Fundamentals of Modern Audio Measurement
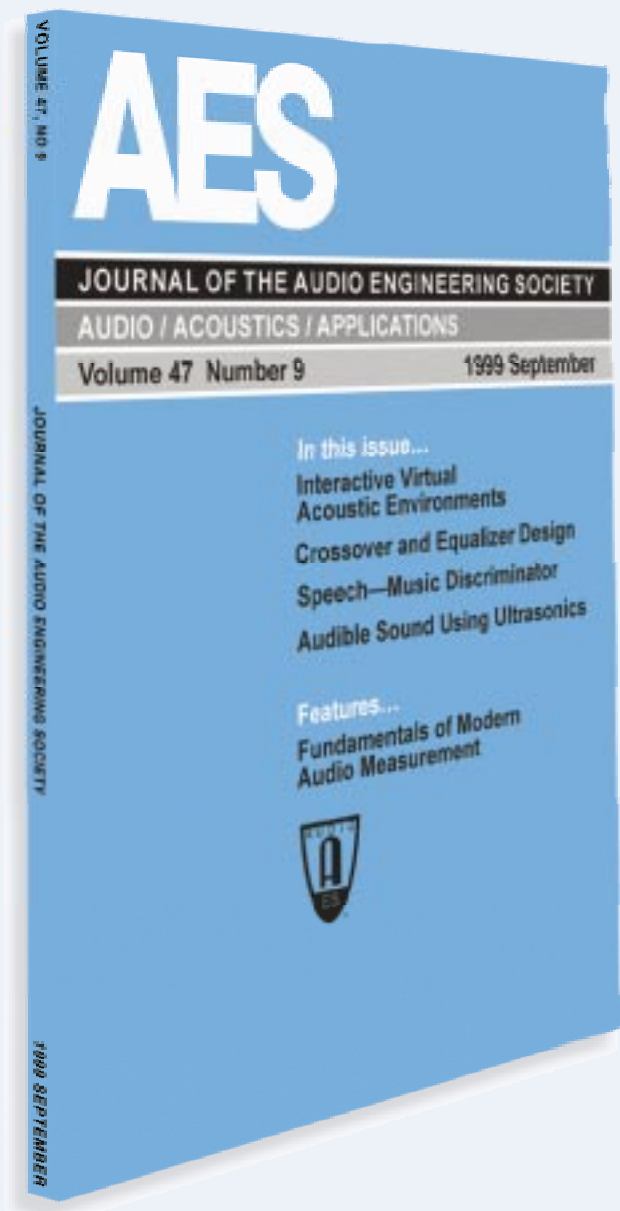
Richard C. Cabot, AES Fellow

Audio Precision, Inc. Beaverton, Oregon 97075, USA

AES

JOURNAL OF THE AUDIO ENGINEERING SOCIETY

AUDIO / ACOUSTICS / APPLICATIONS

Volume 47 Number 9   1999 September

In this issue...
Interactive Virtual
Acoustic Environments

Crossover and Equalizer Design

Speech—Music Discriminator

Audible Sound Using Ultrasonics

Features...
Fundamentals of Modern
Audio Measurement

Audio
precision ®

Fundamentals of Modern Audio Measurement*

Richard C. Cabot, AES Fellow
Audio Precision, Inc., Beaverton, Oregon 97075 USA

Fundamental concepts in testing audio equipment are reviewed, beginning with an examination of the various equipment architectures in common use. Several basic analog and digital audio measurements are described. Tradeoffs inherent in the various approaches, the technologies used, and their limitations are discussed. Novel techniques employing multitone signals for fast audio measurements are examined and applications of sampling frequency correction technology to this and conventional FFT measurements are covered. Synchronous averaging of FFTs and the subsequent noise reduction are demonstrated. The need for simultaneity of digital and analog generation is presented using converter measurements as an example.

Introduction

Characterizing professional and consumer audio equipment requires techniques which often differ from those used to characterize other types of equipment. Sometimes this is due to the higher performance requirements.

Other times it is due to the peculiarities of the audio industry. Other fields deal with some of the same measurements as those in audio. From level and THD to jitter and noise modulation, no other field has the breadth of requirements found in high performance audio.

Performing these measurements requires a knowledge of the tradeoffs inherent in the various approaches, the technologies used, and their limitations. We will examine these measure-

ments and their use in practical engineering and production applications.

Audio has been an analog world for most of its life. The last 15 years have seen a steady increase in the use of digital technology, including the digital recorder, digital effects units, the compact disc, digital mixing consoles and lossy data compression systems. Each has necessitated its own collection of new measurements for the new problems introduced.
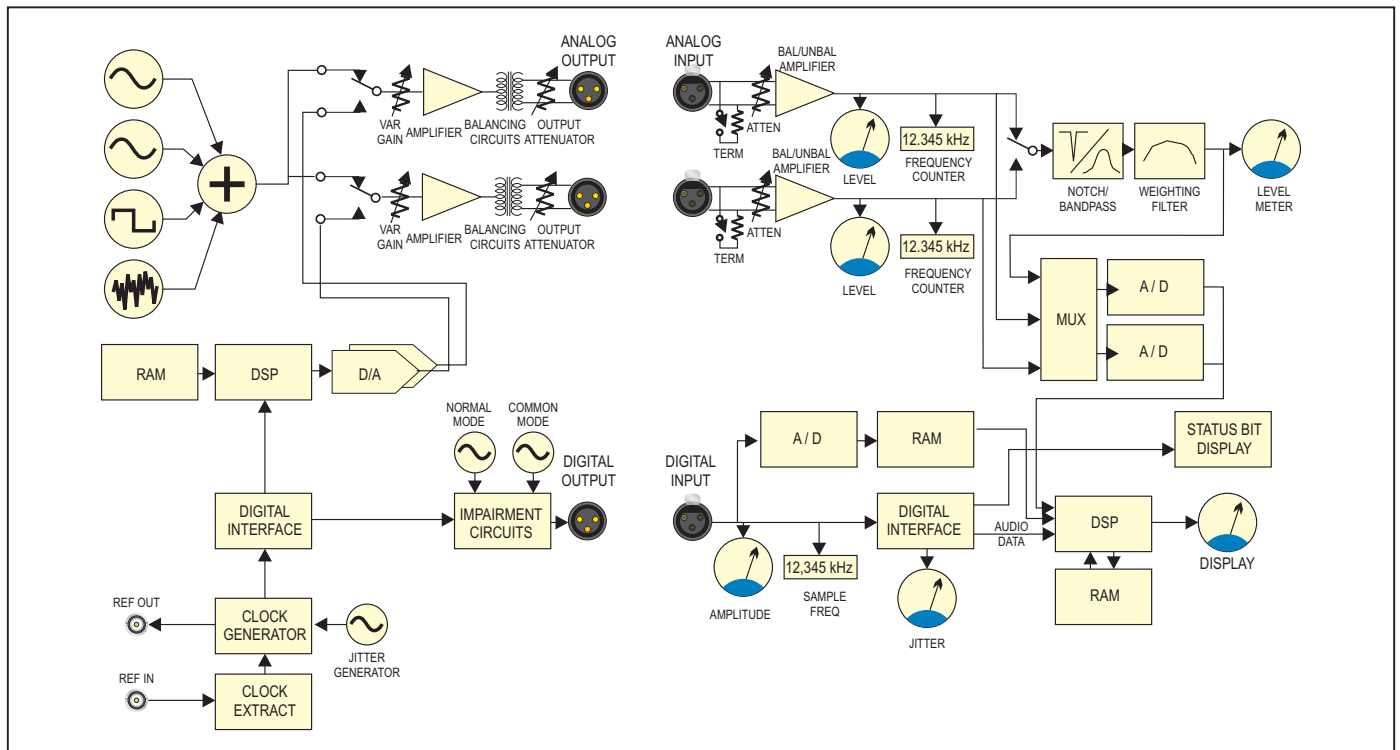

Fig. 1. Dual-domain audio measurement system.

Dual Domain Measurement

Characterizing modern audio equipment requires operating in both analog and digital domains. Measurement equipment and techniques for analog systems are well established (Metzler 1993). Signal generation was usually done with analog hardware signal generators. Signal measurement was usually done with analog filters and ac to dc conversion circuits. In recent years these were connected to microprocessors or external computers for control and display. In 1989, with the increasing prevalence of digital audio equipment, Audio Precision introduced the first Dual Domain[1] audio measurement system. It maintained the traditional use of analog hardware for analog signal generation and measurement, and added the ability to generate and measure digital audio signals directly in the digital domain. This allowed all combinations of simultaneous analog and digital generation and measurement, enabling the measurement of A/D converters, D/A converters, digital processing equipment, etc. in addition to the usual all-analog

[1]Dual Domain and System Two are trademarks of Audio Precision, Inc.

systems. By using internal A/D and D/A converters it also added the ability to perform many analog measurements which were previously not included in the system (such as FFT-based spectrum analysis and fast multitone measurements). This also allowed measurements which were previously impossible, such as bit error measurements on digital processing equipment which only have analog ports available. This was followed in 1995 by the next generation Dual Domain System Two (see Fig. 1).

Other manufacturers have introduced test equipment for measuring combined analog and digital audio equipment. One approach uses an AES-3 digital interface receiver circuit and a D/A converter in front of a conventional analog instrument to allow measuring digital signals. All measurements must go through the digital to analog reconstruction process and suffer the limitations of the converter and reconstruction filter used. This technique, illustrated in Fig. 2, allows an inexpensive, albeit less accurate, method of making measurements on digital domain signals. Some inherently digital measurements cannot be

done this way, such as active bits measurements and bit error rate measurements.

Another approach, used in several commercial instruments, is shown in Fig. 3. All signals are generated in the digital domain through dsp algorithms. If analog signals are needed, they are created by passing the digital signal through a D/A converter. Conversely, all signals are analyzed in the digital domain, and analog signals to be measured are converted by an internal A/D converter. This approach has the advantage of simplicity, since much of the measurement and generation hardware is re-used for all operations.

