (Paris, March 1984) and is used with the permission of the AES.

2. CONVENTIONAL APPROACHES

Receiving equipment must provide demodulation, demultiplexing, decrypting (if program security must by provided), error correction/concealment, digital to audio conversion, and output filtering. Most of the delivery systems proposed to date have employed Pulse-Code Modulation (PCM) for the audio conversion and require on the order of 350 k bits/sec to 1000 k bits/sec of data per audio channel. Well designed systems of this type can provide a quality of reproduction better than any available analogue distribution technique but they carry with them disadvantages. One of these is cost, and another is performance under impaired reception conditions.

In order for a PCM system to produce a dynamic range of 80 dB or more, circuitry is required which can produce an output with a relative tolerance of on the order of 0.01% and such precision is not inherently low cost. In order for a PCM system to operate at a reasonably low data rate it is necessary to reduce the sampling rate to only slightly greater than twice the desired audio bandwidth. This requires the use of relatively elaborate filters, which again are inherently not negligible in cost.

When reception conditions are not optimal, the received digital data will contain some errors. PCM systems are notorious for the extremely objectionable sounds that result from the reproduction of digital audio data containing even small numbers of errors. PCM systems are simply not viable without the addition of error correction or error concealment systems. With the use of such systems acceptable performance can be achieved under moderately adverse reception conditions, but when some some error rate is reached the system will degrade catastrophicly and the audio must be muted.

3. COMPANDED SYSTEMS

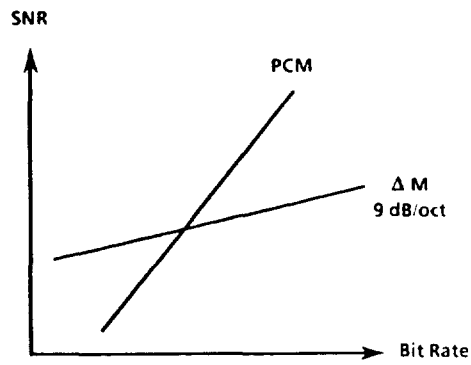
Any digital audio system employed for broadcasting can save channel capacity by employing companding techniques to reduce the required bit rate. Companding can increase the dynamic range of a low bit rate digital audio coding system. If the coding system inherently has adequate instantaneous signal to noise ratio but insufficient dynamic range, then simple companding systems can easily increase the dynamic range to whatever is required. If the SNR is not adequate the increased dynamic range will be accompanied by audible noise modulation.

Typical broadcast digital audio coding systems reduce bit rate to the point that the instantaneous SNR is not quite adequate (for example companding to 10 bit resolution) and increase the SNR by means of emphasis networks. Fixed emphasis networks typically reduce the high frequency noise modulation accompanying low frequency audio program material but actually increase the amount of low frequency noise modulation in the presence of high frequency signals. There is also a loss of high frequency headroom with fixed emphasis networks.

The perceived audio performance of a digital audio system operating at a low data rate will be dominated by the performance of the noise reduction system in use. The sophistication of the coder-decoder will lie in this area. The actual audio-digital-audio coder inside the noise reduction system can be chosen for some desirable characteristics and then a suitable noise reduction system placed around it.

4. DELTA MODULATION

Delta Modulation (DM) has some significant virtues which has lead us to consider it as the core of a digital coding system. DM circuitry does not require any precision components and can be manufactured very economically with today's technology. Since all bits have equal value, isolated bit errors always have a minor audible effect (there is no MSB to create a huge error impulse). As shown in Fig. 1, the inherent SNR of DM is a gradual function of bit rate (9 dB/octave) in sharp contrast to the steep slope of PCM SNR vs. bit rate (where the SNR expressed in dB doubles when the bit rate is doubled). From this simple illustration it is clear why PCM is so attractive for professional recording applications where the data rate is of little concern. Given adequate data rate the performance of PCM is theoretically superb. However, when the data rate is lowered, as in the case of an efficient broadcast system, it is apparent that the performance of PCM will become inferior to that of DM.



SNR of PCM, Delta Modulation vs Bit Rate

Figure 1.

