Audio Engineering Society

Convention Paper

Presented at the 117th Convention
2004 October 28–31 San Francisco, CA, USA

AES Technical Committee on Signal Processing Educational CD Project

Robert C. Maher

Electrical and Computer Engineering, Montana State University, Bozeman, MT 59717-5780 USA
rob.maher@montana.edu

ABSTRACT

The AES Technical Committee on Signal Processing is developing a compact disc with educational material and demonstrations intended for students, educators, and working digital audio engineers. The material includes examples of quantization and dither, basic psychoacoustics, and practical DSP. The multi-mode CD will have both audio tracks and a CD-ROM section with a web-browser interface. The CD will be produced for sale by the AES Publications office.

1. INTRODUCTION

Digital audio systems have become a commodity product in recent years due in large part to the ingenuity and effort of Audio Engineering Society members. However, the maturity of the digital audio signal processing field can present a serious challenge for new students, researchers, software engineers, product testers, and other neophytes who must try—in essence—to interpret the lofty mountain summit without the benefit of having time to scout the approach routes.

Furthermore, it is difficult for new digital audio engineers to understand the relevance and relative importance of various signal processing parameters without being able to *hear* and *see* the details and effects: simply reading a written description is not the best way to learn and understand audio processing issues. It is important for AES to support and encourage education and learning at all levels, and signal processing is a very appropriate topic.

Since early in 2001, the AES Technical Committee on Signal Processing (TC_SP) has been working on a project to produce a CD-ROM containing educational demonstrations of basic digital audio signal processing concepts. Several examples and demonstrations have already been produced over the last two years, and the technical review and revision process will commence soon. This paper provides an update on the tentative contents and status of the CD project.

This effort is modeled after the highly successful CD produced by the AES Technical Committee on Coding of Audio Signals entitled "Perceptual Audio Coders: What to Listen For," edited by Marcus Erne [1]. Like the perceptual coding CD-ROM, the TC_SP CD will contain both a conventional audio (red book) section and a computer-readable data section with HTML-based chapters for viewing with a web browser.

This paper is organized as follows. First, an overview of the CD contents and rationale is given. Next, the tentative contents of each section is described in more detail. Finally, some specific examples of the CD contents are presented.

2. OVERVIEW

The TC_SP CD project was convened by a sub-committee of volunteer contributors and reviewers. The main objective has been to produce a CD-ROM that is useful for individual self-study learning and as a classroom instructional aid.

The basic examples start with simple experiential listening, such as comparisons between proper anti-aliasing prior to sampling and undersampling, various degrees of lowpass filtering, and the effects of A and C weighting filters for sound level measurements. The more sophisticated examples demonstrate the properties of dithered quantizers, frequency masking, and cascaded sample rate conversion effects.

2.1. Why Another Demonstration CD?

There are several notable audio demonstration CDs already available from commercial publishers and professional societies. These include the aforementioned AES-sponsored perceptual coding CD, the Acoustical Society of America "Auditory Demonstrations" CD (1987) [2], the NASA "Auditory Demonstrations in Acoustics and Hearing Conservation" and "Auditory Demonstrations II: Challenges to Speech Communication and Music Listening" compact discs [3], the "Digital Audio CD Resource Pack," a commercial digital audio training product (book and CD) authored by Marcus Erne [4], and numerous manufacturer-produced CDs with audio test and setup signals.

In light of the plethora of existing material, one might reasonably ask why another demonstration CD is needed. The answer is that the TC_SP CD is intended primarily to address issues relevant to signal processing algorithm designers and implementers.

While at one time it was only possible to implement audio signal processing algorithms using special-purpose hardware or arcane assembly language with a fast DSP microprocessor, it is now quite common to do audio processing with general-purpose microprocessors and high-level languages. This has made audio processing accessible to programmers who might not have had extensive experience with real time audio systems. Since a typical characteristic of audio digital signal processing algorithms is the need for sustained, uninterrupted processing, even a single sample dropout or parameter update error can result in an audible artifact. Detecting, diagnosing, and correcting this sort of implementation error often requires experience listening for the defects, and examples of this type are include in the TC_SP CD.

2.2. CD-ROM Organization

The demonstration material is divided into three major sections.

- The first section, entitled *Sampling, Reconstruction, and Simple Audio DSP*, contains material of an introductory nature.

