Audio Engineering Society

Convention Paper

Presented at the 123rd Convention
2007 October 5–8 New York, NY

Filter Reconstruction and Program Material Characteristics Mitigating Word Length Loss in Digital Signal Processing-Based Compensation Curves used for Playback of Analog Recordings

R. S. Robinson

Channel D Corp., Trenton, NJ 08618, USA

Correspondence should be addressed to the author at rrobinson@ieee.org

ABSTRACT

Renewed consumer interest in pre-digital recordings, such as vinyl records, has spurred efforts to implement playback emphasis compensation in the digital domain. This facilitates realizing tighter design objectives with less effort than required with practical analog circuitry. A common assumption regarding a drawback to this approach, namely bass resolution loss (word length truncation) of up to approximately seven bits during digital de-emphasis of recorded program material, ignores the reconstructive properties of compensation filtering and the characteristics of typical program material. An analysis of the problem is presented, as well as examples showing a typical resolution loss of zero to one bits. The worst case resolution loss, which is unlikely to be encountered with music, is approximately three bits.

1. INTRODUCTION

Renewed consumer interest in pre-digital audio music recordings, such as "vinyl" LP records [1-4], along with the associated playback hardware, has spurred efforts to implement playback frequency response compensation curves in the digital domain.

Vinyl records produced within about the last 50 years employ the standardized RIAA emphasis curve. As is generally known to those working in the field, this curve dictates a rising high frequency emphasis of approximately 20 dB at 20 kHz and a similar low frequency attenuation at 20 Hz, applied before transcribing the vinyl disc (Fig. 1). The playback of such material requires the application of a complementary compensation curve to restore the proper frequency balance. This is usually done with circuitry that is integrated into the preamplifier used to properly condition the phonograph signal.

Eschewing compensation in the analog domain for its digital counterpart permits realizing closer agreement with the specified compensation curve, with less cost and effort compared to practically attainable analog circuitry [5]. The accuracy of digital compensation can be engineered to achieve nearly any arbitrary specification limit. Applying compensation in the digital domain also confers immunity to analog filter component tolerances and their variation due to temperature and aging. For a multichannel signal, all channels will have identical compensation (amplitude and phase response). Further, implementing

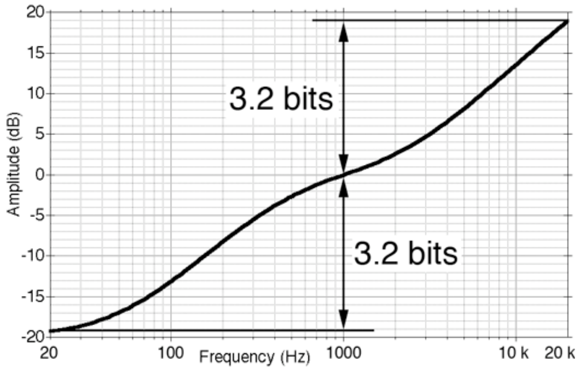

Fig. 1: Plot showing the amplitude *vs.* frequency characteristics of the RIAA emphasis curve used for vinyl recordings. This also describes the potential additional headroom requirement for digital transcription without compensation (*e.g.,* in the analog domain). The magnitudes of high frequency emphasis and low frequency de-emphasis, in digital bits, are indicated.

compensation curves in software makes it easier to accommodate the plethora of antique recordings and their arcane compensation curves, which predate the RIAA standard.

Finally, the restoration (such as the removal of transient clicks or pops) of damaged recordings for archival or other purposes is facilitated. While the music signal has been subjected to a pre-emphasis curve, subsequent physical defects in the recording medium have not. Therefore, the compensation curve distorts and stretches a defect's transient signal. If defects are excised prior to application of the compensation curve, artifacts resulting from their removal will be minimized. While this feature isn't addressed in detail in this paper, it is worthy of consideration. (Also not addressed here, though applicable to the topic of the paper, is the digital playback compensation of magnetic tape recordings.)