Companded DM systems are typically refered to as Adaptive Delta Modulation (ADM) systems. In contrast to companded PCM, noise modulation in an ADM system is caused not by high amplitude signals but by high slope signals. Noise modulation is worst in the presence of high frequency signals but the fact that high frequency signals effectively mask noise makes this characteristic acceptable and perhaps preferable to that of companded PCM where high amplitude low frequency signals will produce noise modulation which may not be masked by the signals. The application of emphasis around a PCM system moves it's characteristics in the direction of a DM system by improving low frequency performance at the expense of high frequency performance.

A judicious choice of fixed emphasis characteristic applied around a low data rate ADM system can provide a good compromise between audible noise modulation with mid frequency signals and high frequency handling capability, but is unsatisfactory when program material contains predominately high frequency energy. The problem is not so much that the signals might overload the system (it is easy to provide adequate dynamic range), but that the low frequency noise which is boosted by the de-emphasis may not be sufficiently masked.

The usual technique for ADM is to use a fixed emphasis characteristic and to code into a single bit-stream both the audio information and the companding (often referred to as step-size) information. The adaption is usually output controlled (operating from the output bit stream) so transient performance is not perfect but may be acceptible. The significance of individual bits to the adaption control signal depends on the nature of the adaption algorithm. Some individual bits will become very significant to the algorithm and a small percentage of data transmission errors which happen to hit these critical gain control bits will cause the gain to significantly deviate from it's proper value. We refer to this as the gain blipping effect. The gain blipping effect may be the most noticeable degradation of such a system operating with data errors. The effect can be reduced in magnitude by making the control signal move more sluggishly at the expense of poorer transient response. The adaption characteristic is therefore a compromise between speed of response (necessary for handling transient signals) and tolerance of errors.

5. OUTLINE OF NEW SYSTEM

A novel approach to the application of delta-modulation to audio coding for broadcast will now be described. A fundamental aspect of this new approach is that the complexity of the encoder (of which few are required for point to multipoint distribution) has been raised in order to lower the cost of a decoder and to remove the performance limitations of a simple ADM system.

The dynamic range of simple DM is improved by means of wide band gain control as in the case of ADM. SNR is improved by the use of emphasis networks but the compromise inherent in the use of fixed networks is avoided by the use of a variation on our 'sliding band' variable emphasis system. This powerful technique gives a larger improvement in noise modulation that fixed emphasis yet does not incur a penalty of reduced high frequency headroom, or low frequency noise emphasis in the presence of predominantly high frequency program material.

The 'gain blipping' effect has been eliminated by sending the noise reduction control signals via a separate low data rate bit stream. A simple algorithm is used to convert this bit stream into a control signal of limited bandwidth and in this algorithm all bits have equal weight. It is thus impossible for a bit error to cause a very significant control signal deviation.

A low data rate control signal suggests a sluggish response which would yield poor transient performance. This is avoided by analyzing the input audio signal, generating the control signals, and delaying the audio main path, so that the control signals respond in advance of the arrival of the audio to be encoded.

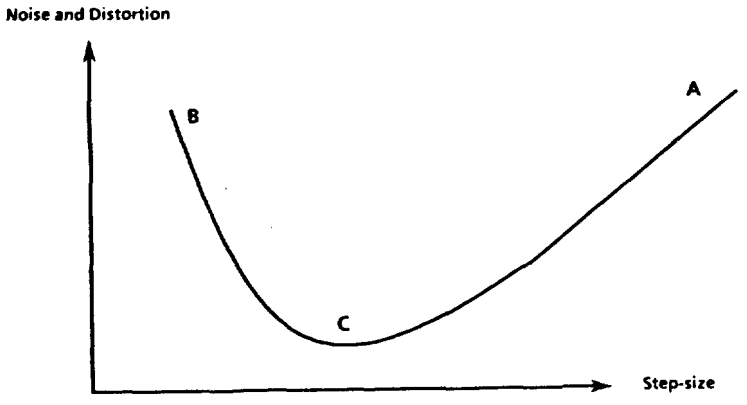
6. LIMITED BANDWIDTH CONTROL SIGNALS

The operation of changing the gain or frequency response of an audio signal in response to a control signal is essentially that of multiplication. When one waveform is multiplied by another, the output signal spectrum is the convolution of both input spectra. In audio signal processing, we consider that the original audio spectra has been contaminated by modulation products, and these products may be highly objectionable. In a complementary noise reduction signal processing system, the introduction of modulation sidebands in the encoder should be accompanied by a subtraction of the same sidebands in the decoder, which is sometimes theoretically possible.