However, hardware simplicity comes at a price. The signal generation performance of current technology D/A converters is not equivalent to what can be achieved with high performance analog electronics. The measurement performance of A/D converters is similarly limited by available devices. Indeed, it is difficult to characterize state-of-the-art converters when the equipment performing the measurements uses commercially available converter technology. These
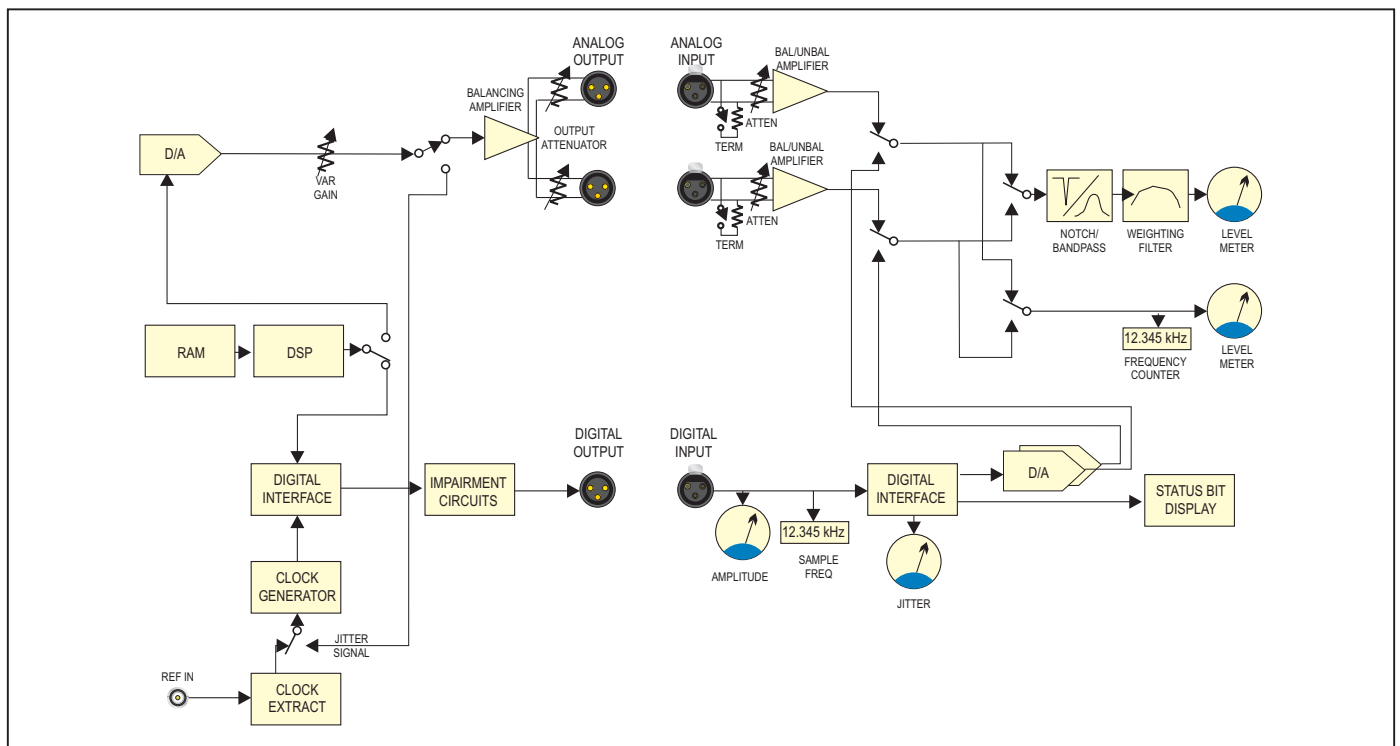

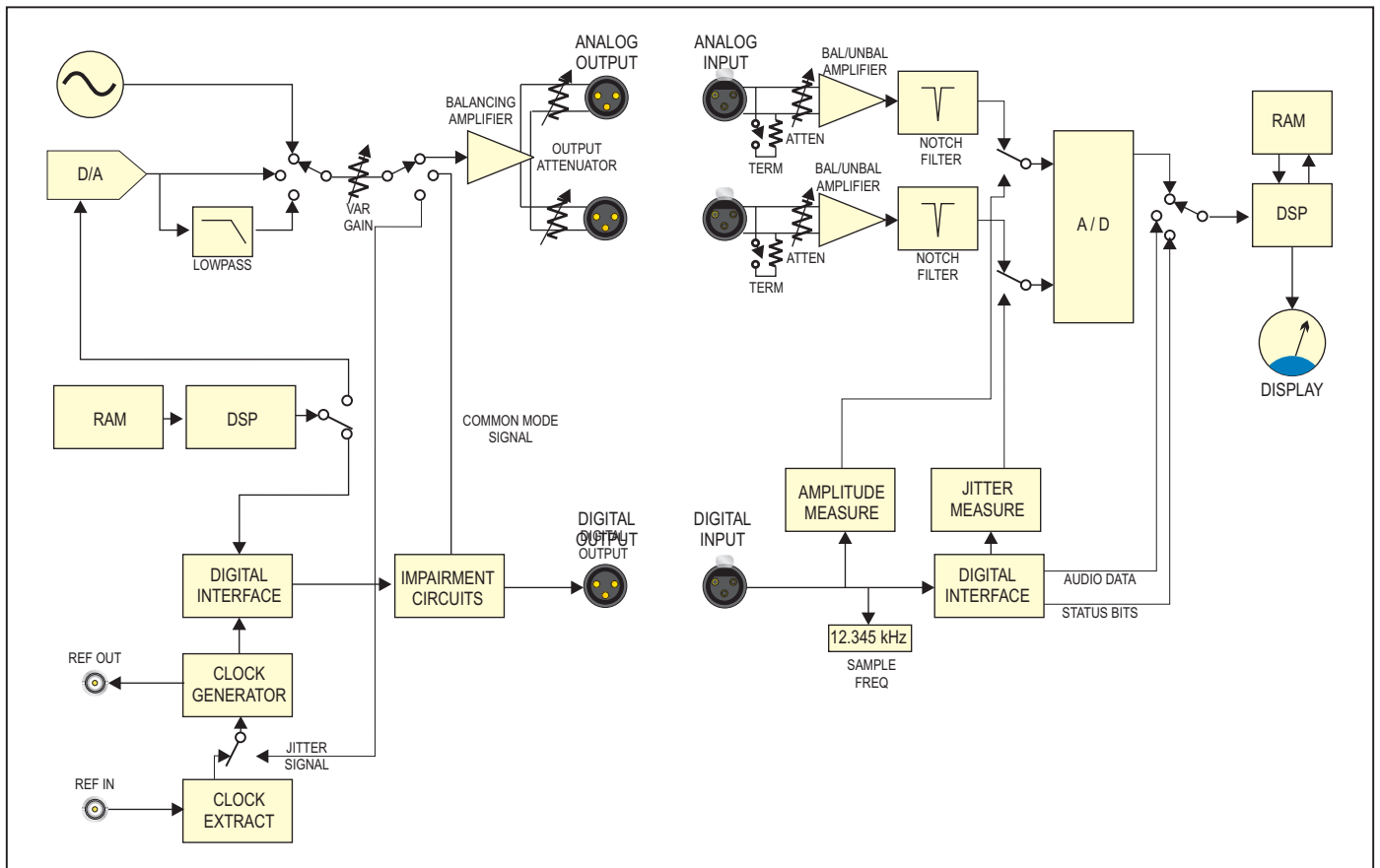**Fig. 2.** Simple Mixed Signal Audio Measurement System.

Fig. 3. Typical Mixed Signal Audio Measurement system.

limitations include frequency response irregularities which *exceed* 0.01 dB and distortion residuals which rarely reach 100 dB THD+N. Consequently, several of the available instruments which use this approach add a true analog signal generator for high performance applications. They also add an analog notch filter in front of the A/D converter for high performance analysis. As we will see later, this negates much of the cost and complexity advantages of the all-digital approach, while retaining most of its problems.

These evolved mixed signal architectures do not qualify as Dual Domain because neither signal generation nor analysis can be done simultaneously in both domains. Simultaneity of signal generation in the analog and digital domains is a critical issue for many types of testing, especially involving converter and interface performance. In many ways the need to simultaneously jitter the active digital audio signal, as well as drive an analog signal, creates a third domain. The mixed signal architecture shown is in-

capable of making interface jitter susceptibility measurements on A/D converters or D/A converters. It cannot generate digital and analog signals simultaneously, nor can it generate a digital signal simultaneous with the jitter embedded on its clock or simultaneous with the common mode interface signal. This prevents testing AES/EBU interface receiver operation under worst case conditions. The Dual Domain approach does allow any cross domain testing without compromise since all signals are simultaneously available, enabling complete characterization of mixedsignal devices under test.

Signal Generation

Audio testing generally uses sinewaves, squarewaves, random noise, and combinations of those signals. The dual domain approach described earlier uses multiple oscillators or waveform generators in the analog domain to optimize performance. Digital to analog converter based generation is used when particular waveform generation is not

easily accomplished by analog means. The D/A converters are used for multitone waveforms, shaped bursts, sines with interchannel phase shift (useful for testing surround sound decoders), etc. With the exception of multitone signals, these waveforms tend to have lower nonlinearity requirements than the other waveforms.

Testing state-of-the-art A/D converters to their performance limit requires a dedicated analog oscillator to achieve adequate THD+N. Several manufacturers have added tunable or switchable lowpass filters to d/a based generators in an attempt to achieve analog oscillator harmonic distortion performance. These have met with varying degrees of success. The trade-off between sharpness of filtering (and the corresponding distortion reduction) and flatness is difficult to balance. Sharper filters need a finer degree of tunability and have more response ripples, making the signal amplitude fluctuate with frequency. These filters also require more switchable elements, which introduce more noise and dis-

tortion. Therefore most high quality audio measurement equipment includes a provision for a dedicated analog oscillator which is used for THD+N testing.

Digital sinewaves may be generated in several different ways. The most common are table look-up and polynomial approximation. The table look-up method is fast but suffers from time resolution limitations driven by the limited length of the table. Commercial direct digital synthesis chips are implemented this way. Theoretical analyses (for example Tierney et al, 1971) have shown that the sine rom length should be at least 4 times the data width output from the rom. This makes the distortion introduced by quantization in the sample timing equal to the distortion introduced by quantization in the data word. Both of these errors may be converted to white noise through proper use of dither or error feedback techniques. The polynomial approximation technique yields sine accuracies dependent on the number of terms in the power series expansion used. Arbitrarily accurate signals may be obtained at the expense of computation time.