2. WORD LENGTH TRUNCATION

When playing or transcribing such recordings digitally, without first applying the usual analog de-emphasis curve, there is an implicit requirement of allowing sufficient recording headroom in the treble to accommodate the 20 dB of signal emphasis. This corresponds to slightly more than three bits of digital dynamic range (upper part of Fig. 1).

Likewise, the de-emphasis of the low frequencies implies that program material will be incapable of using the entire digital dynamic range in the bass region, given the headroom needed in the treble plus the 20 dB of bass de-emphasis. This also corresponds to slightly more than three bits.

This analysis suggests that digital transcriptions made without first applying analog frequency compensation invite a potential loss of nearly seven bits of digital dynamic range. However, this conclusion omits the consideration of the reconstructive effects of the low-frequency emphasis curve (digital filtering), plus the frequency balance of typical musical program material.

3. SIGNAL RECONSTRUCTION

First, consider the reconstruction filter applied to a one-bit, high sample rate (multi-MHz), digital audio data stream (such as the so-called DSD), for proper playback. A lower sample rate, albeit higher resolution (word length) data stream results from the filtering operation. Likewise, applying low frequency emphasis and high frequency de-emphasis filters to a digitally transcribed recording causes a similar signal summation, increasing the effective word length, in proportion to the relative amount of high frequency attenuation applied. This counteracts the bass resolution loss caused by recording uncompensated program content.

This can be demonstrated by a worst case (in terms of resolution truncation) example, starting with a digital full scale, 24-bit 20 Hz sine wave (sample rate of 192 kHz, top waveform shown in Fig. 2). Next, the signal samples were attenuated digitally by 40 dB, simulating the approximate magnitude of the RIAA emphasis curve from treble to bass, causing word length truncation. The samples were normalized to digital full scale by applying a scalar gain (and rounding to the nearest integer during normalization). The truncated signal was then subtracted from the original waveform. Examination of the amplified (by an additional 104 dB) result (Fig. 2, center) shows the expected, non-null error due to word length truncation.

The peak to peak modulation level of the residual error waveform in Fig. 2 (center), after applying the additional 104 dB gain, is -0.3 dB below full scale (corresponding to a modulation level of -104.3 dB for the error signal). Considering that the dynamic range of a 24-bit signal is approximately 144 dB, the *ca.* 40 dB difference corresponds to the magnitude of the word length truncation.

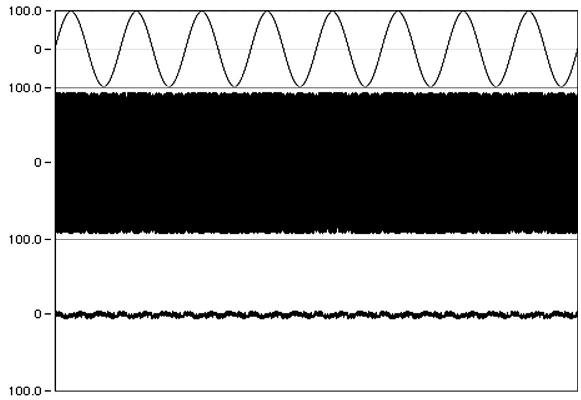

Fig. 2: Percentage modulation *vs*. time (400 ms full scale) of 24-bit, 20 Hz (sample rate 192 kHz) sinusoidal input and amplified (104 dB) residual signals. *Middle,* amplified residual resulting from subtraction of truncated-resolution signal from input signal; *bottom,* amplified residual resulting from subtraction of truncated-resolution, inverse-RIAA filtered signal from input signal. Magnified views of waveforms are shown in Fig. 3.

3.1 Reconstruction from Filtering

Repeating the above procedure, instead of normalizing the samples after truncation, the digital samples were processed by a digitally implemented inverse-RIAA compensation filter. The output of the filter, after carefully matching gain and correcting for the group delay at 20 Hz, was converted from the filter's floating-point output to 24 bit integer format before subtraction. A residual error (difference) waveform with noticeably reduced amplitude is obtained (Fig. 3D and bottom, Fig. 2).

