Pre- and Postemphasis Techniques as Applied to Audio Recording Systems*

LOUIS D. FIELDER**

Ampex Corporation, Redwood City, CA 94063, USA

Audio recorders benefit from pre- and postemphasis, which reshapes the noise spectrum to match human audibility thresholds. A 10-dB increase in apparent dynamic range is realized for some digital audio systems. A first-order boost based on the CCITT J.17 preemphasis standard is shown to be appropriate for dynamic range expansion. A survey of peak acoustic levels present in 36 music performances is also included.

0 INTRODUCTION

The pre- and postemphasis technique is a method that modifies the spectrum of an audio signal from a music performance at the input of an audio channel with inherently flat overall response and then performs the inverse modification at its output to produce a system with a flat low-level frequency response and a modified background noise spectrum. This has been done to match more closely the system background noise to the characteristics of the human ear and the background acoustic noise spectrum of the listening or recording environment in order to produce wider apparent dynamic range. The emphasis technique has found wide application in the past because of the limited dynamic range of audio systems and the fact that music sources produce more energy in the low-frequency region where the ear is less sensitive to noise.

With the advent of digital audio recording and transmission techniques, which have a wider dynamic range, many have felt that pre- and postemphasis was no longer necessary. Quite the contrary is true, however, because the required dynamic range for noise-free reproduction of music is greater than the possible 98 dB of present-day 16-bit pulse-code modulation (PCM) systems. This author [1] has previously shown that a dynamic range capability of at least 118 dB was necessary for subjectively noise-free reproduction of music. Another paper by Manson [2] also confirms the fact that a dynamic range of at least 109 dB is necessary for noise-free reproduction. This paper will confirm this fact and demonstrate that the pre- and postemphasis technique is useful in achieving this wider range.

The use of pre- and postemphasis to develop a subjectively noise-free PCM device for music recording will be the primary concern of this paper. The PCM device is assumed to have an inherently flat frequency response input to output and a white noise floor low enough to be made noise free by pre- and postemphasis. In addition, it will be shown that the use of this technique on the standard 16-bit linear PCM professional recorders produces an improvement, making them subjectively noiseless in approximately twice as many recording situations. It will also be shown that the pre- and postemphasis technique is useful for lower dynamic range systems but not to as great a degree as in systems designed to produce a background noise at or below the threshold of audibility.

1 REVIEW OF PREVIOUS STUDIES

Previous studies of pre- and postemphasis have concerned themselves with the improvement of sound tracks for film, phonograph records, FM radio transmissions, and analog tape recorders. Surprisingly the studies most similar to the work presented here were two very early papers (1941) by Steinberg [3] and Fletcher [4] on the development of a subjectively noise-free optical film sound track. Many of the elements to be discussed in this paper were also presented in these studies. Steinberg used human hearing acuity and environmental noise factors compared to film track noise to determine an

* Manuscript received, 1983 May 16; revised 1985 May 22.

** Present address: Dolby Laboratories, Inc., San Francisco, CA 94114, USA.

optimum postemphasis and then investigated the effect on program peak level when preemphasis was applied. He came to the conclusion that a preemphasis that was a 6-dB-per-octave shelf between 500 Hz and 5 kHz was most appropriate and surmised that preemphasis was useful particularly if some peak overload was permissible.

Later authors wrote papers dealing with the circumstances resulting in lowered but still audible noise floors, and these generally resulted in preemphasis proposals with more extreme high-frequency boosts than those of Steinberg or this author. An example was Stewart [5], who justified the 381-mm/s NAB postemphasis standard by demonstrating that the resultant record high-frequency boost matched the drop in average high-frequency energy present in music. Another was McKnight [6], who produced a preemphasis standard based on information concerning the spectral sensitivity of the human ear. When this was done, an additional record boost resulting from the use of the NAB postemphasis standard in the 1–10-kHz region was suggested (Ampex mastering equalization). McKnight [7] also examined the peak spectral levels of tape recordings and concluded that the worst-case situation had equal peak levels as a function of frequency and allowed no improvement with preemphasis. As a result he stated that preemphasis was useful in only some circumstances unless some peak overload was permissible. In a later paper McKnight and Hille [8] examined the signal and noise capabilities of the phonograph record and matched the master tape recorder to it by modifying the emphasis used. This technique resulted in a system with the overload and noise characteristics of the phonograph record. Next McKnight and Kendall [9] derived a pre- and postemphasis standard matching the overload characteristics of the original 1954 NAB standard tape recorder using the original tape. The resulting record equalization had less high-frequency boost than had been proposed previously in [7].

Other researchers have followed similar arguments as McKnight and coworkers. Bauer [10] extended the work on the peak spectral content in music recordings with the use of wider bandwidth, faster responding filters than McKnight, and some information on the statistical nature of the peak levels. Another researcher, Stuart [11], generated the spectrum for just audible noise in a residential room and combined it with the peak level spectral information available in Sivian, Dunn, and White [12]. From this he concluded that an additional boost between 1 and 8 kHz was required, similar to that originally proposed by McKnight as the Ampex mastering equalization. Boyanova [13] also came to the same conclusion regarding additional preemphasis by comparing the high-frequency overload characteristics of a 381-mm/s analog recorder with his curve of the average spectral content of music.

In summary, past investigations have used spectral examinations of music, overload characteristics of the audio medium in question, and hearing acuity to create preemphasis proposals. Most studies dealt with situations resulting in the lowering of still audible noise, and the resulting preemphasis proposals have had additional boosts in the region above 5 kHz. Only the Steinberg paper arrived at a quantitative measure of improvement obtained when an emphasis system was tied back to hearing acuity.