- The second section, *Basic Psychoacoustics of Digital Audio*, contains a summary of the basic tone-on-tone and tone-on-noise masking phenomena and related critical band effects.

- The last section, *DSP Applied in Practical Digital Audio Systems*, is intended for programmers and hardware designers who must implement and verify audio signal processing algorithms. The examples include the effects of envelope parameter updates, common filter implementation errors, and sample rate conversion effects.

Among the many challenges of this CD project is finding a way to fit the many good ideas and suggestions of the sub-committee into the 650 megabytes available.

3. CONTENTS AND EXAMPLES

As of this writing (2004 August) the CD-ROM material is still in preparation or under review, so it is likely that some of the details will change before the final master is assembled. The preliminary list of topics is as follows.

3.1. First Section: Sampling and Reconstruction

The two most fundamental principles of conventional digital audio systems are (a) the sampling theorem, i.e., the relationship between the passband of the signal and the sampling rate, and (b) the resolution (quantization) of the digital signal representation. Both principles can be demonstrated with audio examples.

3.1.1. Sampling theorem and aliasing

The TC_SP CD presents the implications of the sampling theorem by the typical method of a proper lowpass bandlimited input signal that is sampled at a rate more than twice the signal's lowpass bandwidth, creating an unambiguous discrete-time representation, and then an improperly bandlimited input signal that is sampled at a rate that is less than twice the signal's bandwidth, resulting in audible spectral aliasing in the reconstructed signal [5, 6].

3.1.2. Lowpass, highpass, and bandpass filtering

The bandwidth implications of discrete-time sampling may also represent a rate vs. quality tradeoff. In other words, reducing the spectral bandwidth of the signal reduces the data rate required to represent the information, but often at the expense of the perceived quality of the reconstructed signal. This concept is illustrated by a music example that is processed by lowpass, bandpass, and highpass filters with a range of cutoff frequencies. In some example applications the effect of bandlimiting may be essentially inaudible, while in other cases the bandwidth choice may be more critical.

3.1.3. Amplitude quantization and dither

As long as the input signal remains within the inner levels of the quantizer, the quantization error is bounded by the size of one quantizer level. However, if the amplitude features of the signal are comparable in size to the quantizer level or if the signal is narrowband or repetitive, the quantization error will be highly correlated with the signal. The quantization error will also be audibly tonal, and the error level will be related in a nonlinear manner to the signal amplitude: both characteristics are generally undesirable.

The principles of amplitude quantization are demonstrated with the conventional method. First, a uniform quantizer with a small number of levels is used to represent a low-amplitude decaying sinusoid. As the amplitude decreases relative to the quantization levels, the familiar audible roughness and granularity becomes obvious [7].

In the second group of quantization examples, additive dither is used prior to the uniform quantizer in order to whiten the quantization error and decorrelate it from the signal. The dither examples are intended to be comprehensive, covering a variety of dither probability density functions (rectangular and triangular) and quantizer types (mid-tread and mid-riser, with and without noise shaping) [8, 9].

3.1.4. Hard and soft clipping

If the input signal amplitude exceeds the uniform inner level range of the quantizer, the quantized output value becomes limited to the quantizer's outer levels and the quantization error becomes unbounded. Dither cannot be a remedy in this situation since the signal is clipped and information is unavoidably lost.

This section of the TC_SP CD includes several brief examples demonstrating the audible distortion effects of signal clipping. The first set of examples compare the case of a signal that has been clipped in the continuous-time domain prior to bandlimiting and sampling, and the case of a signal that is clipped in the digital domain after being sampled. In the first case any out-of-band components are removed by the input antialiasing filter, while in the second case the clipping distortion is introduced after the analog input filter. The sound may be subtly different in these two cases.

While it is generally possible to avoid input quantizer clipping by controlling the input signal amplitude, it is more common for an algorithm implementer to encounter clipping as a numerical problem inside a digital signal processing algorithm where coefficients are multiplied and partial sums are accumulated. Examples of this sort are include in the third section on practical DSP (see below).

Finally, an example set illustrating 'soft' clipping is included for completeness (Figure 1). The soft clipping function is a fixed nonlinear warping that may help conceal the onset of clipping—but at the expense of inherent harmonic distortion.