Finger (1986) has shown that proper signal generation in digital systems requires that the generated frequencies be relatively prime to the sample rate. If frequencies are used which are submultiples of the sample rate, the waveform will exercise only a few codes of the digital word. For example, generating 1 kHz in a 48 kHz sample rate system will require only 48 different data values. This may leave large portions of a converter untested. If frequencies are used which are prime to the sample rate then eventually every code in the data word will be used. Using 997 Hz instead of 1 kHz will result in all codes of a digital system (operating at standard sample rates) being exercised. This frequency makes a good "digital 1 kHz" since it is
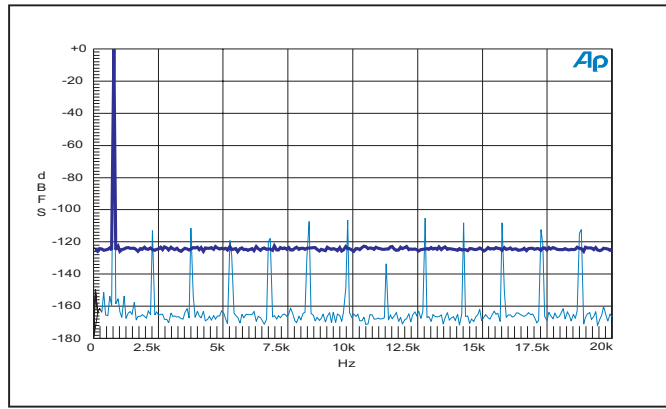

Fig. 4. Illustration of distortion reduction in return for higher noise floor with the addition of dither.
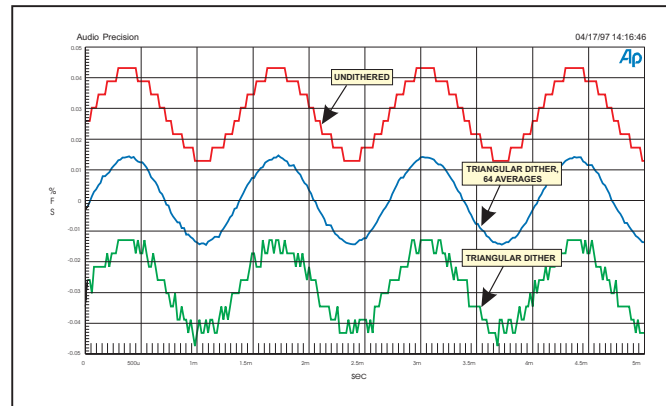

Fig. 5. Effectiveness of dither illustrated with 16 bit quantized signal.

also prime to the 44.1 kHz consumer standard sampling frequency.

Dither is one of the most misunderstood aspects of digital signal generation. When a signal is created in a finite word length system, quantization distortion will be introduced. Vanderkooy and Lipshitz (1987) have shown that the proper addition of dither to the signal before truncation to the final word width will randomize the distortion into noise. This comes at a 3dB (overall) increase in the background noise level. However, it allows the generation of signals below the system noise floor, and it frees large amplitude signals of any distortion products far below the system noise floor. This is illustrated in Fig. 4 which shows two FFTs of a 750 Hz tone overlaid on the same axes. The first is with 16 bit resolution, but no dither. The second is with correct amplitude triangular dither. Dither randomizes the distortion products into a smooth noise floor below the peak level of the distortion.

A smaller amplitude version of this same signal is shown in the time domain in Fig. 5. The upper trace shows the sinewave with no dither. The samples are limited to 16 bit resolution, which results in the familiar digital stair step waveshape. Note that each cycle repeats the same sample values. The lower trace shows the same sinewave with triangular dither. The sample values are different on each cycle, though they still are restricted to the 16 bit system resolution. The middle trace shows the average of 64 of the dithered sinewaves. The same sample values now average out to values between that limited by the 16 bit system. Dither randomizes the limited resolution of the 16 bit system into a smooth waveform with resolution much better than the sample resolution permits.

Complex Signal Generation

The multitone techniques discussed later require a means of generating multiple sinewaves simultaneously. For small numbers of sines this may be done with real-time computation of each sine in a dsp and subsequent summation. For larger numbers of tones rom or ram based waveform generation is normally used. For analog applications this is passed through a D/A converter. The rom size sets the waveform length before repeating, and therefore sets the minimum spacing of tones. The typical size in commercial equipment is 8192 or 16384 points which gives an approximately 6 or 3Hz spacing respectively at a 48 kHz sample rate.

Other waveforms such as those used for monotonicity testing of A/D converters may be created using table look-up techniques, or they may be computed in real time. For signals which do not need control of their parameters such as repetition rate or frequency, the look-up table approach has a speed advantage. It does how-

ever consume more memory or requires downloading from disk. The algorithmic approach offers complete control of waveform parameters, allowing signals such as shaped bursts or walking bit patterns to be adjusted to the use's needs. The available memory size and instrument architecture usually impacts this greatly. At least one commercial piece of audio test equipment derives all waveforms from disk files, though most use the algorithmic approach.

Most audio devices are multichannel. The usual approach to multichannel testing is to use a single generator with a single variable gain stage which is switched between two or more output channels. This can cope with simple crosstalk or separation measurements, but cannot handle more complex versions of these. For example: crosstalk measurements with multitone signals require different frequency tones in the two channels; measuring cross-channel intermodulation requires different frequency sinewaves in the two channels; record/reproduce measurements of tape recorder saturation characteristics requires the ability to make one channel sweep frequency while the other sweeps level so the frequency sweep may be used to identify the original channel's amplitude at each step. The common output amplifier splitting to multiple output connectors also means that there will be a common connection between channels that may affect measured separation. It also prevents adjusting the two channels of a stereo device for maximum output if the gains differ slightly.

Amplitude (Level) Measurement
The most basic measurement in audio is amplitude, or "level". There are many techniques for doing this, but the mathematically purest way is the root mean square value. This is representative of the energy in the signal and is computed by squaring the signal, averaging over some time period and taking the square root. The time period used is a parameter of the measurement, as is the type of averaging employed. The two approaches to averaging in common use are exponential and uniform.

Exponential averaging uses a first order running average (single pole in analog filter terms) which weights the most recent portion of the waveform more heavily than the earlier portion. This is the most commonly used technique for analog based implementations and has the benefit of making no assumptions about the waveform periodicity. It is merely necessary that the signal being measured have a period shorter than a fraction of the averaging time. The fraction sets the accuracy of the measurement, creating a minimum measurement frequency for a given accuracy. For complex signals, not only must each component meet the minimum frequency value, but their spacing in the frequency domain must also meet the minimum frequency requirement. The accuracy of exponential rms converters is better than the measurement repeatability or fluctuation due to ripple in the computed value. This fluctuation may be reduced without increasing the averaging time by post filtering the rms value. The optimum combination of averaging time and post filtering characteristics is well known (Analog Devices 1992).

Uniform averaging computes the rms average of the signal over a fixed time period where all portions of the signal have equal weight. Theoretical analyses of rms amplitude typically make the averaging time a fixed interval, which is then shown to directly affect the error in the measurement. Longer time intervals yield more accurate and repeatable measurements at the expense of measurement time.

This error may be eliminated for periodic signals if the averaging interval is made an integer multiple of the signal period. This technique is normally referred to as "synchronous rms conversion" since the averaging interval is synchronous to the signal. This has

been used in dsp based measurement systems for many years (Mahoney 1987) and has even been included in an analog based audio measurement system (Amber 1986). When measuring simple periodic signals which contain little noise this technique can yield repeatable measurements very quickly. Arbitrarily short measurement intervals may be used with no loss in accuracy, as long as the integer number of cycles constraint is obeyed. However most implementations will yield unstable or inaccurate results for noisy signals or inharmonic signals such as imd waveforms, since the integer averaging constraint is inherently violated. Hence, it must be used with care when measuring complex signals or when used for distortion or signal-to-noise ratio measurements. When this approach is applied to sinewave frequency response sweeps, the resulting speed can be quite impressive. However, because of errors in finding the zero crossings on digitized signals, the repeatability can leave something to be desired. Fig. 6 shows the results of 10 frequency response sweeps of a commercial system which uses this technique. Note that the error is approximately ±0.02 dB over most of the frequency range, rising to ±0.05 dB at high frequencies.

This error can be compensated for if corrections for the fractional portion of the sinewave cycle are computed. These corrections are dynamic, changing from cycle to cycle with the phase of the waveform relative to the sampling instants at both the begin-
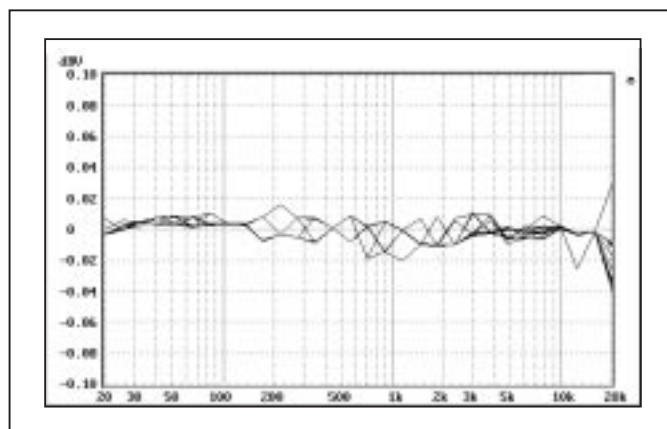

Fig. 6. Frequency response flatness variation due to errors in period computation.

ning and end of the zero crossing. The graph in Fig. 7 illustrates the flatness of a cycle based rms converter using these enhancements. Note the tenfold difference in graph scale compared to Fig. 6.

The simplest technique for amplitude measurement of analog signals, rectification and averaging, is extremely difficult for digital signals. The rectification process is nonlinear and creates harmonics of the signal which will alias based on the finite sample rate. For very low frequency signals this is not a problem, since the harmonic amplitudes decrease with increasing order and are adequately small by the time the folding frequency is reached. However, high frequency signals have enough energy in the harmonics that the average value obtained will depend on the phase alignment of the aliased components and the original signal. The result is beat products between these components which yield fluctuating readings.

Peak measurements have a similar problem with limited bandwidth. The peak value of the signal values is easy to determine in software, and several instruments supply this as an indicator of potential signal clipping. However, the peak value of the analog signal that the samples represent may well be different. This difference increases with signal frequency. When a digital signal is converted to analog (or when an analog signal is sampled) the sample values may not fall on the signal peaks. If the samples straddle a peak, the peak value will be higher, unless the signal is a square wave. This error is directly

proportional to the frequency of the highest frequency component in the spectrum, and to its proportion of the total signal energy. This problem may be reduced to any desired significance by interpolation of the waveform and peak determination on the higher sample rate version.

Quasi-peak amplitude measurements are a variant of the peak value measurement where the effect of an isolated peak is reduced. This technique was developed to assess the audibility of telephone switch gear noise in the days when telephone systems used relays and electromagnetically operated rotary switch devices. The clicks that these devices could introduce into an audio signal were more objectionable than their rms or average amplitude would imply. This technique spread from its origins in the telecom world to the professional audio world, at least in Europe, and has lasted long after the problem it was devised to characterize disappeared. This measurement is implemented with a full wave rectification and limited attack and decay time averaging, similar audio compressor implementations. The implementation techniques in the digital domain are similar.

Any measurement system which implements analog amplitude measurements with dsp techniques by digitizing the original analog signal must consider the effects of converter response ripple. This can be substantial, exceeding 0.01 dB for some commercial devices. The effect of these ripples adds directly to the response error in the rms algorithm itself and may be a significant portion of the instrument flatness specification.

FFT Measurements

With the advent of inexpensive digital signal processing devices, the FFT has become a commonplace audio measurement tool. To obtain accurate measurements, it is essential to understand its operation, capabilities and

limitations. The FFT is merely a faster method of computing the discrete Fourier transform. The discrete Fourier transform determines the amplitude of a particular frequency sinewave or cosinewave in a signal. The algorithm multiplies the signal, point by point, with a unit amplitude sinewave. The result is averaged over an integer number of sinewave cycles. If the sinewave is not present in the signal being analyzed, the average will tend to zero. This process is repeated for a unit amplitude cosinewave, since the sine and cosine are orthogonal. Again, if the cosinewave is not present, the average will tend to zero. If there is some of the sine or cosine wave present, the average will be proportional to the amplitude of the component in the signal. The relative proportion of sine and cosine components at a given frequency, along with their polarities, represents the phase.