2 IDEAL PRE- AND POSTEMPHASIS AT THRESHOLD

To determine the ideal emphasis it is necessary to determine the spectrum of just-audible noise which has maximum levels in each frequency band. This is done by employing the concept of "critical" bands as defined by Fletcher [14], [15] and later discussed by Zwicker [16]. Basically the critical-band concept indicates that the ear acts as a 24-channel real-time analyzer with varying sensitivity in each of the channels, varying bandwidth as a function of frequency, and a sophisticated interconnection between them. Below threshold each channel is independent and cannot share energy with its neighbors. One interesting result of this model is the prediction that at low sound levels, initial perception occurs when one or more of the critical-band thresholds are exceeded. Within each band this threshold energy may come from a combination of noise or other signals. The only essential requirement is that the total energy level within the band exceed a certain prescribed value. Also, at threshold levels it does not matter if one or any number of critical-band threshold levels are exceeded; the noise is still just audible, since only energy within each band is used for sound detection. This independence of critical-band channels at low levels had been observed and examined by Zwicker, Flottorp, and Stevens [17] and Scharf [18], who saw that loudness actually decreased when an equal-energy noise signal's bandwidth was increased beyond a critical bandwidth if the sound levels were low enough. The preceding properties of the critical-band hypothesis lead to the following hypothesis: a just-audible white-noise signal, which has been equalized so that noise energy is just sufficient to exceed the threshold in all critical bands, results in a background noise no more audible than the original signal.

To determine this ideal amount of boost, the critical-band concept was first combined with the spectrum of just audible sine wave levels as determined in the ISO-R226 standard [19], by Robinson and Dadson [20], and much earlier by Fletcher and Munson [21]. In addition the threshold spectrum of just audible octave noise by Robinson and Whittle [22] was used to provide an alternative source to derive the required acoustic noise spectrum level. Hopefully the two resulting curves of the predicted 1-Hz bandwidth spectrum level of noise that was just audible when the noise bandwidth equaled or exceeded a critical bandwidth would be similar, indicating the validity of this approach. In addition, to improve the accuracy of the resulting curves it was necessary to modify them to account for the difference in threshold spectrum between the listening environment and the anechoic chamber used for the threshold determinations. This modification of the detection characteristics of the ear from the anechoic frontal incidence to the diffuse field case was examined by Robinson, Whittle, and Bowsher [23] and applied to these threshold spectra.

Because the preceding boost requirement was derived from an extension of the critical-band concept, experimental verification was very desirable. To do this, 10 listeners aged 24–35, with good hearing, were exposed to a one-third-octave noise signal switched on and off at a 1-s rate. A one-third-octave noise signal was used because it exceeded critical bandwidth for all frequencies tested and was sufficiently narrow to provide good resolution. These experiments were done with a calibrated Spendor (LS3/5a) loudspeaker in a quiet home listening room which had room noise below the threshold of audibility above 500 Hz.

Fig. 1 displays the sine wave, octave noise, and experimental results for the one-third-octave noise spectrum for the threshold of hearing. There is good agreement between all three results. The largest deviations occurred in the 2–5-kHz region between the experimental and derived results from the previously published material. Although less than 4 dB at 3 kHz, this difference might be explained by the greater sensitivity afforded by switching the noise rather than holding it constant. Because of the relatively good agreement between all three curves, an accurate spectrum for the threshold of hearing for noise greater than a critical band in width is obtained by averaging the two curves derived from the literature. Therefore a postemphasis response which boosts the noise level at all frequencies except where the most sensitive critical band is, will create an increased noise background that is no more audible. This is accomplished by designating a filter response for postemphasis that has a minimum gain equal to unity at 4 kHz and greater everywhere else. The response polynomial for this response is given by Eq. (1), and the response is shown in Fig. 2.

$$\text{gain} = 1.86 \left| 1 + \mathrm{j}\frac{200\ \text{Hz}}{f} \right| \left| 1 + \mathrm{j}\frac{1.8\ \text{kHz}}{f} \right| \times \frac{\left| 1 + \mathrm{j}\dfrac{f}{(1.4)\ 4.5\ \text{kHz}} + \left\{ \mathrm{j}\dfrac{f}{4.5\ \text{kHz}} \right\}^2 \right|}{\left| 1 + \mathrm{j}\dfrac{f}{4.5\ \text{kHz}} \right| \left| 1 + \mathrm{j}\dfrac{f}{20\ \text{kHz}} \right|} . \quad (1)$$

In addition an alternate postemphasis curve was also derived. It was very similar to the preceding except that its response was flat above 4 kHz. Although this doen not match the hearing threshold, it will be shown that little reduction in the effectiveness of the resulting emphasis occurs, while at the same time the background noise that is produced is less noticeable for systems with resultant noise floors higher than the threshold of audibility. The response polynomial for this modified postemphasis is given by Eq. (2) and also shown in Fig. 2.

$$\text{gain} = 4.34 \left| 1 + \mathrm{j}\frac{430\ \text{Hz}}{f} \right|^2 \times \frac{\left| 1 + \mathrm{j}\dfrac{f}{(0.8)\ 4\ \text{kHz}} + \left\{ \mathrm{j}\dfrac{f}{4\ \text{kHz}} \right\}^2 \right|}{\left| 1 + \mathrm{j}\dfrac{f}{4\ \text{kHz}} \right| \left| 1 + \mathrm{j}\dfrac{f}{1.1\ \text{kHz}} \right|} . \quad (2)$$

The average noise threshold hearing curve and the resulting two postemphasized noise spectra are shown in this figure. The match to the just audible hearing spectrum is quite good. The basic characteristics of both preemphasis curves are that they have a 12-dB per octave rise from 20 to 500 Hz, a 6-dB per octave rise from 500 Hz to 4.5 kHz, and above 4.5 kHz they differ.