In any practical system the encoder and decoder will not track perfectly because of component tolerances and/or channel errors. The modulation products generated by the signal processing will not be completely canceled. In a system where very fast responding control signals have been employed, discrepancies between encoding and decoding will leave remnants of modulation products splattered all over the audio spectrum; if the noise and distortion of the overall system are not to be audibly degraded, the accuracy of the control signals must be of the same order as that of the audio. In a system with very slowly responding control signals, the remnant spectral energy will be found only close to the original spectral energy and may be audibly masked by the original energy. Hence the slower the control signals, the less audible will be the modulation products of mistracking, or the greater amount of mistracking can be tolerated.

7. OPTIMIZED ADM

The noise and distortion emerging from an ADM codec depend on the audio input signal and the step-size in use; both of which are varying. Consider a codec handling a single sine wave. As a function of step-size, the output noise and distortion will vary as shown qualitatively in figure 2. In the region labelled A, the step-size is too large, which produces excessive quantizing noise. In region B the step-size is too small and the system is in slope overload which produces high noise and distortion. There is an optimum value of step-size for the particular input condition labelled C. For each short time segment of real audio there is a curve like fig. 2, and an optimum step-size. In a conventional output controlled ADM system the step size rarely achieves the optimum value, but remains in region A most of the time, moving into region B on signal transients. Our optimized ADM system essentially always operates in region C, so the core DM is always fully loaded. This is possible because the step-size determination is done in the encoder and is input controlled.



A. Excess noise because system not fully loaded.
B. System in slope overload
C. Optimum operation

Optimum step-size

Figure 2.

8. VARIABLE EMPHASIS

As mentioned above, the use of conventional high frequency emphasis is effective when the predominant spectral components of the input signal are at low or middle frequencies. However with high frequency signals this actually degrades performance. Consideration of the noise levels and spectra delivered by an optimized ADM codec operating at 200 or 300 k bits/sec shows that there are three (overlapping) regimes to be considered if noise shaping is to be effective in eliminating audible noise modulation.

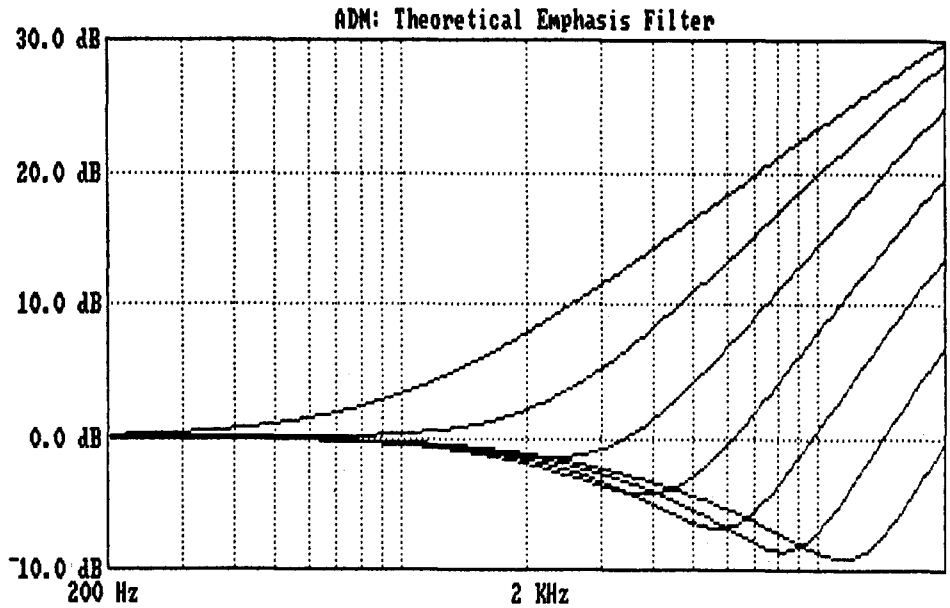


Figure 3: Emphasis Curves.