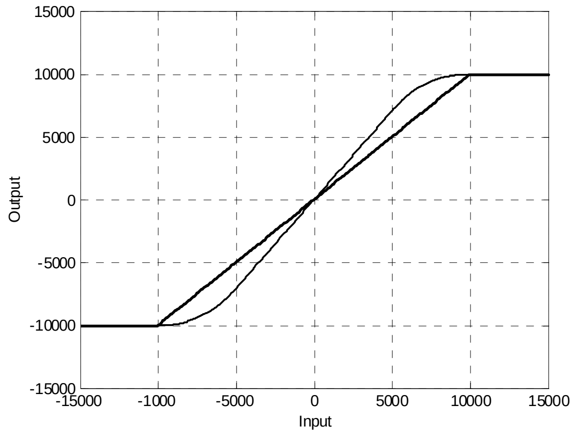

Figure 1: Hard clipping function (straight segments) and soft clipping function (curved segment)

3.1.5. Decibel scales and frequency weighting

The characteristics of audio signals and noise are often specified in *decibels* (dB). The *bel* is defined to be the base ten logarithm of a power ratio. The logarithm compresses the numerical range of its argument, and this is often a convenient feature when one must deal with numbers differing over several orders of magnitude [10].

Decibel measurements may be made using a frequency weighting filter, sometimes called a weighting network, to make the measurement depend on the frequency distribution of the signal's spectrum (Figure 2).

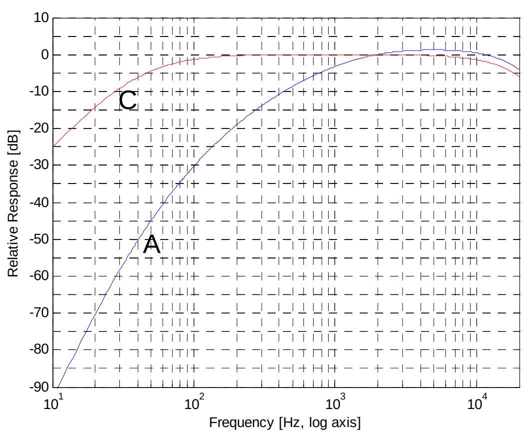

Figure 2: A-weighting and C-weighting filters

Understanding the characteristics of the weighting function is important for algorithm implementers because system performance metrics (e.g., signal-to-noise ratio and transducer sensitivity) can be expressed with weighted dB units [11].

The TC_SP CD examples include unfiltered broadband noise, A-weighted noise, and C-weighted noise. The standard A-weighting filter corresponds roughly to the average sensitivity of the human ear at low to moderate sound levels. A-weighting tends to emphasize spectral components in the 2 to 5 kHz range while reducing the contribution from lower and higher frequencies. The standard C-weighting filter approximates the ear's sensitivity at high sound levels, and therefore has a flatter response over much of the audio frequency band [12].

3.2. Second Section: Basic Psychoacoustics

In almost all cases the ultimate goal of audio signal processing is to produce sound that is pleasing to human listeners. Thus, a good understanding of the strengths and weaknesses of the human auditory system can be useful for algorithm designers and implementers. A complete treatment of psychoacoustical phenomena is beyond the scope (and the available space) of the TC_SP CD, but several basic principles are included.

3.2.1. Masking: tone-on-noise and noise-on-tone

Simultaneous masking is demonstrated using two scenarios: a high level sinusoidal tone masking a low-level narrow band of noise, and a high level narrow noise band masking a lower level sinusoidal tone.

For the tone masking noise case, the signal is a sequence demonstrating the audibility of noise in a narrow bandwidth (somewhat narrower than a critical band) around a sine wave. The sine wave is at 4 kHz and fixed, while the noise band is 4 kHz ± 400 Hz, with uniform amplitude and random phase. The level of the noise energy (total) relative to the sine wave starts at -40 dB and then increases in a sequence of 5 dB steps. The tone effectively masks the presence of the noise for the first several steps.

The noise masking tone example is similar, except the level of the noise is constant and the level of the tone is varied from –15 dB relative to the noise and then increases in several steps to 0 dB. It can be noted that in this case the tone is relatively inaudible compared to the noise: the noise is a more effective masker.