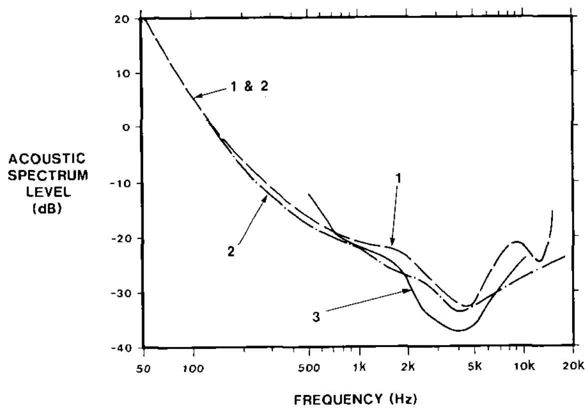


Fig. 1. Just audible noise spectra for noise signals exceeding a critical band in width. 1—derived from sine wave thresholds; 2—derived from octave noise thresholds; 3—average (10 subjects) one-third-octave noise thresholds measured by author. *Note:* 0 dB = 20 μPa/$\sqrt{\text{Hz}}$.

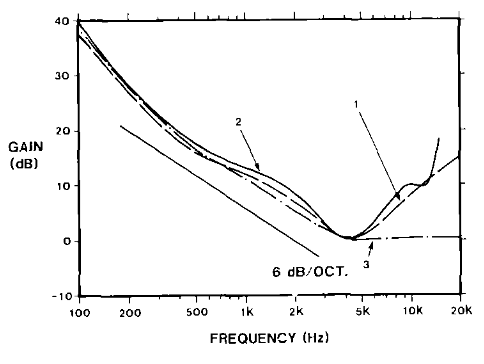


Fig. 2. Proposed deemphasis characteristics versus deduced threshold spectrum. 1—Proposed ideal postemphasis [Eq. (1)]; 2—1-Hz bandwidth noise threshold spectrum (average from Fig. 1); 3—modified postemphasis [Eq. (2)].

To test the validity of the resulting postemphasis schemes, the tests with the previous 10 subjects were continued. In this experiment the subjects were exposed to white noise and two types of postemphasized noise [as per Eqs. (1) and (2)]. The noise signals were adjusted in level until each was just audible. Fig. 3 shows the result of these tests, along with the earlier one-third-octave noise threshold experiment results.

This figure shows that postemphasizing the white noise by a response defined by Eqs. (1) and (2) does not increase its audibility. This is true because the unboosted frequencies near the maximum of the hearing sensitivity at 4.5 kHz are unchanged for all situations. Comparing the just audible broad-band noises with the one-third-octave experimental results determined earlier also shows good agreement, with the difference in hearing sensitivity between the one-third-octave and broad-band results at 3–5 kHz being an artifact of the averaging of the 10 subjects' one-third-octave thresholds into one cumulative result.

In summary, a postemphasis boost defined by both Eqs. (1) and (2) produced a noise signal no more audible than the original white noise. Both postemphases are approximately 6 dB per octave from 500 Hz to 4.5 kHz, with the differences between them occurring for frequencies above 4.5 kHz.

3 APPARENT DYNAMIC RANGE INCREASE—RECORDED MATERIAL

The evaluation of the benefit afforded by the proposed emphasis system was done by examining the peak level reduction afforded by the preemphasis complementary to Eqs. (1) and (2). This was true because the boosting postemphasis was previously shown to produce no audible increase in noise at threshold. A level reduction will always occur because the resultant preemphasis derived from Eqs. (1) and (2) has a maximum gain of 1 at 4.5 kHz and produces an attenuation at all other frequencies. Peak level reduction was examined by comparing the greatest positive or negative levels occurring longer than 30 μs for the nonpreemphasized and preemphasized signals from an entire music performance.

Because of the ease and simplicity afforded by the use of recorded material, the peak level content was determined in 20 popular studio and 20 classical coincident microphone recordings. This was done to examine the two previously defined emphasis system curves with the addition of two other possible simplified emphasis schemes. These simplified emphases were two first-order 6-dB-per-octave boosts shelving either at 5 or at 10 kHz. Both of these simpler preemphasis characteristics differed from the other two in that they had gains greater than 1 above 5 kHz. This was done to line up the response below 5 kHz so that the resulting postemphasized noise would have equal threshold of audibility as those defined by Eqs. (1) and (2). The preemphasis shelf extending to 10 kHz was included to show the disadvantage of extending the boost beyond 5 kHz. Fig. 4 demonstrates the gain characteristics of these preemphasis systems tested.

The 40 tape recording samples were then used to evaluate the peak level reduction afforded by each preemphasis scheme. The results are shown in Table 1. Table 1 presents the average and worst-case level reduction of either channel of the 40 recordings sampled The first observation is that two-microphone audience position classical recording results differed from the popular music studio-derived tapes. This separation reflects the difference in the nature of the two recording techniques. Because the spectral balance in studio recordings had been highly adjusted and tailored for reproduction in the home and automobile, there is more high-frequency energy than in naturally recorded performances.

Examination of Table 1 shows that the level reduction by the two preemphases that most closely match the hearing acuity spectrum were similar and produced the greatest level reduction of the group. Since the preemphasis derived from Eq. (1) will later be shown to produce poor results in situations resulting in an audible noise floor, it will be discarded in favor of the preemphasis complementary to Eq. (2). Considering the other two preemphasis schemes, the first-order preemphasis boosting to 5 kHz produced slightly less level reduction than the previous two, while the first-order 10-kHz shelving preemphasis provided a markedly less effective level reduction. Considering the preemphases complementary to Eqs. (1) and (2), even the worst case recordings had peak levels reduced somewhat, indicating that the design of a below threshold noise-producing audio channel was always assisted by the proper emphasis system. This was true because the preemphasis gain was never more than unity. The benefit afforded by the first-order preemphasis boosting to 5 kHz was only slightly less, being 10 and 5 dB on the average for classical and studio music, respectively, while the preemphasis boosting to 10 kHz produced a significant degradation in some cases.