A. When the predominant audio spectral components lie below roughly 500 Hz, a large high frequency pre- and de-emphasis will reduce noise sufficiently that audible noise modulation will not occur. An example of a practical emphasis characteristic is shown in curve 1 on figure 3.

B. As the predominant spectral component is increased up to 2 or 3 kHz, it is necessary to slide the emphasis upwards in frequency so as to retain high frequency noise reduction relative to the spectral component (curves 2 and 3). Low frequency noise is not yet an audible problem.

C. When the predominant spectral component is above about 3 kHz, noise both at low and at very high frequencies must be reduced. An emphasis curve with a dip at the predominant component will reduce the step-size and hence the broad-band noise emerging from the codec, while the subsequent complementary de-emphasis peak will pick out the wanted signal component, while attenuating high frequency noise and retaining the reduced low frequency noise level which resulted from the smaller step-size. For example, if the predominant signal lies at 6 kHz, curve 5 is a desirable emphasis characteristic.

This explanation has assumed that the predominant components of an audio signal at a particular moment are concentrated in a narrow region of the spectrum; such a signal is in fact the most critical case. When the spectral components are more distributed, their masking properties cover more of the noise, and the emphasis shape is less critical.

Thus a variable emphasis circuit giving a family of response curves of the form shown in figure 3 preceding the delta-modulator, with the complementary de-emphasis following the demodulator will provide efficient subjective noise reduction under all signal conditions, provided that circuitry can be designed to analyze the input audio and to give suitable instructions to the emphasis networks.

9. THE SIMPLE DECODER

Figure 4 illustrates the decoder required in the reciever. Each audio channel of decoding receives three data bit-streams. The audio data is at a moderately high rate, on the order of 200 to 300 k bits/sec. The two control bit streams are at much lower rates, typically about 8 k bits/sec each; for TV sound in sync type systems a convenient rate is half the television line frequency.