If this process is repeated for each hypothetical sinewave and cosinewave whose period is an integer submultiple of the waveform length, several redundancies will occur in the computation. By eliminating these redundancies the number of operations may be reduced. The resulting simplified process is called the FFT.

Since all hypothetical sine and cosine frequencies in the FFT are multiples of the reciprocal of the waveform length, the analysis is inherently equal resolution in the frequency domain. This analysis also presupposes that the signal components are at exact multiples of the reciprocal of the waveform length; serious problems occur when this is violated. Stated differently, the FFT assumes that the waveform being analyzed is periodic with a period equal to the length of the data record being analyzed (Fig. 8). Consequently, if the beginning and end of the record do not meet with the same value and slope when looped back on themselves the discontinuity will result in artifacts in the spectrum. The usual way to deal with this is to "window" the data and drive its value to zero at the end points. This turns the waveform into a "shaped burst", whose spectrum is the convolution of the window spectrum and the signal spectrum.
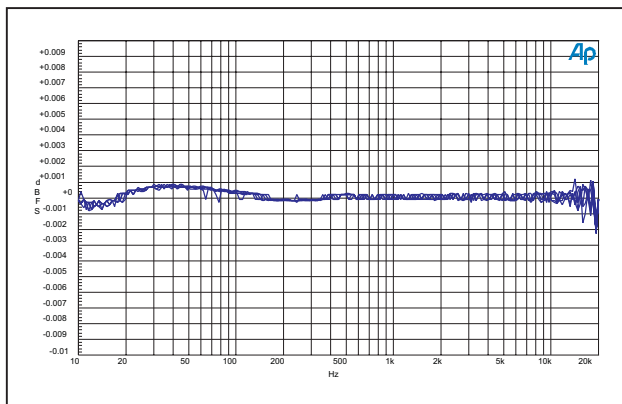
There are approximately as many

Fig. 7. Period-based rms measurement flatness variation with a fractional sample compensation.

different window functions as there are papers about windowing. Everyone has designed their own, probably so they can put their name on it and get a piece of fame. From a practical viewpoint, very few windows are necessary for audio measurements. To understand the advantages, or lack thereof, of the various windows we will start with the performance metrics of windows. Most important are the -3 dB

Fig. 8. Discontinuity in analysis record resulting from asynchronous signal acquisition.

bandwidth (in bins), the worst case amplitude accuracy or scalloping loss, the highest sidelobe amplitude and the sidelobe roll-off. Fig. 9 illustrates these parameters for several representative windows. The -3 dB bandwidth is an indicator of the ability to resolve two closely spaced tones which are nearly equal in amplitude. The scalloping loss is the maximum variation in measured amplitude for a signal of unknown frequency. This indicates the worst case measurement error when displaying isolated tones which may be asynchronous to the sample rate. The highest sidelobe amplitude is indicative of the ability to resolve a small amplitude tone close to a large amplitude tone. The sidelobe roll-off indicates the efficiency of the window at large distances from the main tone.

The simplest window in common use is the Hann window, named after its inventor, Austrian astronomer Julius von Hann (often incorrectly called the Hanning window because of confusion with the Hamming window, named after Richard Hamming). The Hann window does allow good differentiation of closely spaced equal amplitude tones and, because it is a raised cosine wave, is very easy to compute. The Blackman-Harris 4-term 94 dB window (one of the many Blackman-Harris windows) offers a good balance of attenuation (94 dB to the highest sidelobe) and moderate -3 dB bandwidth. The flat-top window offers negligible amplitude error for asynchronous signals and so allows accurate measurements of discrete tones. The Dolph-Chebyshev windows keep all sidelobes an equal distance down from the peak and so offer the optimum resolution of small amplitude tones, but at the expense of somewhat larger -3 dB bandwidth. The Dolph-Chebyshev windows are a family of windows allowing specification of the desired sidelobe level and consequently the worst-case spurious peak in the spectrum (neglecting FFT distortion products, which are discussed below). The Audio Precision 170 dB version specified here as "Equiripple" was chosen to produce spurs comparable in magnitude to the noise floor of 24-bit digital systems.

An approach developed by this author called frequency shifting results in large improvements over the windowing approaches. The FFT assumes that any signal it analyzes has a period that is an integer fraction of the acquisition time. If the record does not contain an integer number of periods, a window must be used to taper the ends of the acquired waveform to zero. The window will smear the sine in the frequency domain, reducing the ability to resolve sidebands on the signal and consequently the ability to resolve low frequency jitter sidebands, noise sidebands or the ability to measure harmonics of low frequency tones. If, after acquisition, the sample rate of the waveform is changed to make an integer number of signal periods fit in the acquired record, there will not be any need for a window. This allows the am-
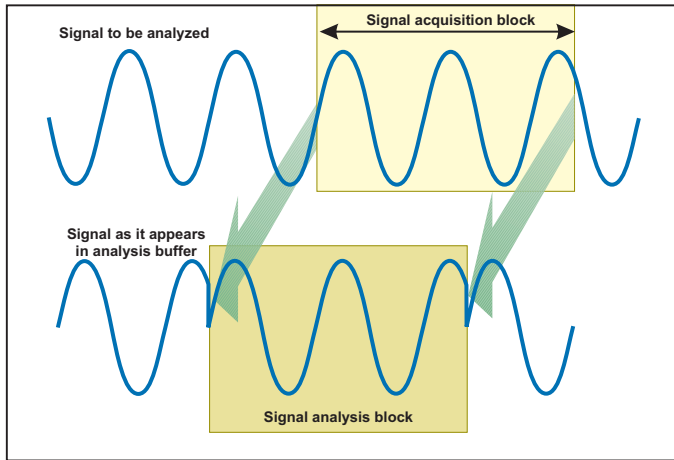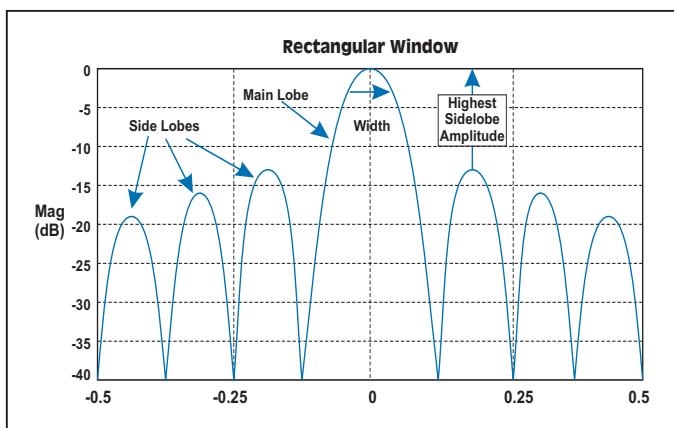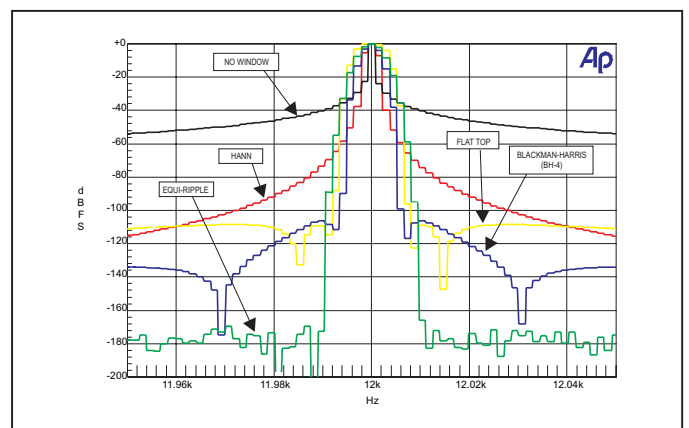

Fig. 9. Illustration of Window Parameters.

Fig. 10. Effective response of various windows.

plitude of neighboring bins to be resolved to the full dynamic range of the FFT and component amplitudes to be correctly measured without scalloping loss. This allows devices such as A/D converters to be tested with signals which are not a submultiple of the sample rate. This maximizes the number of codes tested and maximizes the creation of spurious tones.

Fig. 11 illustrates the operation of this sample rate adjustment for an 18 Hz sinewave. The three traces are the unwindowed version, the equiripple windowed version and the frequency shifted version. Each has been averaged 64 times. Note the complete absence of window-induced spreading and the 150 dB dynamic range ob-

Fig. 11. Selectivity improvement with frequency shifting over windowing.

tained. This reduction in window spreading also results in a substantial improvement in frequency resolution. The typical window width of between 5 and 11 bins has been reduced to one bin, giving a corresponding 5 to 11 times improvement in resolution. This is achieved with no increase in acquisition time or, more importantly, acquired record length. Since the record length is not increased, the ability to resolve semi-stationary signals such as tone bursts is maintained.

When making measurements on semi-stationary signals such as tone bursts or transients it is essential to correlate the time and frequency domains. The exact segment in the time domain which will be transformed must be selectable to allow windowing out unwanted features while retaining wanted features of the waveform. Once the segment boundaries are established, the time domain segment is transformed into the frequency domain. An example of this measurement is the distortion introduced by a compressor on a tone burst during its attack, sustain and release operations. By performing a short FFT every few milliseconds through the acquired record the distortion products may be studied.

Frequency Measurement

There are two basic approaches to measuring frequency: zero crossing based schemes and spectrum peak localization based schemes. Zero crossing counting has been used for decades on analog signals in stand-alone frequency counters. In a simple sense, the number of zero crossings occurring during a fixed amount of time may be counted and reported as the signal frequency. In practice, this approach is never used at audio frequencies because a low frequency signal, such as 20 Hz, would only be counted to a 1Hz (or 5%) resolution with a 1 second measurement time. Instead, the time interval between zero crossings is measured which yields the period. This is reciprocated to get frequency. If the time between successive zero crossings is measured, the measurement rate will be directly proportional to the signal frequency. This leads to excessively fast readings at high frequencies which tend to be sensitive to interfering noise. By measuring the time between zero crossings several cycles apart, this noise may be reduced by averaging. Hence, practical equipment measures the number of zero crossings which occur in a time interval which is approximately constant, independent of signal frequency. The desired reading

rate and corresponding data averaging are used to determine this time interval. At low frequencies, the measurement is typically made over one cycle of signal while at high frequencies, many cycles are used.