In summary, the peak level reduction study using recorded material showed that best performance was obtained with the preemphasis defined by the inverse

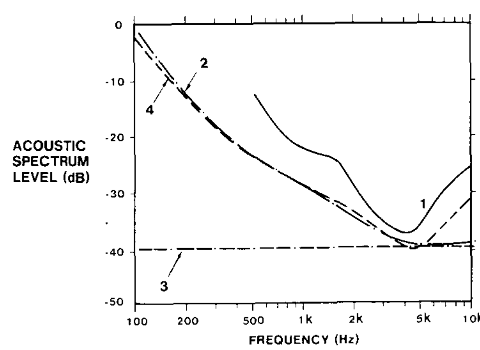


Fig. 3. One-third-octave derived hearing threshold compared to postemphasized and unemphasized white noise threshold signals. 1—one-third-octave derived threshold spectrum (average 10 subjects); 2—just audible post-emphasized noise spectrum [Eq. (2)]; 3—just audible white noise spectrum; 4—just audible postemphasized noise spectrum [Eq. (1)]. *Note:* 0 dB = 20 μPa/$\sqrt{\text{Hz}}$.

Table 1. Effectiveness of various preemphasis schemes to produce a level reduction

	Average level reduction (dB)		Worst-case level reduction (dB)	
Type of preemphasis employed	Classical, 2 microphones over conductor	Popular studio tapes	Classical, 2 microphones over conductor	Popular studio tapes
Preemphasis complementary to threshold noise spectrum to 4.5 kHz; gain = 1 above 4.5 kHz (suggested preemphasis).	11	5	1	1
Preemphasis completely complementary to threshold noise spectrum; maximum gain = 1 at 4.5 kHz.	11	5	2	2
Preemphasis 6-dB-per-octave boost with 10-kHz shelf; gain = 1 at 4.5 kHz.	8	0	−4	−3
Preemphasis 6-dB-per-octave boost with 5 kHz shelf; gain = 1 at 4.5 kHz.	10	5	1	1

of Eqs. (1) and (2), with the first one discarded for its other disadvantages. As a result, the preemphasis derived from Eq. 2 was adopted as the standard "ideal" preemphasis. However, the simplification afforded by the 5-kHz shelving preemphasis made it also desirable in view of only the small reduction in effectiveness and its simple implementation.

4 APPARENT DYNAMIC RANGE INCREASE—SOUND LEVEL SURVEY

Since the validity of the emphasis system was based on the fact that the background noise was just at audibility, it was necessary to have absolute peak acoustic level calibration when investigating the peak level reduction produced by the preemphases discussed previously. To that end a portable acoustic peak level meter using peak detectors similar to those mentioned before was built, incorporating a flat and preemphasized channel as per Eq. (2). It also included a high-quality miniature condenser microphone (Countryman #EM-102) which had a flat frequency response and an overload point greater than 140 dB. This sound level meter was then taken to 25 locations during 36 separate performances to examine the peak levels of 47 different music selections. The survey took place over a one-year period and concentrated on performances in the San Francisco Bay area. An attempt was made to sample a wide range of musical performances, with information obtained from classical, rock, jazz, country, big band, reggae, and folk music. Approximately one-half the samples involved performances employing electronic augmentation; in fact in some musical areas it was impossible to find a nonaugmented example. The sound level pickup point for the survey was always in the audience location and attempted to be at the "best" listening location. The results of this survey are found in Figs. 5–7. It should also be noted that close-miking musical instruments results in even higher levels, perhaps 10–15 dB higher.

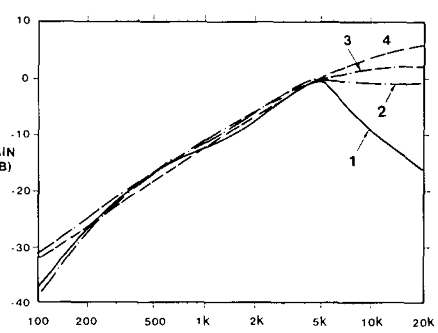


Fig. 4. Various test preemphasis responses. 1—preemphasis complementary to Eq. (1);2—preemphasis complementary to Eq. (2);3—test preemphasis, 6 dB per octave boosts to 5 kHz; 4—test preemphasis, 6 dB per octave boost to 10 kHz.

Fig. 5 is a histogram of the peak acoustic levels in the performances surveyed; it shows several interesting results. The first was the high peak acoustic pressures found in musical performances; typical levels varied from 97 to 127 dB. Even though the group between

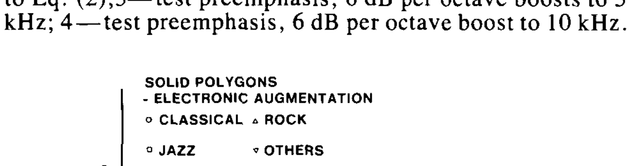


Fig. 5. Peak acoustic levels of various music performances.

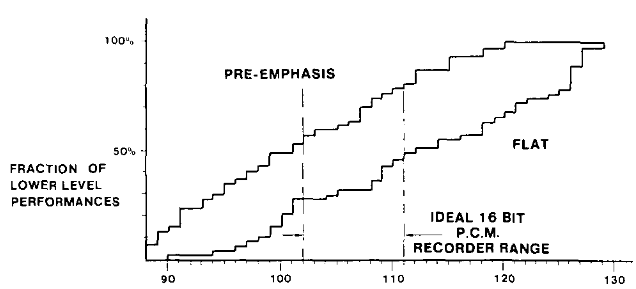


Fig. 6. Cumulative distribution of equivalent acoustic levels before and after preemphasis.