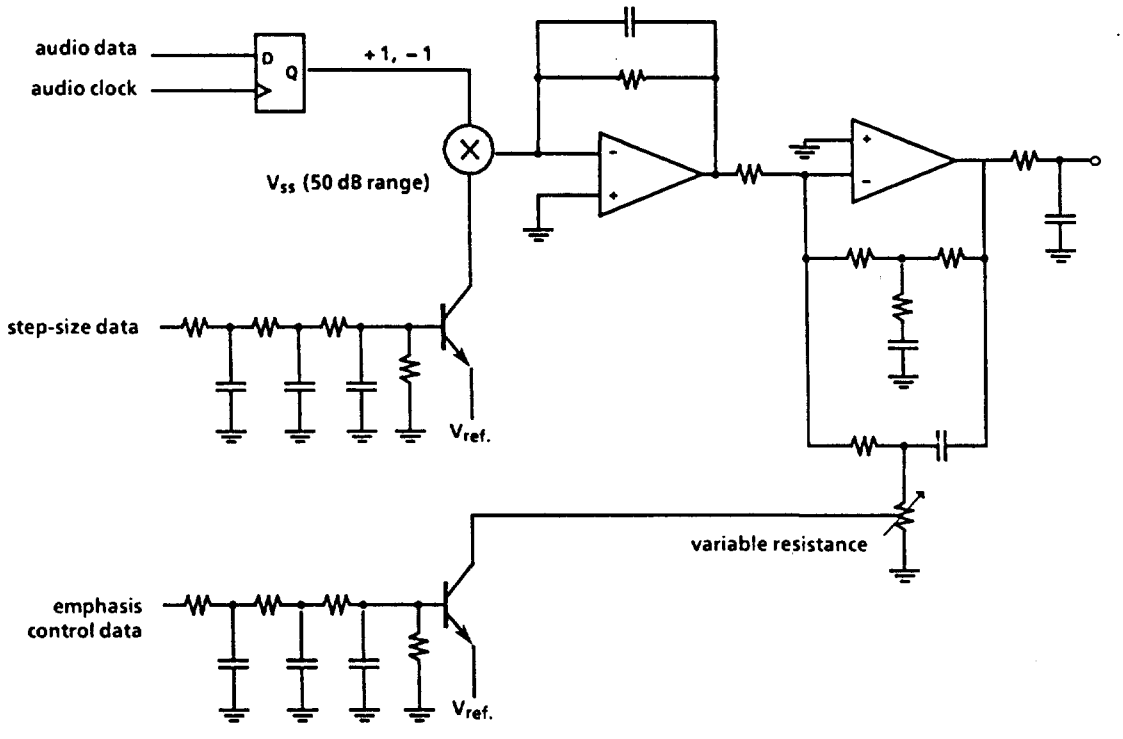


Figure 4: Consumer Decoder.

The product of the audio data and the step-size control signal is integrated in a leaky integrator. The step-size control bit-stream contains the logarithm of the required step-size coded as delta-sigma modulation. The control signal is recovered by low-pass filtering the bit stream and exponentiating the resulting voltage. The de-emphasis control signal is handled identically but instead of a variation in gain, it produces a variable pole frequency in the de-emphasis network. The final single pole of the de-emphasis network yields enough attenuation of clock and spurious out-of-band signals that no output filter is required.

The decoding circuitry is inherently simple and low cost and no precision is required in any of the components. Any component errors produce typically analog types of errors such as incorrect frequency response or level. There are no mechanisms (as in a PCM converter) for the creation of any intolerable sorts of non-linearities; there is little chance that tolerances or long term drift will produce objectionable sounds.

10. THE ENCODER

Figure 5 is a block diagram of the encoder required at the broadcast center. The encoder is more elaborate and expensive than the decoder; since only a few encoders are required in many broadcast systems this is acceptable. It is the sophistication of the encoder which allows such a simple decoder to produce such high audio quality. The fundamental requirement of the encoder is to deliver bit-streams which when decoded by the simple decoder yield the best perceived audio.

