DVD: Pre-emphasis for use at 96kHz or 88.2kHz

Acoustic Renaissance for Audio

This Proposal is addressed to members of the DVD Consortium [1]*, to WG-4 of the DVD Consortium, to members of IFPI, RIAA, RIAJ and of the Audio Engineering Society.*

 jrs@meridian-audio.com

0. Summary

This document supplements other documents from the *ARA* concerning audio for both the video DVD (DVDV) and a pure-audio DVD (DVDA). [2]

This document proposes a new pre- and de-emphasis scheme for material recorded at 96kHz (or 88.2kHz) using linear PCM. The detail of this method was the original work of the late Michael Gerzon.

We suggest that this pre- and de-emphasis scheme be absorbed into any revision of the DVD standard [8], and be incorporated into the thinking behind any pure-audio DVD standard.

We show that the combination of noise-shaping and a new pre-/de-emphasis characteristic for 96kHz (88.2kHz) applications can result in an effective addition of between 2 and 7 bits to the channel – that is to suggest that for these sampling rates a 16-bit channel is sufficient[2].

Compared with other methods of coding, this scheme has the benefits of:

- Very low cost at replay (the de-emphasis can be a simple analogue filter)
- Assured transparency
- Low data rate
- Compatible with existing systems

We believe it should be considered urgently, since the benefits are very substantial.

1. Actions Required

1. We are asking for a change of meaning to the 'Emphasis' flag in the DVDV and DVDA systems to indicate a new pre-emphasis scheme being used for linear PCM when the sampling rate is 88.2kHz or 96kHz.

   The DVD Video Standard 1.0 currently permits the use of the old-fashioned pre-emphasis to the CD Standard at 48kHz (i.e. an analogue model using 50μS and 15μS time-constants). We suggest a new emphasis standard that can be used for 88.2kHz and 96kHz.
2. We suggest that the standard for DVD Audio be sufficiently flexible to include a 14-bit option at 96kHz or 88.2kHz.

---

[1] The DVD standard is published by, and copyright of: *Hitachi Ltd, Matsushita Electric Industrial Co. Ltd, Mitsubishi Electric Corporation, Philips Electronics N.V., Pioneer Electronic Corporation, Sony Corporation, Thompson Multimedia, Time Warner Inc., Toshiba Corporation, Victor Company of Japan Inc.*

[2] Actually a 14-bit channel will give a 21-bit dynamic range. The examples given are based on 16-bit channels since that is the smallest option in the DVD video standard.

2. Extent of discussion and authorship

This document is a proposal from members of *Acoustic Renaissance for Audio*, a free body dedicated to advancing audio quality. The signatories, their advisors and all affiliations are appended in section 9.

3. Introduction to Pre-/De-emphasis

The use of pre- and de-emphasis as a signal-processing method of optimising the subjective dynamic range of analogue channels is quite familiar to audio engineers. In particular, pre- and de-emphasis have been used in channels where the analogue noise level showed an increasing level with frequency, e.g. magnetic tape, shellac or vinyl grooves and FM broadcast. In all these cases, a well documented property of music and speech is exploited; namely that for material of acoustic origin, microphoned at normal listening positions, the average and peak spectrum level declines with frequency above a few kHz. Best use was therefore made of the channel by pre-emphasising high-frequency signals to an extent where they were more likely to occupy the capacity of the channel. De-emphasis is applied on replay or reception and has the dual benefit of reducing noise and distortion from the preceding chain.

Although the origins of pre- followed by de-emphasis were in analogue systems, the techniques can be very usefully applied to digital channels. For applications where a link in the transmission chain requires a smaller word size (e.g. moving a 20-bit recording to CD), it has been shown in [5] that very real benefit could be obtained by:

- performing pre-emphasis in the digital domain, *and*
- quantising with a noise-shaper designed to exploit the pre-emphasis curve, *and*
- performing de-emphasis in the analogue domain (or, otherwise to a large word-size digital channel).

So far, all linear PCM standards for digital audio permit the use of pre- and de-emphasis; the universal characteristic being 50/15µS.

This pre-emphasis characteristic allows an increase in subjective dynamic range by boosting audio frequencies above 3kHz in the transmission channel and attenuating them (and channel noise) on replay. This pre-emphasis characteristic has not been totally popular with the recording industry principally because it looses too much high-frequency headroom, and presents a mastering management issue – ensuring that its use is flagged.