Spectrum peak based techniques have been around since spectrum analyzers were invented. The concept is simple enough: if you know the shape of the filter used to make the spectrum measurement, you can interpolate the exact location of the spectrum peak and therefore determine the frequency. This assumes two things: that there is only one frequency component within the filter bandwidth, and that the filter shape does not change as a function of frequency or signal phase. These limitations are not severe, and this technique offers a significant noise bandwidth advantage over the zero crossing based approaches. If a sinewave is measured in the presence of wideband interfering noise, only the noise which falls within the filter bandwidth will affect the measurement. This technique is especially well suited to FFT based implementation since the window functions normally used provide a predictable window shape. Rosenfeld (1986) describes a window optimized for the task of frequency measurement, though any window shape may be used if appropriate modifications to the software are made. The proprietary scheme developed by Audio Precision for its *FAST*TEST[2] multitone measurement software allows the use of any window the customer chooses. The performance tradeoff simply becomes one of noise bandwidth and selectivity between adjacent tones.

Measurement Dynamic Range

Dynamic range is in itself an interesting issue for both audio measurement equipment and audio processing equipment. The bottom line is usually bits, how many are used and how are they used. The issue of how is not usually so obvious. Data word widths in professional audio range from 16 to 24 bits. However, processing algorithms consume bits by virtue of the truncation or
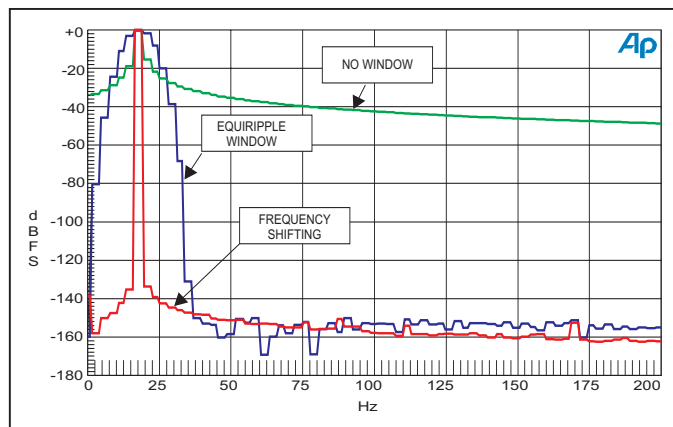
---

[2]*FAST*TEST is a trademark of Audio Precision, Inc.

rounding error introduced with each multiply operation. Consider the effect of multiplying two 24-bit signed words. The result is 47 bits wide (one of the sign bits is redundant). When this is converted to 24 bits again error is introduced in the lsb of the resulting data. When several operations are cascaded this error can grow to unacceptable levels (Cabot 1990). Indeed, for measurement equipment which is intended to test 24-bit systems, any introduction of error in the 24th bit is unacceptable.

The two most common operations in audio measurement are filtering and FFTs. It can be shown that conventional digital filters introduce a noise-like error due to truncation operations which is proportional to the ratio of the sample rate and the filter cutoff or center frequency. For a 20 Hz filter operating at a 48 kHz rate this gives a noise gain of 2400, approximately 67 dB or 11 bits. For a 24-bit processor this filter would give 13 bit noise and distortion performance. There are alternative filter structures which reduce this error, but none can eliminate it. Similarly, it can be shown that the FFT algorithm introduces approximately 3 dB (or one half bit) of noise increase for each pass of the transform. A 16 k transform requires 14 passes ($16k = 2^{14}$), giving a 42 dB noise increase. The result is that a 24-bit 16 k transform gives a 17-bit result. Special techniques can improve
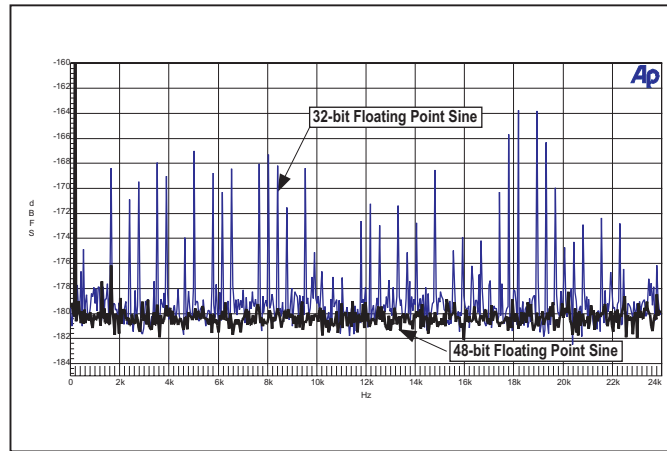

Fig. 12. Comparision of harmonic distortion of 32-bit floating point and 48-bit fixed point sinewaves, quantized to 24-bits.

this Fig. by a few bits at most. Fixed point 48-bit processing allows a theoretical 288 dB dynamic range and resolution, providing considerable margin for loss in the algorithms. Noise problems become even more pronounced in the new 192 kHz sample rate systems.

Floating-point processing is usually touted as being a panacea since the dynamic range of 32-bit floating-point numbers is many hundreds of dB. Most floating point formats consist of a 24-bit mantissa and an 8-bit exponent. For major portions of a waveform, even those as simple as a sine, the mantissa resolution actually sets the performance of the processing. This is because the exponent is zero for 240 degrees of the cycle. The FFT in Fig. 12 shows two 187.5 Hz sinewaves (at 48 kHz sample rate). One was generated by a commercial audio measurement system which uses 32-bit float-

ing-point processing, while the other was generated with 48-bit fixed point computations in a System Two Cascade.

Measurement Averaging

Many audio measurements are made on noisy signals. It helps to be able to average several measurements together to reduce the effects of noise. The mathematically correct way to do this is either with power law or with vector operations. Each has its place. Power law averaging takes successive data points, squares them, and accumulatess them into a running sum. This reduces the measurement variability, since the variance of the final measurement result is the variance of the original measurements divided by the square root of the number of data points averaged. Fig. 14 illustrates this improvement for a typical distortion and noise spectrum of an A/D converter. The upper trace is a single FFT of the A/D converter under test. The trace immediately below it is a power law average of 64 FFTs. Note that the variability is drastically reduced. The trace is smooth and general trends are clearer.

Power law averaging is inherently phase insensitive. Vector averaging considers both a magnitude and phase of each data point. Instead of operating on the FFT results, successive acquisitions are averaged before transforming. This is equivalent to vectorially averaging the FFT results
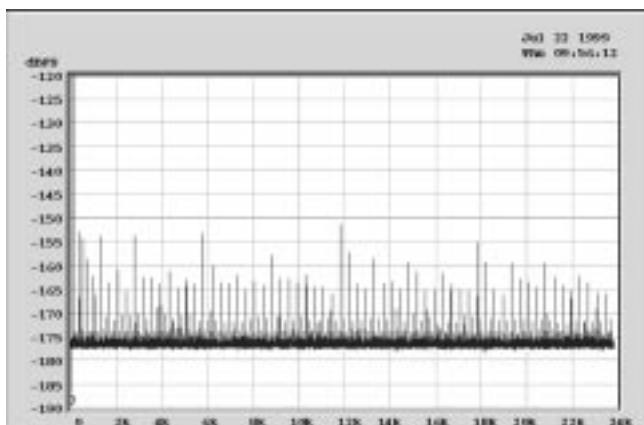

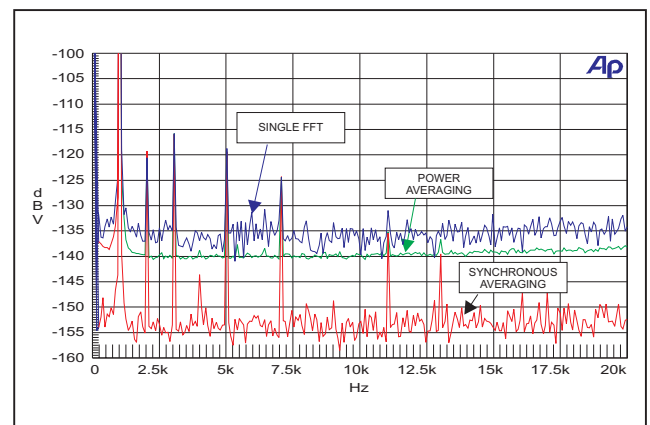**Fig. 13.** Residual distortion of a 32-bit floating point FFT.

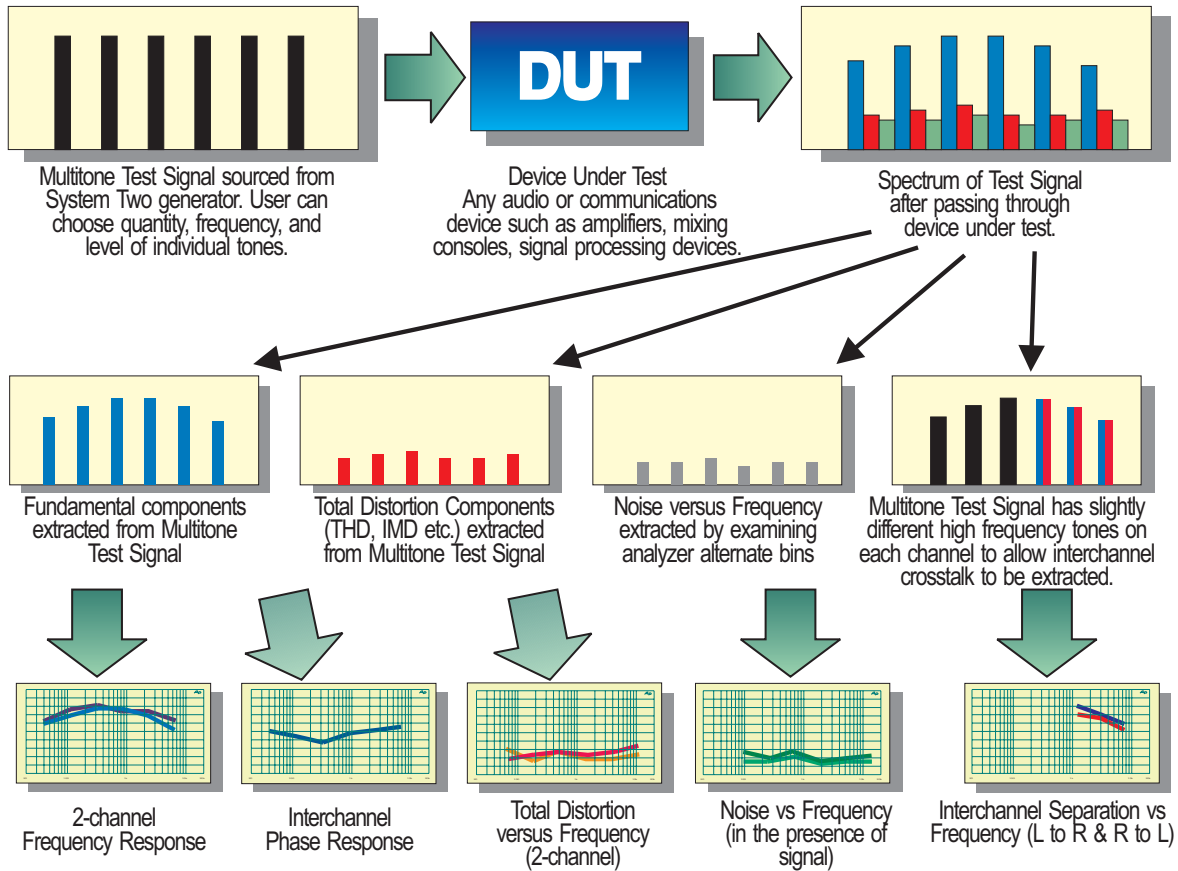**Fig . 14.** A/D converter noise and distortion improvement with averaging.