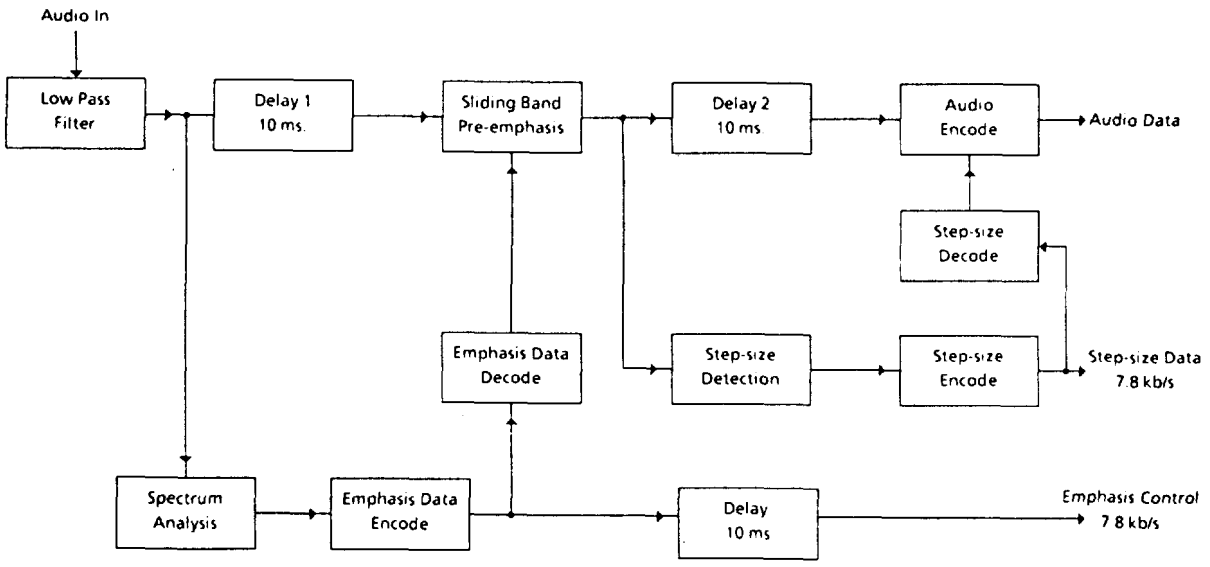


Figure 5: Professional Audio Encoder. Block Diagram.

The emphasis control block analyzes the spectrum of the input audio to determine the optimum emphasis characteristic to minimize the audibility of codec noise in the presence of that signal spectrum. This information is coded into the low bit rate emphasis control data stream. The conversion of this bit stream back to a signal suitable for operating on the variable emphasis network includes limitation of the bandwidth to about 50 Hz, corresponding to a rise-time of 10 ms. Hence the audio is delayed by this time before entering the variable emphasis circuit.

Emphasized audio is passed to the step-size detection block which measures the slope of the signal to determine the optimum step-size (point C on the curve in Fig. 2). The logarithm of this value is coded into the step-size control data. Again conversion back to actual step-size voltage restricts the rise-time so the audio is again delayed before entering the actual delta-modulator. An equal delay is applied to the emphasis control data so that all data will arrive properly timed at the decoder.

11. AUDIO PERFORMANCE

The coding has been designed so that the dynamic range of the system easily may exceed 100 dB, though actual circuit implementations may not reach this figure. Frequency response may in theory extend to one-half of the audio clock, but it is wise to low pass filter incoming audio at 16 kHz to eliminate spurious signals. The input filters may be much gentler than those required with low sample rate PCM. Non-linear distortions are not inherent in DM but practical circuits yield levels on the order of 0.1 to 0.2%. With an adequate audio data rate the audible performance of the codec is superb. What one considers adequate can be the subject of some debate. Our feeling is that a very high level of performance is achieved with an audio clock at around 300 kHz. We typically recommend operation at about 200 kHz , as performance at this rate is still so good that it is only with the most critical listening conditions and program material that one can begin to detect noise modulation. For some applications a clock rate as low as 157 kHz may yield more than adequate performance.

12. ERROR CONCEALMENT

The effects of bit errors on PCM or DM systems are very different. An error in a PCM system produces an impulse lasting one sample period, with an amplitude dependent on the significance of the bit in error. MSB errors will produce impulses of half of full scale. An error in a DM system produces a small step which lasts for a long time (the step decays by the integrator leak time constant). The spectra of the errors is such that the PCM impulse is essentially always audible, whereas the DM step will often be audibly masked by the program material. Reproduced errors, while certainly somewhat objectionable, can be tolerated with DM systems. They cannot be tolerated with PCM systems because of their possibly huge amplitude.

Error concealment in PCM works by reducing the amplitude of the error impulse. To work, errors must be reliably detected, and a previous or interpolated word substituted for the word in error. As long as errors are reliably detected concealment can give fair results but at some error rate error detection will become unreliable, concealment will break down and many unconcealed errors will be reproduced. Because the reproduction of these errors is intolerable, the audio will have to be muted in this situation, although audio which erratically mutes is also intolerably annoying leading the listener to turn it off.

Error concealment in DM works by reducing the length of the small step which the error created. The approximate location and polarity of an error is determined and an opposite polarity error is artificially introduced near the original error; this terminates the step the error created. The step typically is shortened to a length similar to a PCM sample period. The performance of PCM or DM concealment successfully operating is similar since both produce small steps of similar duration. When high error rates break down the error concealment in a DM system some unconcealed errors are reproduced. Since these errors may be tolerated, there is no need to mute or turn off the audio. A DM system is typically usable at an error rate something like an order of magnitude (or more) higher than a PCM system.