More recently, recordings are being made with larger word-size than the distribution medium – e.g. it is now commonplace to record 20-bit on sessions.

4. Channel optimisation at 96kHz

Linear and psychoacoustically correct coding methods are known which can improve the performance of linear-PCM channels, particularly if the distribution channel has a smaller word-size than the original. The principal methods use noise-shaping at the word-length reduction process to maintain a highly effective dynamic range.

Noise shaping can also be very effectively combined with pre-/de-emphasis – and is particularly effective if the noise-shaper design takes account of the pre-emphasis curve. [1] [3] [5] [6] [7]

This document shows that the combination of noise-shaping and a new pre-/de-emphasis characteristic for 96kHz (88.2kHz) applications, can result in an effective addition of up to 7 bits to the channel capacity.

For linear PCM systems operating at high sample rates, this technology is important, it ensures e.g. that a 16-bit channel operating at 96kHz can provide an effective dynamic range equivalent to 23 bits in a normal 48kHz PCM channel.

5. Pre-/De-emphasis for linear PCM at 88.2kHz or 96kHz

We suggest it is more appropriate for the pre-/de-emphasis characteristic to be:

- described in the digital domain, and
- implemented by a noise-shaper.

The pre-emphasis and noise-shaper characteristics can be implemented with the structure shown in Fig. 6 and the responses are shown in Fig. 1, and described by:

$$P = \frac{k}{1 + \frac{34}{16} z^{-1} + \frac{23}{16} z^{-2} + \frac{2}{16} z^{-3} - \frac{2}{16} z^{-4}} \qquad (1)$$

k = 73/16 for the pre-emphasis curve and k = 1 for the shaper.

The advantages of this characteristic are as follows:

- Substantially improved high audio-frequency headroom compared to the current pre-emphasis. (The reduction of headroom is only 2dB at 15kHz compared to 9dB with the current standard.)
- The pre-emphasis method includes a noise-shaper which gives a 2.2-bit increase in overall audio dynamic-range when used as a word-length-reduction device.
- Because the noise-shaper has the same shape as the pre-emphasis curve, the output (i.e. de-emphasised) noise spectrum is ‘white’.
- The de-emphasis can be accurately applied by digital replay devices using a simple FIR filter. Such a filter has the benefit of not propagating transmission errors.
- The de-emphasis can be very closely matched by analogue replay devices.
- Can be usefully combined at the user’s discretion with appropriately matched high-advantage noise-shapers, like that shown in Fig. 4.
- The pre-emphasis can be implemented using *Autodither* or *Lossless* processing – in which case their use should be flagged,

Figures 2 and 3 show the overall channel dynamic-range expressed in ‘bits’. The headroom is reduced at ultrasonic frequencies – note the flat or ‘white’ spectrum of the final noise.

6. Optional application with additional Noise-Shaping

It is not proposed that any additional noise-shaper be standardised, since the replay device does not have to know the method or that it is being used.

The combination of pre-emphasis and noise-shaper can be implemented with the structure shown in Fig. 7.

Fig. 4 gives the response of a suitable shaper that could be used in combination with the pre-/de-emphasis described by Equation 1. This shaper has a characteristic given by:

$$N = \frac{1 - \frac{22}{16}z^{-1} + \frac{9}{16}z^{-2}}{1 + \frac{31}{16}z^{-1} + \frac{26}{16}z^{-2} + \frac{9}{16}z^{-3} - \frac{1}{16}z^{-4}} \quad (2)$$

Fig. 5 clarifies the way in which the suggested pre-emphasis and example noise-shaper provide increased dynamic range. The headroom curves at the top show the de-emphasised response normalised for 16, 20 and 24-bit channels. The lower curve is the noise-spectrum for the example shaper after correction to allow for the 2.19 bit gain given by the pre-emphasis scheme (see Fig. 2 and 3). This figure indicates how a 16-bit channel at 96kHz can give an effective dynamic range of 23 bits in the critical 4kHz region; note also that the channel is still offering 19-bit performance at 20kHz.