Fig. 15. *FAST*TEST multitone measurement concept.

(considering both magnitude and phase of the data values). If two values are equal magnitude but opposite in phase they average to zero. Power law averaging would give the same magnitude as each of the two original magnitudes. The result is that vector or "synchronous" averaging reinforces coherent signals and reduces the variability of their amplitude and phase, just as power law averaging reduces variability of their magnitude. However, synchronous averaging reduces the amplitude of noncoherent signals but not their variability. Consequently the fundamental and its harmonics are more easily visible because the noise floor moves down. This is shown in Fig. 14 as the lowest trace. Note that the variability of the background noise is the same as the unaveraged case but its amplitude is 18 dB lower (8 times or the square root of 64).

Multitone Measurements

Multitone measurements allow very fast measurement of linear errors such as amplitude and phase response vs.. fre-

quency, interchannel crosstalk and noise, as well as nonlinear effects. Originally developed to allow very fast measurements of broadcast links, the technique has also found wide application in production test, because of its high speed, and in tape recorder testing, since it does not need synchronization between source and receiver. *FAST*TEST is the trade name for the implementation and enhancements of the basic multitone concept developed and described by Cabot (1991). Classic multitone measurements are detailed by Mahoney (1987).

The operation of the *FAST*TEST measurement technique is illustrated in Fig. 15. The excitation is the sum of several sinewaves whose frequencies are typically distributed logarithmically across the audio range. The device under test output spectrum is measured and the amplitudes and phases of the components at the original stimulus frequencies provide the linear amplitude and phase vs. frequency response. Additional measurements such as crosstalk and noise may

easily be obtained from the measurement by appropriate choice of signal and analysis frequencies.

The number of individual sinewaves in the *FAST*TEST signal, their frequencies and the individual amplitudes may be set by the user. The only restriction is that they be a multiple of the basic FFT analysis length. In the typical configuration with an 8192 point waveform at a 48 kHz sample rate this results in 4096 bins of 5.96Hz frequency resolution spanning the dc to 24 kHz range. This flexibility may be used to adjust the test signal spectrum to simulate the typical frequency distribution of program material. The phases of the sinewaves comprising the test signal may also be adjusted to control the crest factor. For instance, if all tones are set to a cosine phase relationship the peaks will add coherently, producing a maximum amplitude equal to the sum of the individual sinewave peak amplitudes. The test signal rms amplitude will be the power sum of each sinewave rms amplitude, and the resulting crest factor will be
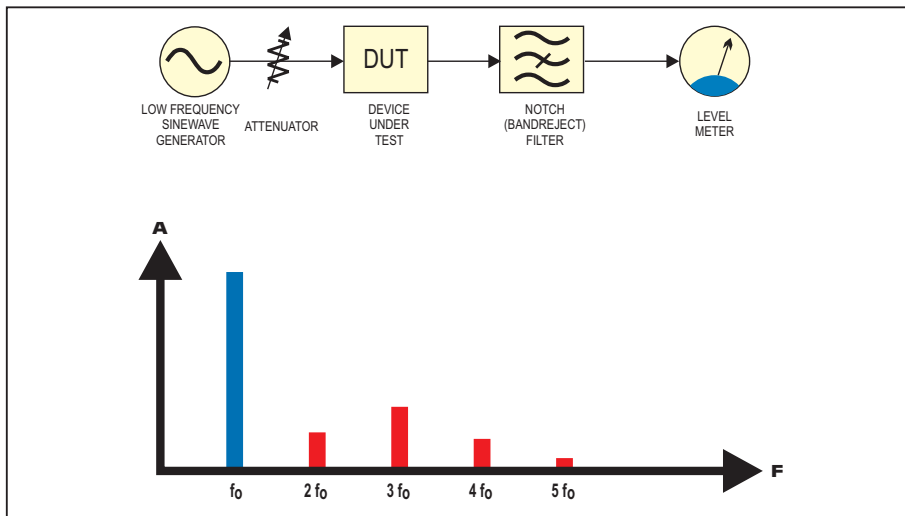
Fig. 16. Total Harmonic Distortion (THD).

proportional to the square root of the number of tones. This is the maximum possible for a given signal spectrum. Alternatively, the phases may be adjusted to minimize the crest factor. This will typically result in a crest factor which increases as the fourth root of the number of tones. Typical crest factors for $1/3^{rd}$ octave-spaced tone signals are around 3.5, approximately 2.5 times that of a single sinewave.

Harmonic Distortion

Harmonic distortion, illustrated in Fig. 16 is probably the oldest and most universally accepted method of measuring linearity (Cabot 1992). This technique excites the device under test with a single high purity sine wave. The output signal from the device will have its waveshape changed if the input encounters any nonlinearities. A spectral analysis of the signal will show that in addition to the original input sinewave, there will be components at harmonics (integer multiples) of the fundamental (input) frequency. Total harmonic distortion (THD) is then defined as the ratio of the RMS voltage of the harmonics to that of the fundamental. This may be accomplished by using a spectrum analyzer to obtain the level of each harmonic and performing an RMS summation. This level is then divided by the fundamental level, and cited as the total harmonic distortion (usually expressed in percent). Alternatively a distortion analyzer may be used which removes the fundamental component and measures the remainder. The remainder will

contain both harmonics and random noise. At low levels of harmonic distortion, this noise will begin to make a contribution to the measured distortion. Therefore measurements with this system are called THD+N to emphasize the noise contribution.

Low frequency harmonic distortion measurements suffer a serious resolution limitation when measured with FFT techniques. Measuring a 20Hz fundamental requires the ability to separate a 40 Hz second harmonic with a dynamic range equal to the desired residual THD. Since the FFT yields a linear frequency scale with equal bin sizes, an 8192 point FFT gives approximately 6 Hz bins at a 48 kHz sample rate. To resolve a 100 dB residual 2nd harmonic requires a window attenuation of 100 dB only 3 bins away from the fundamental. This is not achievable. The FFT length may be increased to reduce the bin width, but this will lengthen the measurement time.

A sine wave test signal has the distinct advantage of simplicity, both in instrumentation and in use. This simplicity has an additional benefit in ease of interpretation. If a notch type distortion analyzer (with an adequately narrow notch) is used, the shape of the residual signal is indicative of the shape of the nonlinearity. Displaying the residual components on the vertical axis of an oscilloscope and the input signal on the horizontal gives a plot of the transfer characteristic deviation from a best fit straight line. Examination of

the distortion components in real time on an oscilloscope will immediately reveal such things as oscillation on the peaks of a signal, crossover distortion, clipping, etc. This is an extremely valuable tool in design and development of audio circuits, and one which no other distortion test can fully match. Viewing the residual components in the frequency domain also gives much information about the distortion mechanism inside the device under test. This usually requires experience with the test on many circuits of known behavior before the insight can be obtained.

Another advantage of the classic filter based approach to harmonic distortion measurement is the opportunity for listening to the distortion products. This will often yield significant insights into the source of the distortion or its relative audible quality.

The frequency of the fundamental component is a variable in harmonic distortion testing. This often proves to be of great value in investigating the nature of a distortion mechanism. Increases in distortion at lower frequencies are indicative of fuse distortion or thermal effects in the semiconductors. Beating of the distortion reading with multiples of the line frequency is a sign of power supply ripple problems, while beating with 15.625 kHz, 19kHz or 38kHz is related to subcarrier problems in television or FM receivers.

The subject of high frequency harmonic distortion measurements brings up the main problem with the harmonic distortion measurement method. Since the components being measured are harmonics of the input frequency, they may fall outside the passband of the device under test. An audio device with a cutoff frequency of 22kHz will only allow measurement of the third harmonic of a 7kHz input. THD measurements on a 20kHz input can be misleading because some of the distortion components are filtered out by the recorder. Intermodulation measurements do not have this problem and this is the most often cited reason for their use. THD measurements may also be disturbed by wow and flutter in the device under test, depending upon the type of analysis used.

SMPTE Intermodulation

Intermodulation measurements using the SMPTE method (originally standardized by the Society of Motion Picture and Television Engineers, hence its name) have been around since the 1930s. The test signal consists of a low frequency (usually 60Hz) and a high frequency (usually 7kHz) tone, summed together in a 4 to 1 amplitude ratio as shown in Fig. 17. Other amplitude ratios and frequencies are used occasionally. This signal is applied to the device under test, and the output signal is examined for modulation of the upper frequency by the low frequency tone. As with harmonic distortion measurement, this may be done with a spectrum analyzer or with a dedicated distortion analyzer. The modulation components of the upper signal appear as sidebands spaced at multiples of the lower frequency tone. The amplitudes of the sidebands are added in pairs, root square summed, and expressed as a percentage of the upper frequency level. Care must be taken to prevent sidebands introduced by frequency modulation of the upper tone from affecting the measurement. For example, loudspeakers may introduce Doppler distortion if both tones are reproduced by the same driver. This would be indistinguishable from intermodulation if only the sideband powers were considered. If the measurements are made with a spectrum analyzer which is phase sensitive, the AM and FM components may be separated by combining components symmetrically disposed about the high frequency tone.