120 and 130 dB was primarily electronically augmented, the high peak pressures recorded probably came from use of the drum set, not the electronic augmentation. More examples of natural acoustic performances would exist between 120 and 130 dB if it were customary to play without electronic augmentation. The high sound levels encountered in live performances have great significance in the design of recording and transmission systems as well as in other parts of the audio chain because it is not widely recognized how high the peaks are in actual music performances.

Unfortunately very little work exists determining the peak levels in music performances. This author [1] earlier measured levels of 122 and 124 dB for percussive classical music and country music, respectively, while Sivian, Dunn, and White [12] provided peak sound level data for classical music with a piano solo, a 15-piece orchestra, a 75-piece orchestra, and a pipe organ performing and reported levels of 103, 112, 113, and 116 dB, respectively. Manson [2] of the BBC in a paper on PCM coding resolution also gave data for peak sound levels in recording studios at 111, 113, and 129 dB for recital, orchestra, and dance band music, respectively. Another work, by Lebo and Oliphant [24], focusing on hearing damage, presented peak data of 114 (popular), 122 (popular), and 101 dB (classical), using a C frequency weighting (reasonably close to a flat system) and a statistical level analyzer they developed. The 101-dB value for the classical music is lower than the other reported values, possibly due to the selection of a quieter performance or the incorporation of a slower responding peak meter.

The high sound levels imply wide dynamic range requirements for an audio recording or transmission system. Fortunately the application of pre- and post-emphasis reduces this requirement significantly. When the previously chosen preemphasis as per Eq. (2) is applied, a significant reduction in peak levels occurs. This effect is shown in Fig. 6, which is a cumulative distribution versus peak sound level for the flat and preemphasized acoustical signals for the music performances surveyed.

Examining Fig. 6 it should be noted that the two vertical lines marked indicate the spread in maximum possible recorded sound levels for noise-free operation when system background noise detection is either limited by hearing acuity (line at 102 dB) or by audience background noise (line at 111 dB), as shown by the author [1] in an earlier paper on dynamic range requirements for noise-free reproduction. Several interesting facts arise from examination of Fig. 6. The first is that approximately a 10-dB reduction in recorded level and therefore apparent dynamic range increase occurs when preemphasis is employed. The worst-case requirement of 129 dB SPL is reduced to 120 dB. Also shown is the "noise-free" maximum sound level recording capability which was provided by the use of an ideal 16-bit linear PCM recorder. Examination of the cumulative distribution in Fig. 6 indicates that an unemphasized 16-bit PCM recorder would be sufficiently quiet in only 28–50% of the situations surveyed, depending on whether background noise detection was limited by hearing acuity or audience background noise. After the application of emphasis these percentages increase to 53–83%. Although preemphasis is shown to be quite useful, the design of an audibly noise-free recorder requires a wider dynamic range than possible with a 16-bit PCM system. In fact, a 107–116-dB dynamic range is still required, depending on the background noise in recording circumstances.

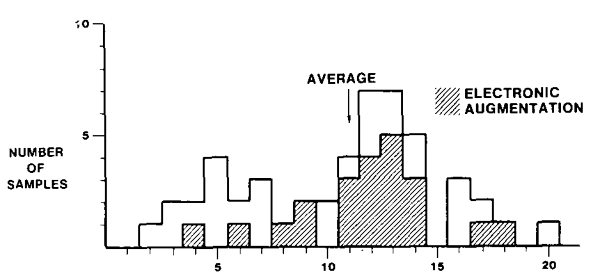


Fig. 7. Level reduction with preemphasis for the performances in sound survey.

At this point it is beneficial to compare these results with the previous studies on emphasis techniques. The work of McKnight [7] and Bauer [10] would indicate that no significant worst-case level reduction would occur when employing preemphasis of this kind because of the relative flatness at their peak spectral curves. The data obtained from this sound survey lead to a very different conclusion because the worst-case sound level of 129 dB is reduced to 9 dB. This difference is primarily a result of the lack of absolute acoustic reference in the previous spectral examinations. In this survey it was found that the loudest performances contained significantly less high-frequency energy than lower level ones. To give support to this conclusion, the sound survey data were examined without reference to absolute level and are presented in Fig. 7. Examination of Fig. 7 shows that the worst-case situation now only indicates a 2-dB improvement when preemphasis is employed. In fact, however, this selection was a relatively low level music performance where little preemphasis was required for noise-free reproduction of music.

In summary, we have shown that an audio system with a white-noise floor and flat power bandwidth can have its apparent dynamic range improved by between 5 and 11 dB for average recordings if the resulting background noise is just at audibility. The greatest improvement is available for recordings accurately representing the audience sound field conditions, while the lesser improvement manifested itself for processed studio recordings. For natural recording of all the performances surveyed, an original system dynamic range of 107–116 dB was required, depending whether microphone or audience noise were limiting factors.

5 EXTENSION TO NONTHRESHOLD SYSTEMS

The wide dynamic range requirement of 107–116 dB is presently beyond the reach of systems in use today. Because of this fact, an attempt to modify the previous results for nonthreshold-noise-producing systems was made. The fact that the noise would now be audible invalidated the original assumption of the independence of the critical bands. Therefore the emphasized background noise seemed subjectively louder than similarly raised white noise for levels above threshold. This was true since audible postemphasized noise excited many more critical bands than the original white noise, therefore it seemed louder. Thus an optimum preemphasis shape would depend on the total subjective audibility of the postemphasized noise as well as the peak level reduction afforded by the complementary preemphasis. Rather than explore this optimization for each recording condition and dynamic range of the system, a simpler approach was used, which employed the previously derived emphasis for systems with inaudible resultant noise floors and evaluated the reduction in emphasis system effectiveness due to the extra perceived loudness of the postemphasized noise relative to the original white noise. This correction factor was obtained experimentally by exposing the original subjects, as mentioned before, to white and postemphasized noise as defined by Eqs. (1) and (2). As the noise levels were increased above threshold in 5-dB steps, the postemphasized noise became subjectively louder relative to the raised white noise. Fig. 8 shows the results of this experiment.