7. Representing 96kHz Pre-emphasis on DVD

7.1 Pre-emphasis indication in Linear PCM Audio

Changes to DVD Spec. 1.0:

1. Table 5.4.2.1-1:

- Change entry under ‘Emphasis’ and 96kHz
- Suggestion: Can be applied according to equation (1) above.

8. Rights

To the best of our knowledge this material is within the state of the art, and is not under any patent or licensing rights.

9. Authorship

Bob Stuart. Chairman and Technical Director, Meridian Audio Ltd. Visiting Fellow Essex University, Fellow *Audio Engineering Society*, Member *Acoustical Society of America*, Chairman *Acoustic Renaissance for Audio*, Member *XtraBits*, Member *Technical Subcommittee National Sound Archive*, Member *IEE* and *IEEE*. Member *ADA* committee of *Japan Audio Society*.

Peter Craven. Consultant. Member *Audio Engineering Society*, Member *XtraBits*.

Takeo Yamomoto. Chief Auditor, Pioneer Electronic Inc., Member *Audio Engineering Society*, Founder *Japan Acoustic Renaissance for Audio*, Member *ADA* committee of *Japan Audio Society*.

Professor Malcolm Hawksford. University of Essex. Fellow *Audio Engineering Society*, Fellow *Institute of Acoustics*, Fellow *IEE*., Member *Acoustic Renaissance for Audio*.

Tony Griffiths. Consultant. Fellow *Audio Engineering Society*, Member *Acoustic Renaissance for Audio*, Chairman *Technical Subcommittee National Sound Archive*, Member *IEE*, Member *Royal Television Society*.

Michael Gerzon. (Deceased). Gold Medallist and Fellow *Audio Engineering Society*, Former Member *XtraBits*, Former Member *Acoustic Renaissance for Audio*.

10. Bibliography and references

1 Akune, M., Heddle, R.M. and Akagiri, K. 'Super Bit Mapping: Psychoacoustically Optimized Digital Recording', *AES 93rd Convention* San Francisco, preprint 3371 (1992)

2 *Acoustic Renaissance for Audio,* 'A Proposal for High-Quality Application of High-Density CD Carriers', private publication (April 1995)

3 Craven, P.G. and Gerzon, M.A. 'Compatible Improvement of 16-Bit Systems Using Subtractive Dither' *AES 93rd Convention* San Francisco, preprint 3356 (1992)

4 Hawksford, M.O.J. 'Extended-Definition Digital Audio systems for High Capacity CD' *IEE Colloquium on Audio Engineering,* digest 1995/089, pp 2-1 to 2-12 (May 1995)

5 Gerzon, M.A., Craven, P.G., Stuart, J.R. and Wilson, R.J. 'Psychoacoustic Noise Shaped Improvements in CD and Other Linear Digital Media' *94th AES Convention*, Berlin preprint 3501 (March 1993)

6 Stuart, J.R. and Wilson, R.J. 'Dynamic Range Enhancement Using Noise-shaped Dither Applied to Signals with and without Pre-emphasis' *AES 96th Convention*, Amsterdam, preprint 3871 (1994)

7 Stuart, J.R. and Wilson, R.J. 'Dynamic Range Enhancement using Noise-Shaped Dither at 44.1, 48 and 96 kHz' *AES 100th Convention*, Copenhagen, (1996)

8 Toshiba et al., 'DVD Specifications for Read-Only Disc', Version 1.0 (September 1996)

11. Acknowledgements

DVD is a registered trademark of *Toshiba* and the ten companies.

12. Contact Addresses

Acoustic Renaissance for Audio, Meridian Audio, Stonehill, Stukeley Meadows, HUNTINGDON PE18 6ED, England. **Phone**: +44 1480 52144 **Fax**: +44 1480 451587