A dedicated distortion analyzer for SMPTE testing is quite straightforward. The signal to be analyzed is high pass filtered to remove the low frequency tone. The sidebands are demodulated using an amplitude modulation detector. The result is low pass filtered to remove the residual carrier components. Since this low pass filter restricts the measurement bandwidth, noise has little effect on SMPTE measurements. The analyzer is very tolerant of harmonics of the two input signals, allowing fairly simple oscillators to be used. It is important that none of the harmonics of the low frequency oscillator occur near the upper frequency
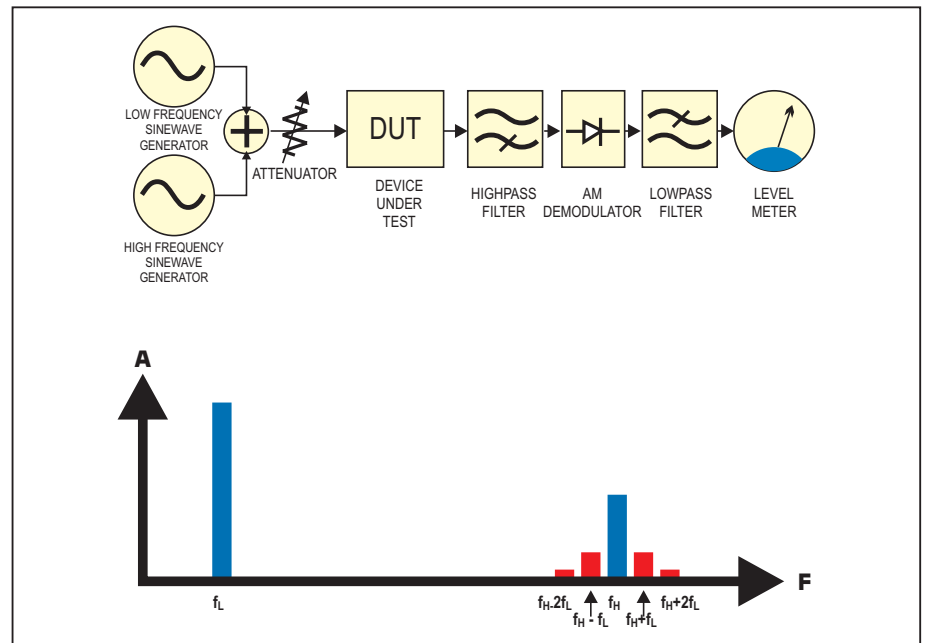

Fig. 17. SMPTE Intermodulation Distortion.

tone, since the analyzer will view these as distortion. After the first stage of high pass filtering in the analyzer there is little low frequency information left to create intermodulation in the analyzer itself. This simplifies design of the remaining circuitry.

A major advantage of the demodulator approach to SMPTE distortion measurement is the opportunity for listening to the distortion products. As with listening to harmonic distortion, it often yields insights into the source of the distortion or its relative audible quality.

Considering the SMPTE test in the time domain helps understand its operation. The small amplitude high frequency component is moved through the input range of the device under test by the low frequency tone. The amplitude of the high frequency tone will be changed by the incremental gain of the device at each point, creating an amplitude modulation if the gain changes. This test is therefore particularly sensitive to such things as crossover distortion and clipping. High order nonlinearities create bumps in the transfer characteristic which produce large amounts of SMPTE IM.

SMPTE testing is also good for exciting low frequency thermal distortion. The low frequency signal excursions excite thermal effects, changing the

gain of the device and introducing modulation distortion. Another excellent application is the testing of output LC stabilization networks in power amplifiers. Low frequency signals may saturate the output inductor, causing it to become nonlinear. Since the frequency is low, very little voltage is dropped across the inductor, and there would be little low frequency harmonic distortion. The high frequency tone current creates a larger voltage drop across the inductor (because of the rising impedance with frequency). When the low frequency tone creates a nonlinear inductance, the high frequency tone becomes distorted. A third common use is testing for cold solder joints or bad switch contacts.

One advantage in sensitivity that the SMPTE test has in detecting low frequency distortion mechanisms is that the distortion components occur at a high frequency. In most audio circuits there is less loop gain at high frequencies and so the distortion will not be reduced as effectively by feedback. Another advantage of the SMPTE test is its relatively low noise bandwidth, allowing low residual measurements.

The inherent insensitivity to wow and flutter has fostered the widespread use of the SMPTE test in applications which involve recording the signal. Much use was made of SMPTE IM in
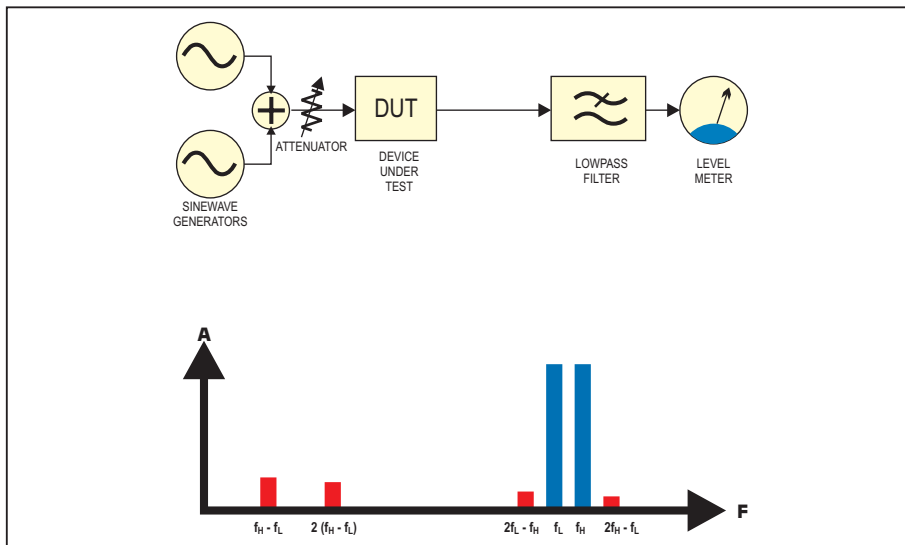
Fig. 18. CCIF Intermodulation Distortion. (Also called DFD or Difference Frequency Distortion).

the disc recording and film industries. When applied to discs, the frequencies used are usually 400Hz and 4kHz. This form of IM testing is quite sensitive to excessive polishing of the disc surface, even though harmonic distortion was not. It also has found wide application in telecom and mobile radio areas because of its ability to test extremes of the audio band while keeping the distortion products within the band.

CCIF (DFD) Intermodulation

The CCIF or DFD (Difference Frequency Distortion) intermodulation distortion test differs from the SMPTE test in that a pair of signals close in frequency are applied to the device under test. The nonlinearity in the device causes intermodulation products between the two signals which are subsequently measured as shown in Fig. 10c. For the typical case of input signals at 14kHz and 15kHz, the intermodulation components will be at 1kHz, 2kHz, 3kHz, etc. and 13kHz, 16kHz, 12kHz, 17kHz, 11kHz, 18kHz, etc. Even-order or asymmetrical distortions produce the low "difference frequency" components while the odd-order or symmetrical nonlinearities produce the components near the input signals. The most common application of this test only measures the even order difference frequency components, since this may be achieved with only a multi-pole low pass filter.

This technique has the advantage that signal and distortion components can almost always be arranged to be in the passband of a nonlinear system. At low frequencies, the required spacing becomes proportionally smaller, requiring a higher resolution in the spectrum analysis. At such frequencies a THD measurement may be more convenient.

Recent versions of IEC standards for DFD have specified the results in spectral terms. Previous versions of the IEC standard specified the reference level computation differently. This introduces a 6 dB difference between the two versions of the standard for DFD measurements. This re-definition also conflicts with accepted practice for difference tone distortion measurements and with usage of the technique in other IEC standards.

Level Linearity

One method of measuring the quantization characteristics of converters is to measure their amplitude linear-

ity. If a signal, for example at -20 dBFS, is applied to an audio device the output will depend on the gain. If for this example the output is also -20 dBFS the device gain is 0 dB. If the input is changed to -40 dBFS the output should follow. In other words, the gain should be constant with signal level. For typical analog equipment except compressor/limiters and expanders this will be true. At low levels crossover distortion will make this not the case. It is common for A/D and D/A converters to suffer from a form of crossover distortion due to inaccurate bit matching. To measure this, we apply a sinewave to the input and measure the amplitude of the output with a meter. The input is changed by known amounts and the output level is measured at each step. To enable measurements below interfering noise in the system, a bandpass filter tuned to the signal frequency is placed ahead of the measurement meter. The measurement block diagram is shown in Fig. 19. Frequencies used for this testing are normally chosen to avoid integer submultiples of the sample rate, for example 997 Hz in digital systems with standard sample rates. This maximizes the number of states of the converter exercised in the test.

Graphing the device gain vs. input level gives a level linearity plot. For an ideal converter this would be a horizontal line whose value is the device gain. In practice this gain will vary as the level is reduced. Examples of typical device measurements are shown in Fig.s 21 a, c, and e. The level linearity plot is a standard fixture of most consumer digital audio equipment test reports.

Noise Modulation

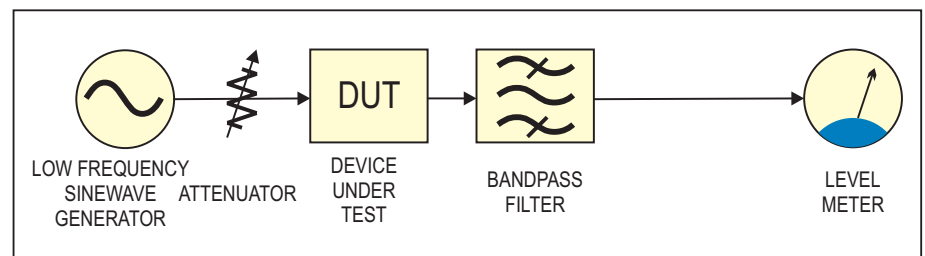Fielder developed a technique for characterizing digital conversion systems

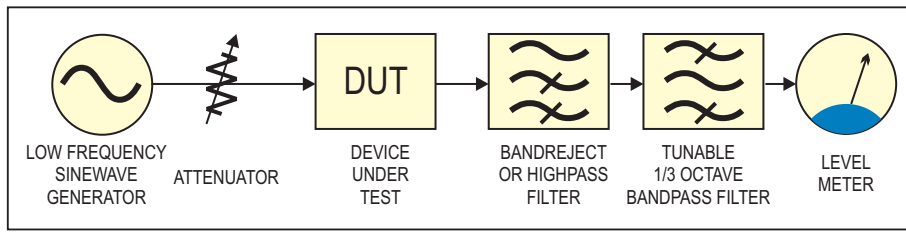**Fig. 19.** Level linearity measurement block diagram.

Fig. 20. Noise modulation measurement block diagram.

called noise modulation which has been shown to correlate well with perceived quality. It measures the variation in the 1/3rd octave band noise floor with variations in signal level. If the noise floor varies by more than 1 dB in any band, the converter will likely have audible modulation of the noise floor with changes in signal level. This will manifest itself as an audible shift in the level or tonal balance of the background noise as music such as piano notes decays into the noise floor. Fig. 20 illustrates the test setup for this measurement. The device under test is stimulated with a low frequency sinewave. This is removed with either a notch filter or high pass filter and the spectrum of the remaining signal is measured in 1/3rd octave steps with a 1/3rd octave bandpass filter. The signal level is changed and the measurement is repeated. The amplitude is typically changed in 5 dB steps beginning 40 dB below full scale. The deviation in the noise spectrum is the parameter of interest, so the peak variation between traces in each band is the noise modulation. Fig.s 21 a-f compare the level linearity and noise modulation for three different channels of a digital multitrack tape recorder. These measurements and the theory behind them are detailed in Cabot (1991).

If the amplitude sweep of a noise modulation measurement is sped up, it offers the opportunity for listening to the background noise and the resulting shifts of timbre. Being essentially a test of shifts in noise spectrum balance, this ability to listen may offer insights into relative audible quality.