Examination of Fig. 8 indicates that as two postemphasized noise signals increase along with the unemphasized white noise, the postemphasized noise, as per Eq. (2), eventually appeared 5 dB greater in subjective level. For the 16-bit PCM recorder employing emphasis and reproducing the loudest performance surveyed, the resulting noise is up to 20 dB above threshold, and therefore emphasis based on Eq. (2) would be 3–4 dB less effective. Also included in Fig. 8 are the results from a similar experiment using postemphasized noise as specified in Eq. (1). Only four subjects were used, and the experiment was discontinued because of the much greater penalty induced by boosting the frequencies above 4.5 kHz. The increased subjective loudness was up to 10 dB greater rather than 5 dB, and the resulting noise was perceived as much more unpleasant than the deemphasized noise with no boost above 4.5 kHz. As a result, an emphasis system fully

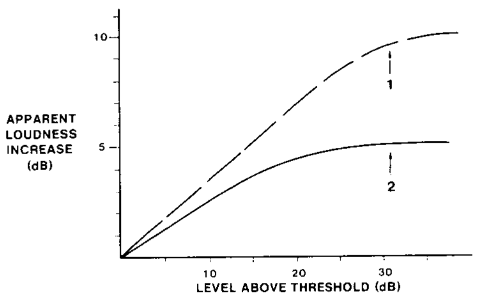


Fig. 8. Apparent relative loudness increase in the deemphasized noise as level is raised above threshold. 1—deemphasized noise [Eq. (1)]; 2—deemphasized noise [Eq. (2)].

complementary to hearing acuity is very undesirable for systems producing noise above threshold, and the modified emphasis defined by Eq. (2) that levels off above 4.5 kHz is much more desirable.

Therefore audio channels with flat power bandwidth and white noise floor, which will remain audible after emphasis is employed, are still benefited but to a lesser degree than in threshold systems. Considering an emphasis system based on Eq. (2), the effectiveness of emphasis as a dynamic range improver is reduced due to the greater subjective loudness of the postemphasized noise. For 16-bit PCM recorders in applications with resultant audible noise floors and employing emphasis as per Eq. (2), the dynamic range improvement will be as much as 4 dB less, thus yielding average improvements of 1–7 dB, rather than 5–11 dB, as was shown in Table 1. Since emphasis produced little average benefit for the recording of studio-derived material played back at high sound levels, it would be desirable to switch off the emphasis in these situations. In other words, unlike audio channels with a background noise just at or below audibility, emphasis in some situations can cause degradation in the perceived dynamic range if the music is played loud enough and the resulting background noise is too high. Situations requiring the reproduction of studio music from an emphasized 16-bit PCM recorder played at levels greater than 110 dB should be avoided.

6 SIMPLIFICATION OF EMPHASIS SYSTEMS

As previously shown, a preemphasis system based on a 6-dB-per-octave rise to 5 kHz was very nearly as effective as one defined by Eq. (2) and would most likely find practical applications. One further simplification required is the reduction of attenuation at low frequencies for the preemphasis. Since the improvement afforded by emphasis is approximately 10 dB, reducing the input signals significantly more than 10 dB at low frequencies relative to 5 kHz would provide little additional dynamic range expansion at the expense of greater implementation difficulty. A practical compromise can be accomplished by the use of two possible simplified preemphasis responses based on an existing preemphasis standard proposed by the CCITT. The European standard organization CCITT, in standards proposal J.17 [25] for emphasis of group sound links, defines a 6-dB-per-octave shelf filter between 477 and 4134 Hz, which is a good match to the simple preemphasis rising to 5 kHz. As a result it is proposed as a practical, already existing standard to implement and is shown in Fig. 9 along with the preemphasis complementary to Eq. (2). It should be noted, however, that any preemphasis response with additional low-frequency attenuation is equally or even slightly more effective in extending the dynamic range for systems designed to be "hiss-free." One such modified preemphasis is also shown in Fig. 9. It was proposed to serve the dual function of dynamic range extension and the lowering of modulation noise audibility in a companded digital audio system as introduced by the author [26].

Observation of this figure shows that the major differences exist below 1 kHz, where the exact preemphasis shape will have little effect on dynamic range extension. The curve defined by Eq. (2) has slightly more attenuation between 1 and 3 kHz and slightly less boost above 5 kHz because of its higher order. These differences manifest themselves as a slightly higher degree of effectiveness for this preemphasis. The absolute gain was adjusted so that the boost at 5 kHz equaled the average gain reduction observed in level reduction experiments in Sec. 3 for naturally recorded music. Therefore the conversion to pre- and postemphasis would not, on the average, necessitate an input gain setting change.

Two simplified preemphasis systems have thus been proposed. They derive from hearing sensitivity characteristics at threshold, are based on an already existing standard, and have their gains normalized so that recordings made using "natural" recording techniques require no gain adjustment when the emphasis is switched on. An additional benefit afforded by this gain normalization is that the pre- and postemphasis defined by CCITTJ.17 has unity gain at 1 kHz.