Email: ara@meridian-audio.com **Web:** http://www.meridian-audio.com/ara

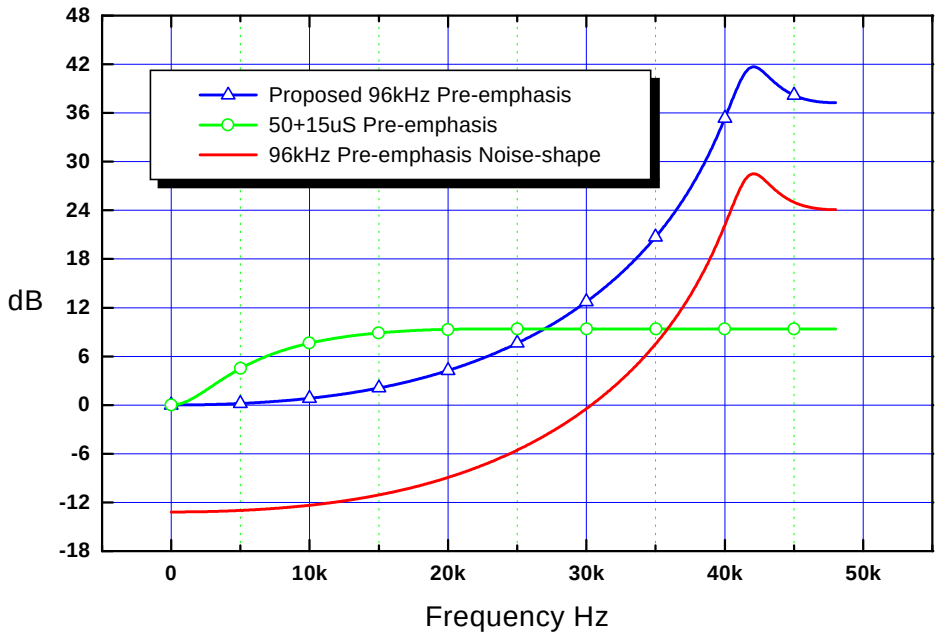

Figure 1 Showing the proposed pre-emphasis compared to 50µS / 15µS, and the noise-spectrum resulting from the pre-emphasis noise-shaper.

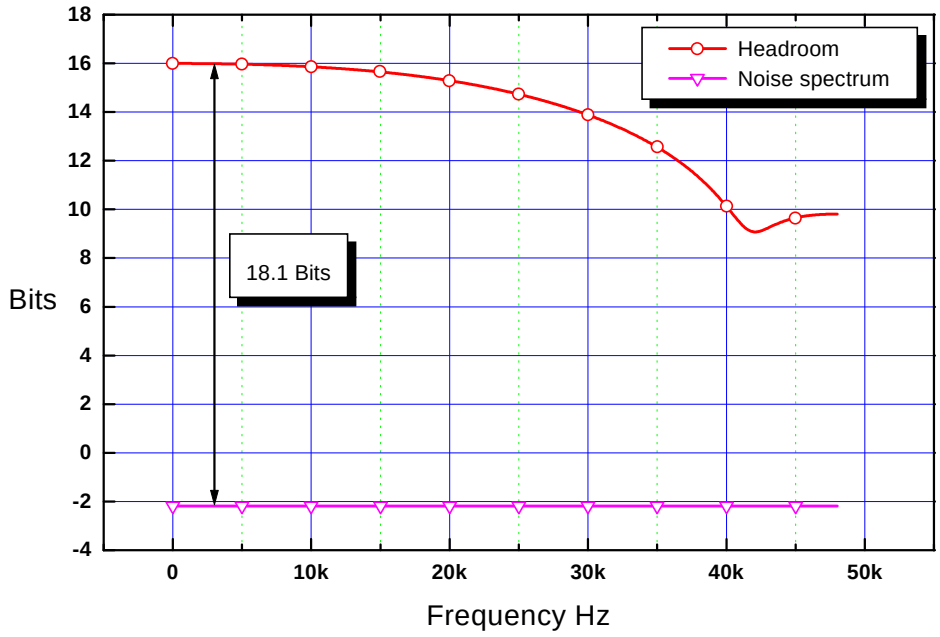

Figure 2 Showing the output noise spectrum and headroom for a channel after applying the proposed pre- and de-emphasis. The graph expresses dynamic range in bits. This example illustrates a capacity of 18.1 bits at 4kHz for a 16-bit channel – i.e. a perceptual gain of 2.1 bits.