Fig. 3 shows magnified views of the residual signals from Fig. 2. The curves in Figs. 3A and 3B were created at 16-bit resolution to better illustrate the relationship between the original, truncated and residual error waveforms. The positive polarity peak amplitude for the original and truncated waveforms in Fig. 3A is 32767. Close examination of Fig. 3B (the difference) shows that this peak occurs at only one point, at the wave crest.

Fig. 3C is similar to Fig. 3B, except transitions in the difference signal are spaced more closely together because of the smaller steps of the 24-bit waveform. The peak to peak amplitudes of Figs. 3B and 3C are similar, despite starting with different waveform resolutions, because the amount of word length reduction (40 dB) was the same in both cases.

Fig. 3D shows a highly magnified view of the residual error (difference) waveform. The plot scale is ten times more sensitive than Fig. 3C, to better show detail. The magnitude of the error is on the order of only a few LSBs, out of 24 bits. Each clearly resolvable amplitude step in the waveform is just one LSB out of a maximum range from 8388607 to -8388608. The peak to peak modulation of the error waveform (Fig. 3D and bottom, Fig. 2) is significantly (24.6 dB) smaller than the uncompensated error waveform (Fig. 3C and middle, Fig. 2).

This proves that the word length truncation in the bass will be over four bits *less* than that expected from a first order consideration of the digital de-emphasis process. One obtains not only complete recovery of lost bass amplitude resolution (caused from the RIAA bass de-emphasis part of the curve alone) *via* the digitally applied RIAA compensation filter, but even a slight enhancement of resolution in the bass.

When beginning with 16 bit data, the word length recovery of the truncated waveform by the compensation filter was slightly less effective, about 20.0 dB (now unimportant, given the widespread availability of modern 24 bit A/D converters). Finally, the above example considers a frequency of 20 Hz. Frequencies between 20 Hz and 1 kHz, where the compensation filter gain gradually diminishes to unity, would have reduced word-length recovery, but also would have reduced truncation from de-emphasis to contend with, as well.

4. PROGRAM CHARACTERISTICS

In the high frequency region of the spectrum, the degree of extra headroom needed for recording uncorrected program material will depend on the frequency balance of the program material. The amplitude of program content above 1 kHz is critical, because that frequency range is pre-emphasized and will impact the effective dynamic range of lower frequencies. Therefore, it's useful to know about the frequency balance of typical, actual music LP recordings that have been emphasized with the RIAA compensation curve.

Efforts were made to find examples of recordings with the greatest amount of high frequency emphasis, as well as more generally representative recordings, from