7 EXTENSION TO ANALOG TAPE RECORDERS

Finally the use of the preceding emphasis techniques for analog tape recorders was considered. This discussion focused on the 381-mm/s (15-in/s) tape speed with NAB equalization, using two tracks on a 6.3-mm (¼-in) tape width, with Ampex 456 tape on an Ampex ATR-102 tape recorder. We measured the rms noise spectrum and the peak playback level experiencing 1 dB compression (approximately 3% distortion) of this system. Then the recorder was modified to produce a flat frequency overload spectrum, and the CCITT J.17 pre- and postemphasis was applied. The resulting curves are shown in Fig. 10.

Examination of this figure showed that the noise levels in the 4-kHz region were reduced and therefore would produce a 9-dB apparent increase in the dynamic range at the threshold of audibility rather than the 10 dB previously seen for white noise floor systems. This was due to the fact that the analog recorder, equalized for flat overload characteristics, did not have a white noise floor. Since analog recorder applications almost always result in recordings that produce noise substantially above threshold, the emphasis effectiveness was 5 dB less effective. This was true because the apparent loudness increase of the postemphasized noise was at the asymptotic limit of 5 dB for apparent loudness increase of postemphasized noise. Therefore it was concluded that the analog recorder subjective dynamic range could be increased only 4 dB for naturally recorded performances. Studio applications using emphasis would not be useful due to the existence of a greater high-frequency content, and therefore the emphasis techniques proposed here would not be beneficial for analog recorder applications in the studio. These results apply for a comparison between a NAB standard machine and one modified by this emphasis proposal, designed for threshold-noise-producing systems having a white noise floor.

8 CONCLUSION

In conclusion, the technique of pre- and postemphasis was examined to determine its usefulness for audio recording and transmission systems. It was found that without pre- and postemphasis a dynamic range requirement of up to 125 dB existed. This was true because of the high peak acoustic levels found in the music performance survey combined with the knowledge that white noise signals as low as 4 dB SPL were audible to the listener. Fortunately pre- and postemphasis was useful in reducing this figure to 116 dB. Even so, 16-bit PCM recording and transmission systems were found to be inadequate for "noise-free" operation in the worst practical cases. Despite this limitation, emphasis was shown to produce an average apparent dynamic range of 7–9 dB in 16-bit PCM systems for recording applications accurately sampling the audience sound field depending how audible the resultant background noise

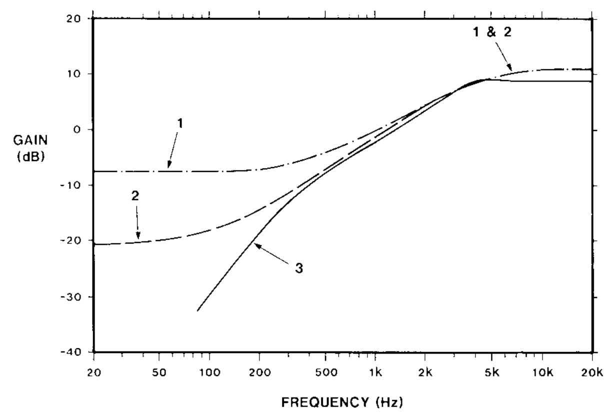


Fig. 9. Proposed preemphasis. 1—CCITT J.17 preemphasis recommendation, gain adjusted so that 1 kHz = 0 dB; 2—preemphasis extension to CCITT J.17 with more attenuation at low frequencies; 3—more ideal preemphasis, complementary to Eq. (2).

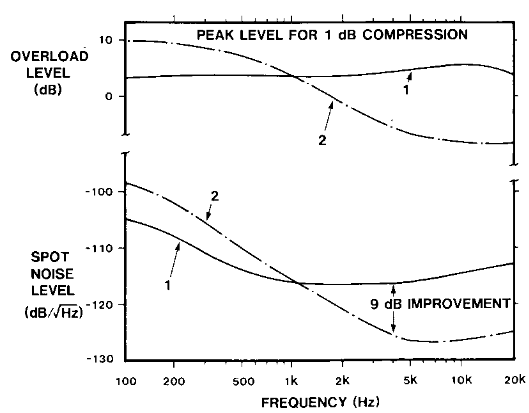


Fig. 10. Comparison of signal and noise characteristics at 381 mm/s (15 in/s) for NAB and proposed equalization. 1—adjusted to NAB standard; 2—after implementation of proposed pre- and postemphasis (ATR-100 as above). *Note:* 0 dB = original rms level producing 3% distortion at 400 Hz.

was. Examining the use of pre- and postemphasis for studio-type sound sources revealed the existence of substantially more high-frequency energy. This reduced the effectiveness of preemphasis by 5 dB in all applications when compared to naturally recorded performances. Pre- and postemphasis as defined previously and applied to analog recorders produced a smaller apparent dynamic range improvement of 4 dB for naturally recorded performances, while studio recording applications were not benefited.

The pre- and postemphasis curve shapes were arrived at by considering the development of a system with a just-audible background floor. The curve shapes were derived from the hearing characteristics of the human ear and were then adjusted and simplified for practical use. When this was done, the most practical emphasis standard was selected. It was the CCITT J.17 preemphasis standard, which boosted the noise floor between 477 Hz and 4134 kHz at a 6-dB-per-octave rate using a first-order filter. Interestingly enough, this 6-dB-per-octave boost closely followed the preemphasis proposed by Steinberg in 1941 [3], derived from similar arguments as those of the author. Unfortunately much of the later work extended the boost of the preemphasis above 5 kHz, which is not justifiable in view of the hearing characteristics of the human ear and the level reduction afforded by the preemphasis.

In summary, this study showed pre- and postemphasis to be useful in extending the dynamic range capabilities of audio systems. For the development of a subjectively noise-free recording system, improvements on the average of 10 dB were obtained for sound recordings sampling the sound field at the audience position, while studio-processed recordings fared less well. From the sound survey, the achievement of accurate music reproduction was shown to be made more difficult by the high peak sound levels of up to 130 dB encountered if no peak overload in the audio range was allowed.