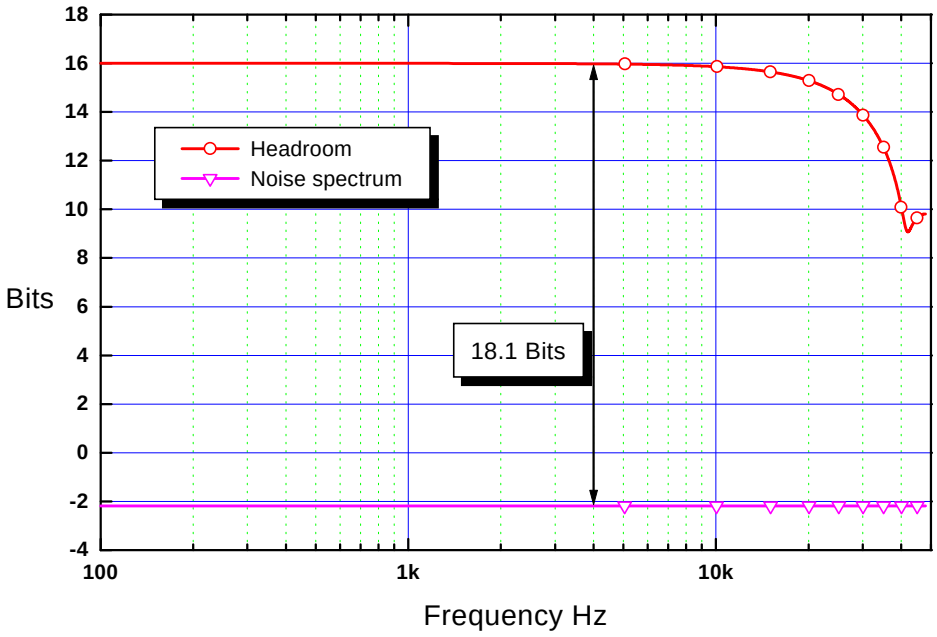

Figure 3 Showing Figure 2 on a log frequency scale.

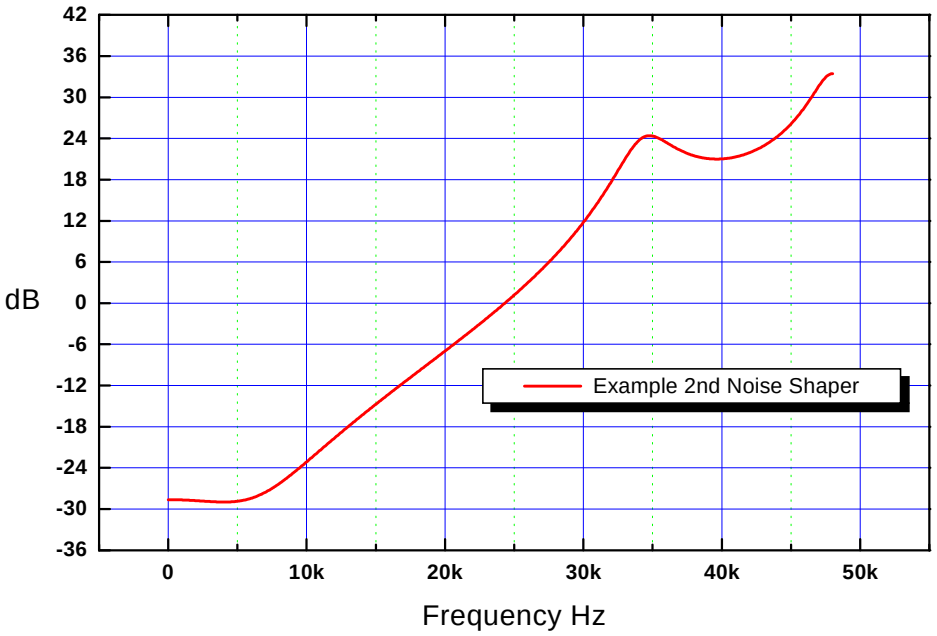

Figure 4 An example of a 6th-order noise-shaper that can be combined with the proposed pre-emphasis scheme.