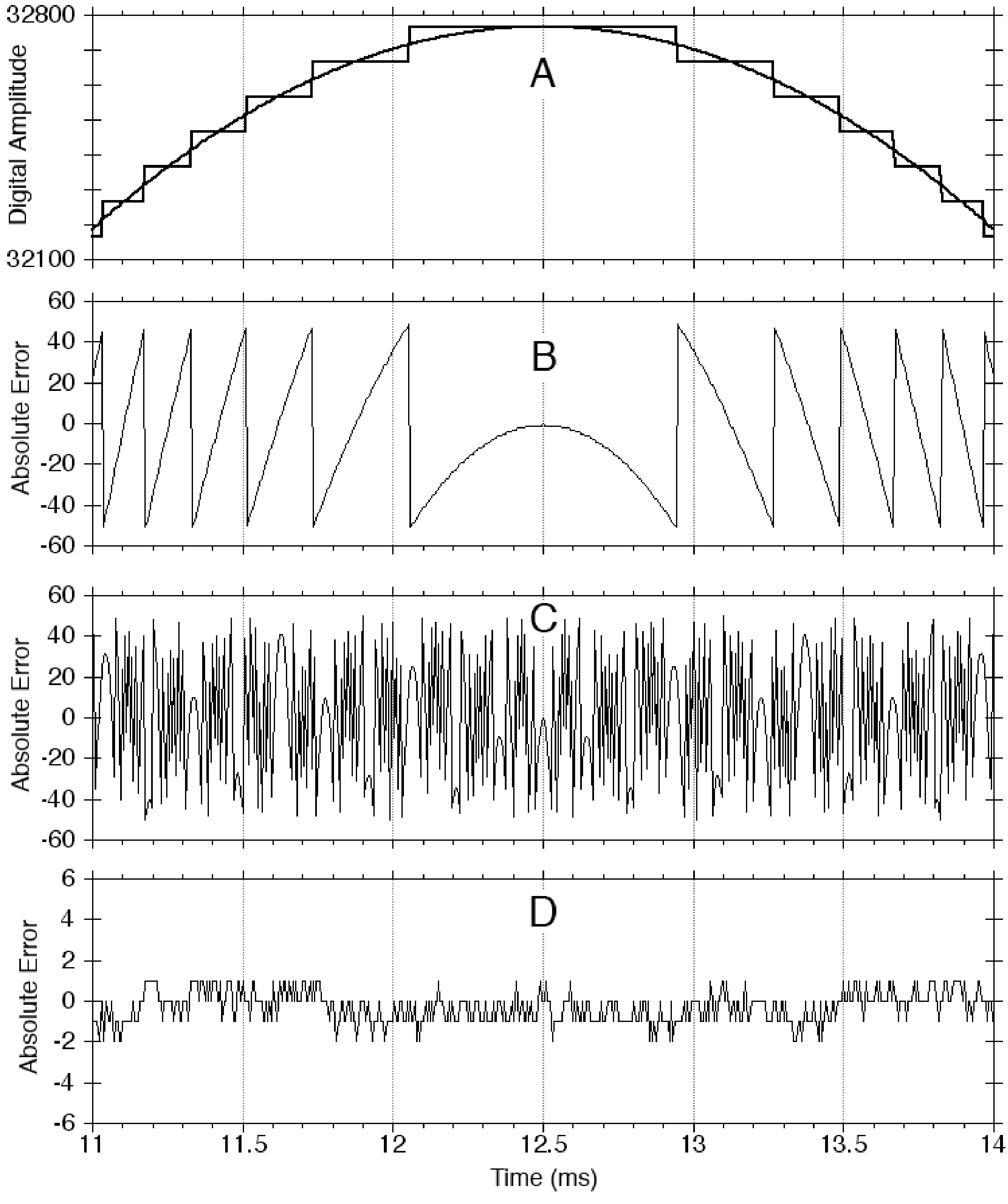

Fig. 3: Single-sample resolution (192 kHz sample rate) test signals. **A** is a magnified view of the crests of 16 bit resolution, 20 Hz sinusoidal signals. The sine wave's positive-going, zero-crossing point is located at time coordinate zero. The stair-stepped, truncated-resolution signal was created from the smooth original, as explained in the text. **B** shows the absolute (digital sample amplitude) difference resulting from subtracting the truncated signal from the original. **C** and **D** are magnified views of the residual (difference) waveforms from Fig. 2, which were created from 24 bit source data. **C** and **D** are similar to **B,** except for the original signal resolution. In **D,** the truncated source signal was digitally inverse-RIAA filtered instead of being linearly normalized.