9 REFERENCES

[1] L. D. Fielder, "Dynamic-Range Requirement for Subjectively Noise-Free Reproduction of Music," *J. Audio Eng. Soc.*, vol. 30, pp. 504–511 (1982 July/Aug.).

[2] W. I. Manson, "Digital Sound: Studio Signal Coding Resolution for Broadcasting," BBC Research Rep. 1980/15 (1980 Dec.).

[3] J. C. Steinberg, "The Stereophonic Sound Film System—Pre- and Post-Equalization of Compander Systems," *J. SMPTE*, vol. 37, pp. 366–379 (1941 Oct.).

[4] J. Fletcher, "The Stereophonic Sound Film System—General Theory," *J. SMPTE*, vol. 37, pp. 1–22 (1941 Oct.).

[5] W. E. Stewart, "Why the NARTB Curve for Magnetic Tape?" *Radio Telev. News*, vol. 53, pp. 40–41, 112 (1955 June).

[6] J. G. McKnight, "Signal-to-Noise Problems and a New Equalization for Magnetic Recording of Music," *J. Audio Eng. Soc.*, vol. 7, pp. 5–12 (1959 Jan.).

[7] J. G. McKnight, "The Distribution of Peak Energy in Recorded Music, and Its Relation to Magnetic Recording Systems," *J. Audio Eng. Soc.*, vol. 7, pp. 65–72 (1959 Apr.).

[8] J. G. McKnight and P. F. Hille, "Master-Tape Equalization Revised," presented at the 42nd Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 20, p. 418 (1972 June), preprint 856.

[9] J. G. McKnight and T. Kendall, "Proposed Equalization for 15 in/s Studio Master Recording on High-Output Low-Noise Tapes," presented at the 45th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 21, p. 482 (1973 July/Aug.), preprint 920.

[10] B. B. Bauer, "Octave-Band Spectral Distribution of Recorded Music," *J. Audio Eng. Soc.*, vol. 18, pp. 165–172 (1970 Apr.).

[11] J. R. Stuart, "Tape Noise Reduction," *Wireless World*, pp. 104–110 (1972 Mar.).

[12] L. J. Sivian, H. K. Dunn, and S. D. White, "Absolute Amplitudes and Spectra of Certain Musical Instruments and Orchestras," *J. Acoust. Soc. Am.*, vol. 1, pp. 330–371 (1931 Jan.).

[13] M. Boyanova, "Equalization in Magnetic Recording," *dB*, vol. 5, pp. 36–41 (1974 Sept.).

[14] H. Fletcher, "Loudness, Masking, and Their Relation to the Hearing Process and the Problem of Noise Measurement," *J. Acoust. Soc. Am.*, vol. 9, pp. 275–293 (1938 Apr.).

[15] H. Fletcher, "Auditory Patterns," *Revs. Mod. Phys.*, vol. 12, pp. 47–55 (1940 Jan.).

[16] E. Zwicker, "Subdivision of the Audible Frequency Range into Critical Bands," *J. Acoust. Soc. Am.*, vol. 33, p. 248 (1961 Feb.).

[17] E. Zwicker, G. Flottorp, and S. S. Stevens, "Critical Bandwidth in Loudness Summation," *J. Acoust. Soc. Am.*, vol. 29, pp. 548–557 (1957 May).

[18] B. Scharf, "Critical Bands and the Loudness of Complex Sounds near Thresholds," *J. Acoust. Soc. Am.*, vol. 31, pp. 365–370 (1959 Mar.).

[19] "Normal Equal-Loudness Contours for Pure Tones and Normal Thresholds of Hearing under Free Field Listening Conditions," ISO Recommendation R226 (1961 Dec.).

[20] D. W. Robinson and R. S. Dadson, "A Redetermination of the Equal-Loudness Relations for Pure Tones," *Brit. J. Appl. Phys.*, vol. 7, pp. 166–181 (1956 May).

[21] H. Fletcher and W. A. Munson, "Loudness, Its Definitions, Measurement and Calculation," *J. Acoust. Soc. Am.*, vol. 5, pp. 82–108 (1933 Oct.).

[22] D. W. Robinson and L. S. Whittle, "The Loudness of Octave-Bands of Noise," *Acustica*, vol. 14, pp. 24–34 (1964).

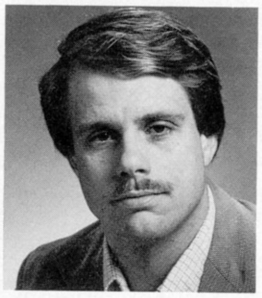
[23] D. W. Robinson, L. S. Whittle, and J. M. Bowsher, "The Loudness of Diffuse Sound Fields," *Acustica*, vol. 11, pp. 397–404 (1961).

[24] C. P. Lebo and K. S. Oliphant, "Music as a Source of Acoustic Trauma," *J. Audio Eng. Soc.*, vol. 17, pp. 535–538 (1969 Oct.).

[25] "Pre-emphasis Used on Sound Program Circuits in Group Links," CCITT, Red Book III, Fascicle III. 4, Rec. J.17 (1972).

[26] L. D. Fielder, "The Audibility of Modulation Noise in Floating Point Conversion Systems," *J. Audio Eng. Soc.*, Submitted for publication.

THE AUTHOR



Louis Fielder received a B.S. degree in electrical engineering from California Institute of Technology in 1974 and an M.S. degree in acoustics from UCLA in 1976. From 1976–1978 he worked on electronic component design for custom sound-reinforcement systems at Paul Veneklasen and Associates. From 1978–1983 he was involved in digital-audio codec and magnetic recording physics research at the Ampex Corporation. Recently Mr. Fielder has been working on digital-audio development at Dolby Laboratories.