13. APPLICATIONS

The system is being applied to the transmission of audio with video and audio only. Application of the system requires that the encoder output data be formatted into blocks and provision made for synchronization, so that the receiver can

determine the function of each received bit. Two general types of formats are used: those with bursty data and those with continuous data. An example of a bursty format is the transmission of audio data in the horizontal interval of an NTSC or PAL TV signal, or in the B-MAC format. Since the choice of data rate is very flexible (the quality of sound varies gradually as the bit rate is changed) it is common to choose a rate conveniently related to video frequencies. Systems have used one-half of the horizontal line rate for control data and a multiple of the line rate for audio data. The most common audio data rate in use is 13 fh or 204 k bits/s. Error concealment blocks are conveniently chosen to be groups of one or two lines worth of data.

Continuous data formats are used where the digital data is sent without interruption over a link. Sound only transmissions or those with video but using an unrelated carrier are examples. Common bit rates for stereo audio transmissions will include 512 k bits/sec (or 1/3 of T1 rate) and 772 k bits/sec (or 1/2 of T1 rate).

14. CONCLUSION

The digital audio coding method described here fills the need for a rugged, tolerant, efficient, practical system which can be implemented today. Numerous application areas are being developed including consumer delivery systems via DBS, cable TV and terrestrial broadcasting. Professional applications include network broadcast sound distribution and STL links. Integrated circuit decoders are under development. New techniques will significantly lower the cost of the encoding hardware which will allow more applications to become feasible.

REFERENCE

MASAMURA, T., SAMEJIMA, S., MORIHIRO, Y. and FUKETA, H. [June, 1979] Differential detection of MSK with nonredundant error correction. *IEEE Trans. Comm.*, Vol. COM-27, 6, 912-918.

BIBLIOGRAPHY

CLARK, D. S. [1985] Demodulation techniques for 20 Mbit/s 2-4 PSK signals. IBA E & D Report No. 133/85.

* For the same bit error ratio, these results could also be valid for the D2-MAC/packet system which uses the same coding law and the same error protection as the C-MAC/packet system. The relation between bit error ratio and C/N is shown in Fig. 8 of Annex IV for the C-MAC/packet system and in Fig. 4 of Annex III for the D2-MAC/packet system.

ANNEX V

SUBJECTIVE TEST RESULTS ON ADAPTIVE DELTA MODULATION FOR HIGH-QUALITY SOUND CODING

1. Subjective measurements conducted in Australia

The Australian Broadcasting Corporation (AuBC) has conducted subjective tests on the adaptive delta modulation (ADM) system [AuBC, 1985] in accordance with Recommendation 562 to relate subjective audio quality to bit error ratio, with the eventual objective of relating this to C/N ratio.

The supplied ADM equipment had a random error generator card to introduce errors into the data stream in a known switch-selectable ratio. The accuracy of the switch settings was checked and found to be accurate within normal statistical variations.

A carefully controlled and measured environment was used to replay the material in sequences according to Recommendation 562.

For evaluation of the audio part of the system, seven different segments of material were chosen from compact disc or original 16 bit PCM recordings and recorded in sequence in 16 bit PCM on a videotape machine for subsequent replay. The seven programme segments are listed in Table III.

TABLE III – *Reference programme segments*

Item (and curve) No.	Description
1	Male solo
2	Male spoken word
3	Piano
4	Flute, then boys' choir
5	Female spoken word
6	Orchestral
7	Mixed modern group

Figure 12 indicates that for a given BER, the impairments vary widely but consistently with the programme material, but that a mean impairment rating of 3.5 is achieved for a BER around 10^{-3}.

The overall impairment rating of 3.5 for a BER of 10^{-3} was thought to be equivalent to a C/N less than 9 dB. However, this would be degraded when using four levels of data as proposed for Australia. It was not possible to make measurements on this mode of operation, but results suggest a degradation of approximately 1 dB for the outer levels of data and approximately 1.8 dB for the inner levels. This may have implications as to which programmes should be carried on which data channels. However, facilities have not been available to the AuBC to corroborate these figures.

Independent correlation is yet to be established between C/N and BER obtainable from receivers in the field.