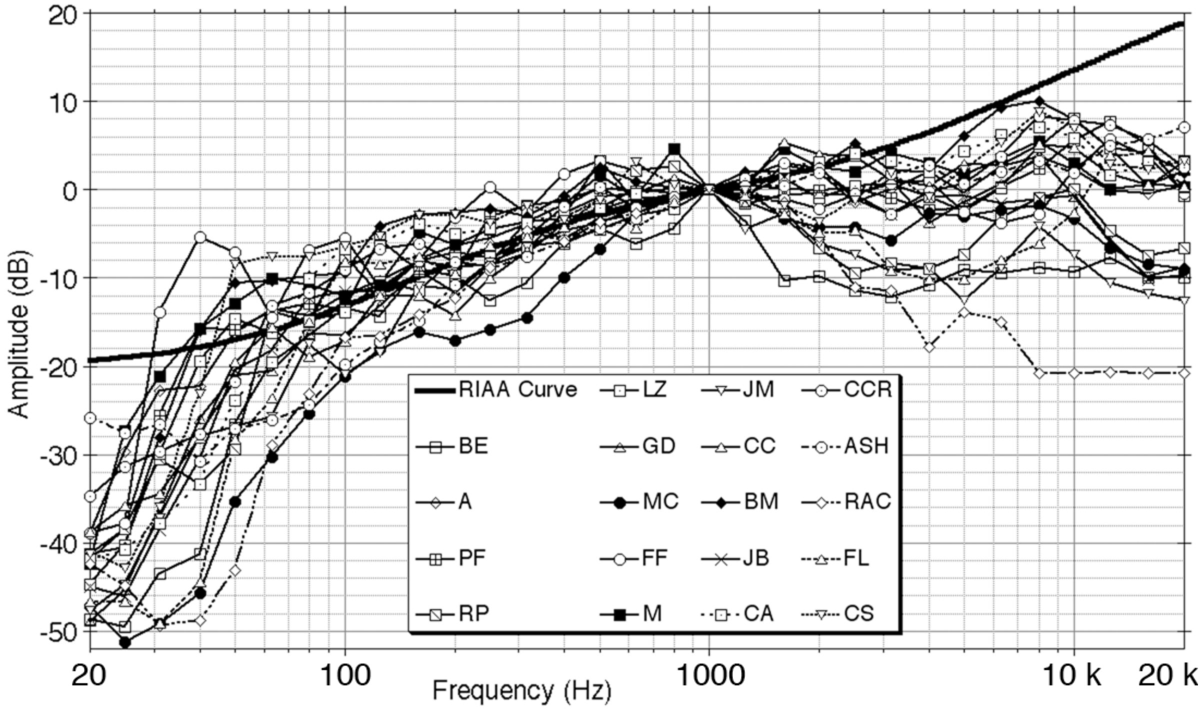

Fig. 4: Normalized (to the amplitude at 1 kHz) frequency balance of digital recordings made from vinyl LPs, measured with peak-responding, peak-hold third-octave audio analyzer software. The recordings were made with a linear preamplifier (without RIAA compensation). All recordings listed below in Table 1 are plotted here, to show the frequency balance envelope. Selected recordings from Fig. 4 are examined in more detail in Fig. 5.

Key	Artist	Album	Label	Peak (dB)
BE	Bill Evans	Live at the Village Vanguard	Riverside (reissue)	9.1 (Left)
A	Ambrosia	Ambrosia	20th Century	13.5 (Left)
PF	Pink Floyd	Wish You Were Here	Columbia	14.6 (Right)
RP	Rebecca Pidgeon	The Raven	Chesky	11.3 (Left)
LZ	Led Zeppelin	Houses of the Holy	Classic (reissue)	15.0 (Right)
GD	Grateful Dead	Workingman's Dead	Warner Brothers	13.3 (Right)
MC	Maria Callas	Lucia di Lammermoor	Angel	9.3 (Left)
FF	Frederick Fennell	Holst, Handel, Bach/Cleveland Symphonic	Telarc Digital	14.0 (Left)
M	Magazine	Secondhand Daylight	Virgin (U.K.)	15.2 (Left)
JM	Joni Mitchell	Blue	Rhino (reissue)	8.2 (Right)
CC	Clifton Chenier	Clifton Chenier’s Very Best	Blue Thumb	14.9 (Left)
BM	Bob Marley & the Wailers	Natty Dread	Island	17.5 (Right)
JB	Jeff Beck	Blow By Blow	Epic	11.6 (Left)
CA	The Cars	The Cars	Elektra	15.9 (Right)
CCR	Creedence Clearwater Rev.	Cosmo's Factory	MFSL (reissue)	16.0 (Left)
ASH	Ash	1977	Infectious (Germany)	14.8 (Left)
RAC	Svatioslav Richter	Rachmaninoff (Op. 23 & 32 Preludes)	MHS	7.6 (Right)
FL	Nicholas Zumbro	Liszt Piano Concerto No.1 in E-flat Major	MHS	12.1 (Right)
CS	Cat Stevens	The Teaser and the Firecat	Universal (reissue)	15.7 (Left)

Table 1: List of recordings used in Fig. 4, with Key corresponding to curves shown in Figs. 4-7. The Peak column indicates the peak *overall* level of the recording, normalized with respect to the 1 kHz band.