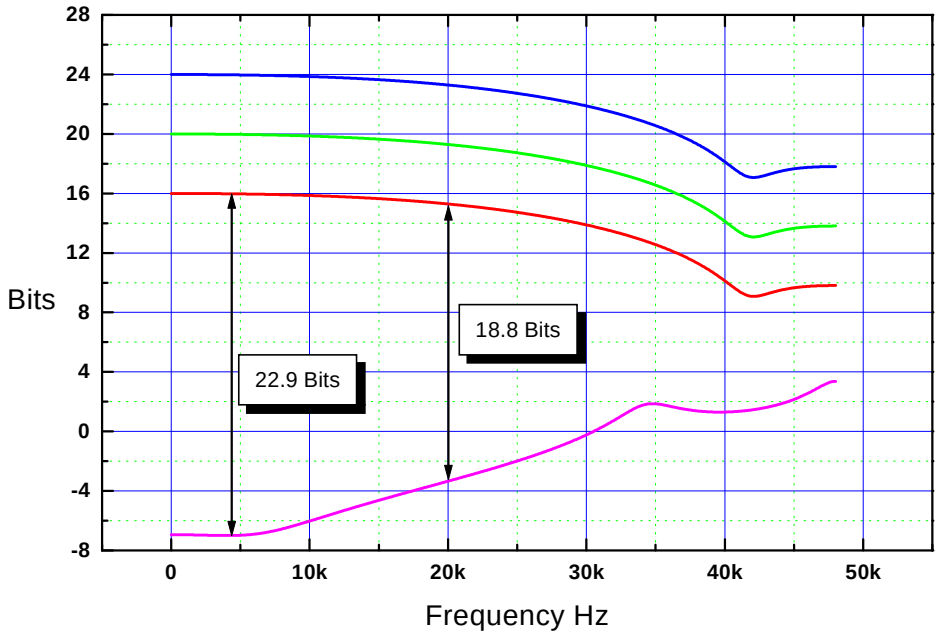

Figure 5 Showing the output noise spectrum and headroom for a channel after combining the example 2nd noise-shaper with the proposed pre- and de-emphasis. The graph expresses dynamic range in bits. This example illustrates a capacity of almost 23 bits at 4kHz for a 16-bit channel – i.e. a perceptual gain of 7 bits.

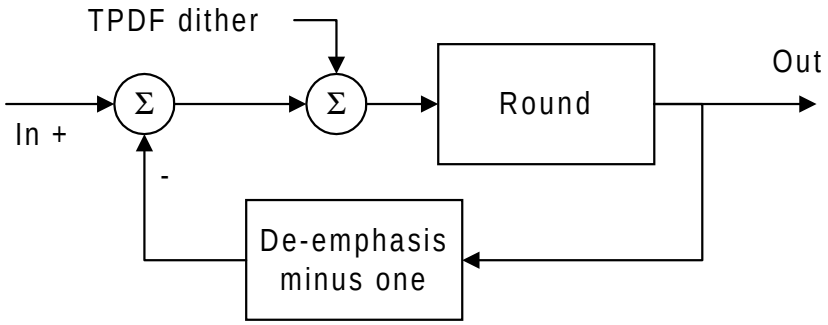

Figure 6 Computing structure for the proposed pre-emphasis scheme.

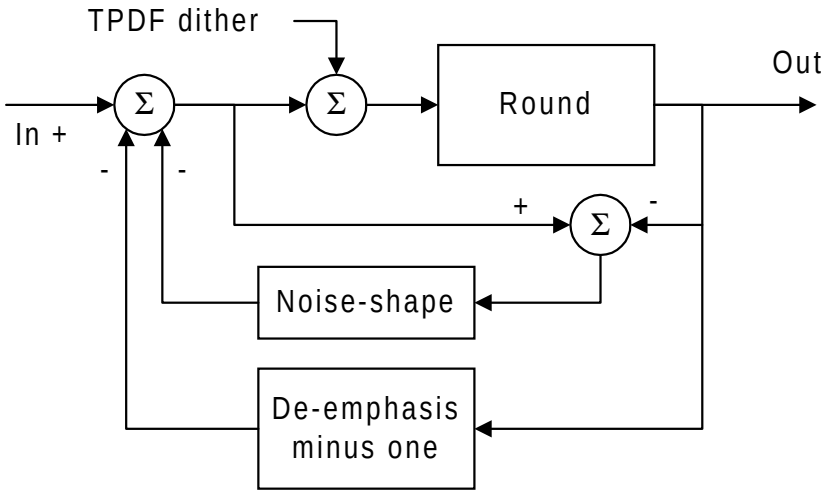

Figure 7 Computing structure for combining additional noise-shaping with the proposed pre-emphasis scheme.