3.2.2. Audibility of phase

Basic signal processing elements such as filters and equalizers are usually specified in terms of their frequency and level accuracy only, since in many cases the human ear is thought to be relatively insensitive to phase differences. However, there are situations in which the effects of phase can become audible and obvious.

The demonstration signal contains a carrier frequency of 1 kHz with sidebands at ±25 Hz on either side of the carrier. The amplitude of the two sidebands is 0.25 that of the 1 kHz signal. In the first case the two sidebands are in phase, creating the equivalent of an amplitude modulation (AM) signal. In the second case the signal is constructed with the two sidebands out of phase, creating the equivalent of a frequency modulation (FM) signal. Although the power spectrum of the two signals is identical, they are easily distinguishable when listening.

3.2.3. Audibility of level change

Two signal sequences are used to demonstrate the audibility of level differences. First, a sustained tone with gradually increasing amplitude fluctuations is used to show the degree to which the level changes can be detected. The signal envelope is shown in Figure 3.

The second level audibility example is similar to the first, except a time gap is inserted between each level (Figure 4). The result is that the audible level comparisons with the time gap are less sensitive than for the sustained tone.

3.3. Third Section: Practical DSP

The concluding section of the TC_SP CD contains examples that are intended to be particularly relevant to algorithm designers and implementers. These implementation issues include parameter update rates, table lookup details, and diagnosing dropped samples and intra-filter overflow.

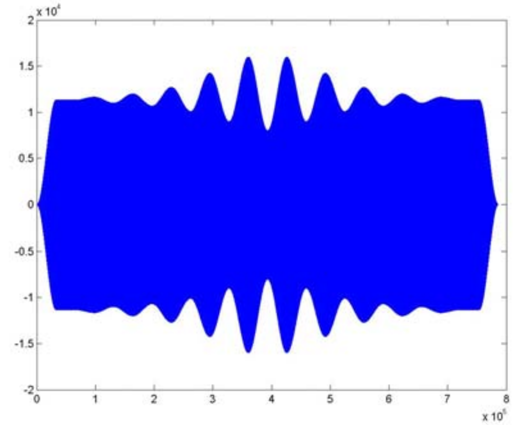

Figure 3: Continuous level variation example (amplitude vs. time)

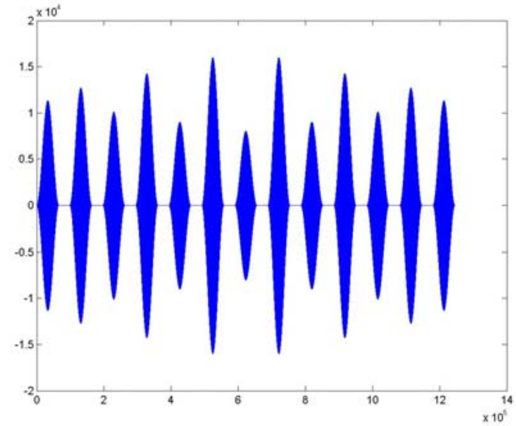

Figure 4: Level variation with time gaps (amplitude vs. time)

3.3.1. Updating time-varying parameters

For most signal processing algorithms at least one control parameter will need to change. The parameter could be a filter coefficient, pan setting, channel gain, etc., that is altered due to user input or some other external control signal. This leads to an implementation question regarding how quickly the parameter can be allowed to change without leading to an undesirable audible click or numerical instability. The CD examples show that even for a parameter as simple as gain, care must be taken to avoid audible artifacts.

The examples demonstrate a gain increase of 18 dB occurring in a set of discrete steps over a 100 millisecond period. At one extreme the gain is increased in four abrupt steps, and at the other extreme the gain is increased with 64 smooth ramp steps, each over 32 samples (approximately 0.726 milliseconds). Both linear and logarithmic ramps are presented. The examples demonstrate that coarse gain adjustments can cause audible increments known as *zipper noise* in the output signal. The tradeoff is the amount of computation required to update the parameter values smoothly versus the time required to perform the update.

3.3.2. Table lookup and interpolation

Many digital audio signal processing applications use tables of stored data, control envelope tables, and wavetables for signal synthesis. Common implementation decisions have to do with how large the data tables need to be, whether the stored data can be interpolated, and the spectral implications of repetitive wavetable sampling [13]. This section includes examples of wavetable signal generation and implementation issues.