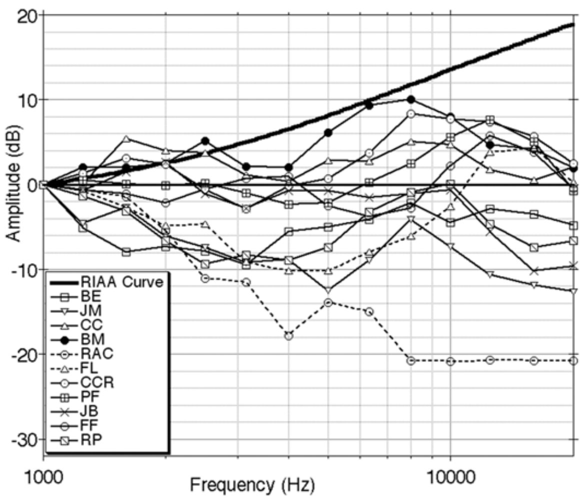

Fig. 5: Peak responding third-octave frequency balance of recordings from Fig. 4, focusing on the region above 1 kHz.

a selection of thousands of vinyl LP recordings available to the author. From those, a selection of 40 or so LPs was made, and digital transcriptions of those LPs were analyzed for frequency balance. From these measurements, useful observations could be made about program content that bear upon the subject of digitally implemented compensation curves.

4.1 Peak Responding / Peak Hold

Shown in Fig. 4 are examples of third-octave peak-responding, peak hold amplitude *vs.* frequency spectra of recordings representing various genres of musical program material, recorded without RIAA compensation, reflecting the signal presented to the A/D converter. (In all examples presented here, the most demanding of the two stereo channels of the recordings, in terms of having the most high frequency content relative to 1 kHz, is shown.)

The abbreviations used to identify the curves in Figs. 4 to 7 are listed in Table 1. Fig. 4 is self explanatory in showing that most music recordings lack significant (more than 6 dB greater than the 1 kHz amplitude) high frequency content, compared to the 1 kHz region, despite the emphasis from the RIAA curve (which also is shown). 6 dB of extra recording headroom corresponds to the loss of only 1 bit of digital dynamic range. The frequency balance of recordings selected from Fig. 4 showing extreme (not representative) and moderate behavior are shown in more detail in Fig. 5, focusing on the content above 1 kHz.

The boundary of the RIAA emphasis curve is useful from the perspective of considering the effect of the treble emphasis and bass attenuation on the program material. However, it also establishes a boundary that delineates the excess headroom required due to the RIAA treble emphasis. A signal having a frequency balance greater in amplitude than the treble emphasis curve has no additional impact on the treble headroom required for digital conversion/compensation, because the excess also will raise the treble balance (above the 1 kHz level) of a signal that has been compensated before digital conversion. In a similar way, having an excess of bass amplitude (“above” the RIAA curve) in the uncompensated signal will improve the word length resolution in the bass. In such a signal that has compensation applied before analog to digital conversion, this will limit the resolution in the treble, as explained below. Incidentally, many of the recordings in Fig. 4 exhibit this kind of frequency balance.

4.2 Long-Term Average

There are a few additional issues to consider regarding the peak responding analysis. First, a long-term RMS averaged spectrum might be viewed as more indicative of the statistical incidence of resolution truncation caused by high frequency program content. However, the overall available digital headroom is determined *only* by the highest peak in the program material, so the level of truncation can’t be inferred from a long-term averaged spectrum analysis.

4.3 Frequency Balance at Position of Peak Overall Signal Amplitude

A disadvantage of the peak responding / peak hold spectrum analysis is that the frequency balance information represented at particular locations (time coordinate) in the recording, such as at a time-domain amplitude peak, is masked by the aggregate spectrum.

Program peaks result from the coherently summed contributions of bass, midrange and treble. This is important, because a program peak with significant high frequency content will quantitatively determine the resolution truncation from treble emphasis for that particular recording (because frequency emphasis then will affect the level of the peak, which in turn will determine the overall linear analog gain to be

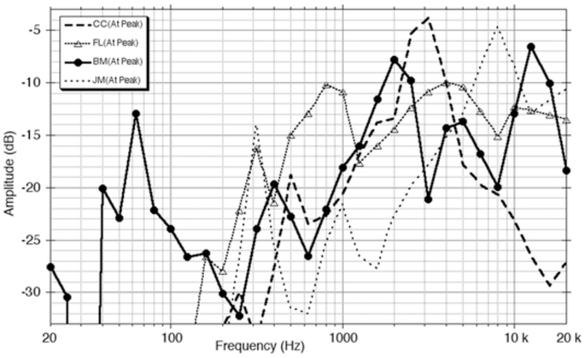

Fig. 6: Frequency balance measured at the specific time offset of the overall peak signal amplitude in the indicated recordings, with the curve amplitudes normalized to the overall (integrated) peak signal amplitude value.

employed ahead of the A/D converter when making a digital transcription).

Fig. 6 shows the peak frequency balance measured at the specific location (in time) of the peak amplitude of those recordings which were found to exhibit greater degrees of high frequency emphasis (as shown in Figs. 4 and 5). The plot is scaled to show the uppermost 30 dB of amplitude, and the curves were normalized to the overall peak level of the recording (because their coherent sum determines the peak level). Generally, for recordings exhibiting greater high frequency emphasis, the overall program peak correlated with a peak in high frequency content.

The Bob Marley and the Wailers recording is unique in that the *program peak level* simultaneously contains significant amounts of energy in more than one frequency region: low bass (in the 63 Hz, third-octave band), midrange (2 kHz) and treble (12.5 kHz). The Liszt piano concerto also was unusual in that the emphasis of the high frequency balance at the program peak position is at a lower frequency (4 kHz), compared to that for the entire recording (16 kHz, Fig. 5). Most of the other recordings also have significant high frequency content at the signal peak, but exhibit only one prominent frequency peak in the signal at the location of the overall signal peak, as shown, for example: Clifton Chenier (3.15 kHz), Joni Mitchell (8 kHz), Led Zeppelin (10 kHz), Pink Floyd (10-12.5 kHz); and those not shown in Fig. 6: Magazine (6.3-8 kHz), Ash (8 kHz), Creedence Clearwater Revival (10 kHz). This supports the validity of the peak responding / peak hold analysis presented above; all peaks in Figs. 4 and 5 represent instantaneous values at particular frequencies at different times in the recording and don't necessarily sum coherently and simultaneously; otherwise, the result might be more akin to noise, not music.

4.4 Upper Bound of Treble Headroom

Considering the foregoing, the upper bound of the extra treble headroom needed will be determined by the coherent sum of the instantaneous amplitudes of the frequencies in the recording. This can be estimated by comparing the amplitude difference between the 1 kHz frequency band (where the emphasis gain is unity) with the overall peak level of the recording. This is reported in the last column of Table 1, as the peak level (after normalizing to the level in the 1 kHz band). The *upper bound* to treble-induced truncation will be the number of bits corresponding to that peak level. For the exceptional example (Bob Marley and the Wailers) among the recordings surveyed, and shown here, this was a maximum of 17.5 dB (2.9 bits). There also is the possibility that other, more extreme such examples exist and could be located, but they will be uncommon.

Given the results from the variety of recordings surveyed (most of which aren't presented here), and considering that the above figure represents an *upper bound,* a more likely typical loss of digital dynamic range due to treble emphasis, from the peak responding / peak hold analysis, is less than one bit (6 dB). It appears to be unlikely, in practice, to exceed this figure with LP music recordings on a consistent basis.

5. ANALOG COMPENSATION

Finally, it's worth considering the effect of using analog de-emphasis compensation in the signal chain prior to transcribing recordings digitally, such as might typically be done for archival / restoration purposes or format conversion. Accordingly, with the above recordings, a similar analysis was performed, except the recordings were analyzed after applying RIAA compensation.