3.3.3. Diagnosing dropped-sample problems

Among the most common debugging issues in a real-time signal processing system is a failure to process all signal samples in a timely manner. As the real-time input/output cannot stall while waiting for the signal processing system to catch up, at least one output sample will be incorrect and the output will contain an audible discontinuity. This section contains several examples of occasional and regular dropped sample problems that algorithm implementers are likely to encounter and need to diagnose.

3.3.4. IIR overflow propagation and limit cycles

Simple signal clipping is covered in the first section of the TC_SP CD, but another clipping problem can occur at a storage location within the recursive loop of an IIR filter if the numerical size of the signal samples grows to exceed the word size at a filter node. If the conditions are just right (or wrong), the IIR filter can become temporarily or permanently unstable—at least until the filter is re-initialized.

The recursive nature of the filter means that this internal clipping error will recirculate and linger in the filter for some period of time. An overflow limit cycle is a rather rare occurrence in IIR filters. What typically happens when an internal value overflows is that the output is distorted without throwing the filter into chaos. However, the distortion is bound to sound different than that of magnitude overflow outside of a filter. The distortion generated by the internal overflow is itself subject to filtering before it reaches the filter output

Examples of IIR overflow limit-cycle oscillations and internal clipping are included in this section.

4. CONCLUSION

The AES Technical Committee on Signal Processing expects to assemble the CD-ROM material in the next few months. The project participants—all volunteers—are completing the draft audio demonstrations and the audio example material has been donated by the copyright holders. Information on the complete table of contents and availability of the final CD-ROM will be forthcoming.

5. ACKNOWLEDGEMENTS

This work was supported by many organizations and volunteers. The contributions and efforts of several individuals are gratefully acknowledged: Ronald Aarts, Allison Banfield, John E. Bregar, Sydney Bulman-Fleming, Marcus Erne, James Johnston, Stanley Lipshitz, Vicki Melchior, Christoph Musialik, Werner Oomen, TC Electronic (Søren Nielsen), Duane Wise, and Bill McQuaide (AES HQ).

6. REFERENCES

[1] Erne, M., “Perceptual Audio Coders: What To Listen For,” Preprint 5489, presented at the AES111th Convention, New York, NY, USA, 2001 November 30 – December 3.

[2] Acoustical Society of America, *Auditory Demonstrations*, compact disc and guide booklet, Houtsma, A.J.M., Rossing, T.D., and Wagenaars, W.M., eds., 1987.

[3] NASA Glenn Research Center, *Auditory Demonstrations in Acoustics and Hearing Conservation*, and *Auditory Demonstrations II: Challenges to Speech Communication and Music Listening*, compact discs, 2004.

[4] Erne, M., *Digital Audio CD*, Focal Press, 1998.

[5] Oppenheim, A., Schafer, R., and Buck, J., *Discrete-Time Signal Processing*, 2nd ed., Prentice Hall, 1999.

[6] McClellan, J., Schafer, R., and Yoder, M., *DSP First: A Multimedia Approach*, Prentice-Hall, 1998.

[7] Maher, R.C., "On the nature of granulation noise in uniform quantization systems," *J. Audio Eng. Soc.*, vol. 40, pp.12-19, 1992.

[8] Lipshitz, S.P., and Vanderkooy, J., "Digital Dither: Processing with Resolution Far Below the Least Significant Bit," *Proc. AES 7th International Conference: Audio in Digital Times*, paper 7-014, 1989.

[9] Lipshitz, S.P., Wannamaker, R.A., and Vanderkooy, J., "Quantization and Dither: A Theoretical Survey," *J. Audio Eng. Soc.*, vol. 40, pp. 355-375, 1992

[10] Kinsler, Lawrence E., Frey, Austin R., Coppens, Alan B., and Sanders, James V., *Fundamentals of Acoustics*, 4th ed., Wiley & Sons, 1999.

[11] IEC 61672-1:2002: Electroacoustics-Sound Level Meters, IEC, Geneva, for A- and C-weighting filters.

[12] Eargle, John M., *Handbook of Recording Engineering*, 4th ed., Kluwer Academic Publishers, 2002.

[13] Massie, D.C., "Wavetable Sampling Synthesis," Chapter 8 in *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, eds., Kluwer Academic Publishers, 1998