Keeping in mind that the peak signal level will determine the maximum recording level at the input of the A/D converter, a different result is obtained with the pre-compensated signal. Fig. 7 shows that in all cases considered here, the overall gain structure of the digital transcription will be determined not by the treble, as in the uncompensated case, but by the program content in the low bass to midrange

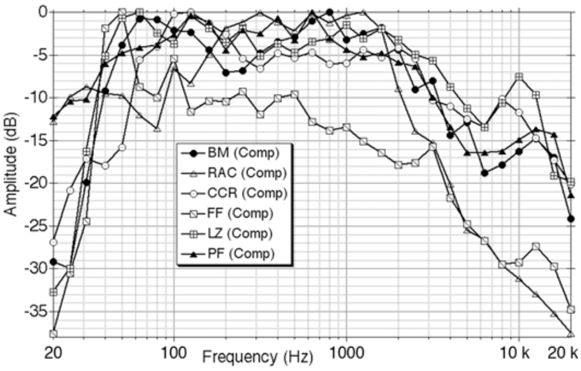

Fig. 7: Peak responding third-octave analysis of the frequency balance of selected LP recordings from Fig. 4, after applying the RIAA compensation curve.

region. Therefore, the argument made regarding the loss of bass resolution because of the RIAA emphasis curve, also applies: but in reverse, because the de-emphasized treble is being recorded at lower amplitude (and lower resolution) than the bass/midrange region. Here, the effective word length will be truncated at *treble* frequencies (which cannot benefit from the word length reconstruction of the digital compensation filtering). Since human hearing has been demonstrated to have a higher sensitivity in the midrange and treble [6, 7], this actually comprises a less desirable situation than word length truncation in the bass region.

Setting aside the benefits of word length recovery in the bass because of digital compensation, as proven above: if truncation does occur, bass resolution truncation represents a more desirable alternative to resolution truncation in the midrange and treble; further, the perception of any “disadvantage” to word length truncation in the bass is unfounded, because of being more than compensated by the feature of digital resolution *enhancement* in the treble.

6. CONCLUSION

Digitally implemented de-emphasis filtering, used for playback of vinyl LP recordings, enhances the digital word length at low frequencies, similar to the reconstruction filters used for the playback of one-bit digital audio streams. The reconstructive properties of digital de-emphasis filtering, as shown here, in conjunction with the characteristics of most program material, will cause a typical overall bass resolution truncation of only one bit or less, significantly lower than the nearly 7 bits predicted by a casual analysis of digitally realized frequency emphasis playback compensation. Exceptional, uncommonly encountered program material may cause a worst case bass truncation of approximately three bits, which is negligible considering the 24 bit resolution capability of modern analog to digital converters. In cases where bass word length truncation does occur, the disadvantage is balanced by the complementary *enhancement* of digital resolution, due to treble pre-emphasis, in the frequency range where human hearing is at its most sensitive.

7. REFERENCES

[1] Sepulvado, John, "Putting a New Spin on Vinyl Records," National Public Radio broadcast, April 16, 2007; reports of sales of new vinyl records up 10% in the first quarter 2007.

[2] Record Industry Association of America sales figures current as of 2006 show that vinyl sales, while steady as of the date of the 2006 report, have rebounded by between a factor of 2 to 3 (as a percentage of total music sales) compared to their market share minimum in 1993.

[3] "Renewed energy in the vinyl and turntable business," quote from a speech by National Association of Recording Merchandisers president Jim Donio, April 30, 2007.

[4] "Back in the Groove: young music fans ditch downloads and spark vinyl revival," Guardian (U.K.), July 16, 2007.

[5] An informal survey of high quality phonograph preamplifiers specifying close agreement with the RIAA curve (less than plus or minus 0.1 dB), where available, indicates a cost comparable to but sometimes significantly greater than a high quality linear preamplifier and analog to digital interface.

[6] D. W. Robinson and R. S. Dadson, *Brit. J. Appl. Phys.* **7**, 166-181 (1956).

[7] H. Fletcher and W. A. Munson, *J. Acoust. Soc. Am.* **5**, 82-108 (1933).