FASTTEST Total Distortion

Most of the distortion products in a multitone signal will fall between the original stimulus frequencies and will include both harmonics and inter-

modulation products of these frequencies. The FASTTEST total distortion measure (Cabot 1991) is a summation over frequency of the powers in the distortion products. If the summation is done in segments, such as those represented by the space between the original tones, the result may be displayed as a distortion vs. frequency plot. This graph is not the usual sensitivity of the distortion measure to signal frequency but represents the distribution of distortion products with frequency. This distinction is important since it is not an equivalent display. If the summation is done over the entire frequency band a single value will be obtained. As with other distortion measures, this value may be graphed as a function of stimulus amplitude or device output amplitude.

The average slew rate of a FASTTEST signal will be dependent on the distribution of energy with frequency. Including more tones at high frequencies will increase the average slew rate, making the test more sensitive to frequency-dependent nonlinearities. Including more tones at low frequencies will make the test more sensitive to inverse frequency dependent nonlinearities.

If the sinewave frequencies are chosen to be in the FFT bin centers, the transform results contain no spillage into neighboring bins. This maximizes the dynamic range and frequency resolution by avoiding the use of windowing. However, if the generator and analyzer are not driven by the same clock, it may be difficult to place the sinewave frequencies in the bin centers because of differences between the generator and analyzer electronics. Similarly, there can be a problem if the device under test stores and replays the signal, as is the case with a tape recorder. The record/playback

speed error can shift the frequencies away from the nominal bin centers. The traditional approach to this problem is to synchronize the sampling clock to a received signal by phase locking the sampling clock to one component of the signal (Mahoney 1987). This requires a non-trivial amount of time compared to the total test time of the measurement. The FASTTEST multitone software contains a provision for correcting this frequency error after the signal is acquired, not before.

Sample rate correction in FASTTEST relies on the ability to accurately measure the frequencies in the multitone signal. The measured frequencies are compared to the known generator frequencies. The ratio of the measured frequencies to the generator frequencies represents the amount of frequency shift which must be corrected. FASTTEST uses this frequency measurement to perform a sample rate conversion on the acquired signal, shifting the frequencies to their correct values.

If the frequency shift is not measured accurately, then sample rate conversion will fail to fully "synchronize" the FFT to the acquired data. The result will be skirts around the fundamental tones in the frequency domain. These skirts fall between the tones and will produce an elevation in the total noise and distortion plots. If the frequency shifts are significant, there will also be an effect on the measured amplitudes of the fundamental tones. This is due to scalloping loss in the FFT as the tones shift substantially off the bin centers. Cabot (1996) shows that in the presence of interfering noise, the frequency measurement technique used in FASTTEST will reduce the distortion caused by synchronization errors to below the amplitude of the interfering noise.

FASTTEST is also capable of making rudimentary measurements of distortion audibility by computing the masking curve created by a particular multitone test signal. It can then compare the total distortion measurement to the perceptual limit imposed by the masking curve to assess the audibility of the distortion. This is detailed in the
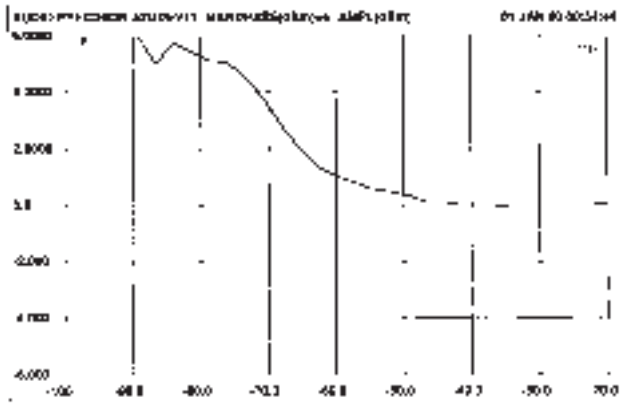
Fig. 21a. Deviation form linearity for 1 channel of a digital multitrack. 6dB worst case deviation through -90dB. Poor linearity performance.
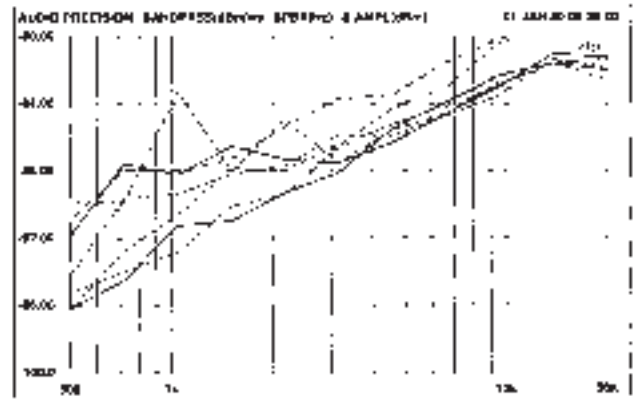

Fig. 21b. Corresponding noise modulation. 10dB noise modulation at low frequencies. 3dB noise modulation at high frequencies. Poor noise modulation performance.
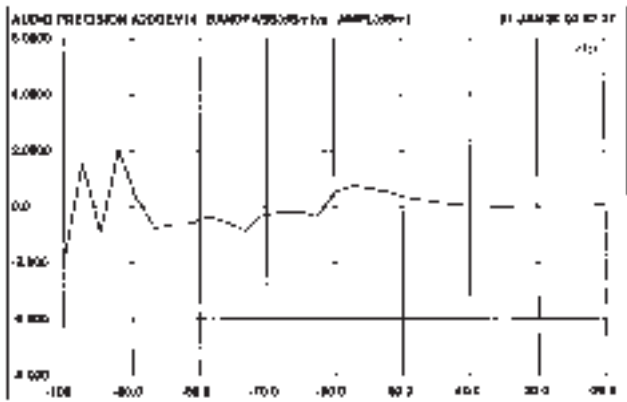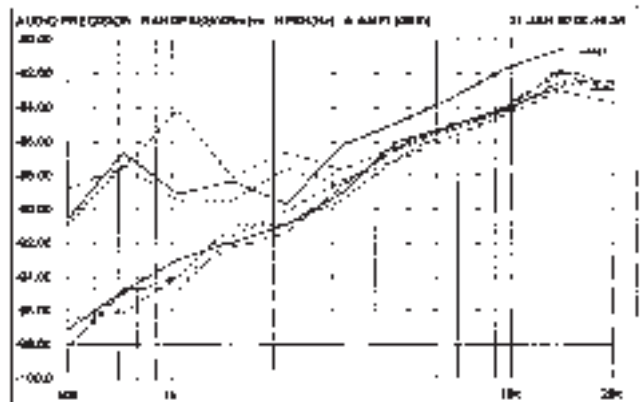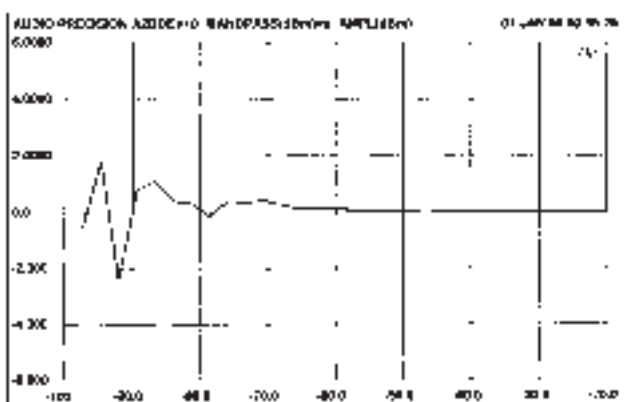

Fig. 21c. Deviation from linearity for 1 channel of a digital multitrack. 1dB worst case deviation through -90dB. Good linearity performance.

Fig. 21d. Corresponding noise modulation. 10dB noise modulation at low frequencies. 3dB noise modulation at high frequencies. Poor noise modulation performance.

Fig. 21e. Deviation from linearity for 1 channel of a digital multitrack. 1dB worst case deviation through -90dB. Good linearity performance.
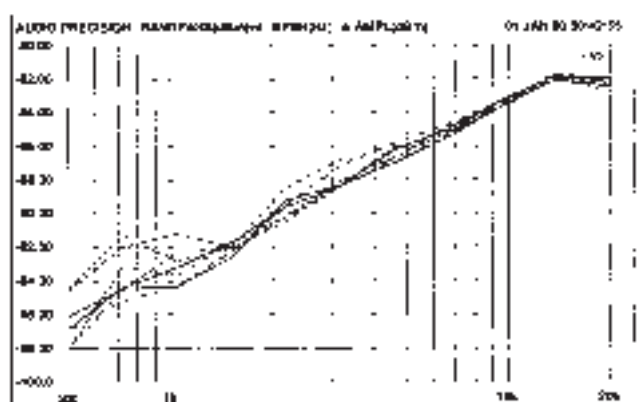

Fig. 21f. Corresponding noise modulation. 4dB noise modulation at low frequencies. 1dB noise modulation at high frequencies. Good noise modulation performance.

low bit rate coder measurements paper by Cabot (1992).

Interface Measurements

Most audio equipment today is interfaced through the AES/EBU serial digital interface or its fraternal twin for consumer use standardized by the EIAJ. When dealing with equipment connected through one of these interfaces, there are three broad areas of concern. First are problems with interface waveform parameters which affect the ability of the interface to reliably pass data from one device to the next. Second are problems which, although allowing error-free communication, affect the ultimate audio performance. Last are problems which manifest themselves only in a system environment when multiple devices are interconnected in arbitrary combinations.

As described in Cabot (1990), the AES interface is a self-clocking, polarity-independent, Manchester coded interface. The data, clocking and synchronization information are all contained in the edge timing of the stream. This makes the proper detection of edges and their location crucial not only to the interface functionality, but to its performance as well.

The AES standard specifies minimum and maximum waveform amplitudes for signals transmitted on the interface. It also specifies a minimum amplitude at which any properly functioning AES receiver must correctly receive data. This is intended to insure that all devices have adequate interface signal to correctly recover clock and data information, without being overloaded by excessive amplitude signals. Although the interface signal amplitude conveys no information, it is all in the edges, inadequate levels increase the receiver susceptibility to noise. Reduced amplitudes also increase jitter in the recovered signals due to errors in the slicing level. Proper testing of an AES interface requires the ability to control the interface waveform amplitude over a range at least as wide as that specified in AES-3. To insure margin for error when the device is used in practice, testing over a wider range is desirable.

Similarly, a received signal's ampli-tude must be checked to see that it is within the acceptable range for proper recovery. Readouts of this amplitude are normally measured peak-to-peak with a wide bandwidth peak ac-to-dc converter. Since peak measurements are used, it is essential that waveform fidelity be maintained through the path leading to the detector. Otherwise waveform tilt and overshoot will create incorrect readings. Modest amounts of tilt or overshoot are no cause for concern since the information is conveyed in the waveform edges, but the increased level they imply may mask serious problems with inadequate minimum amplitude in the body of the waveform. The only way to see the true effects of amplitude reduction, without being obscured by such waveform artifacts, is with eye patterns or histograms, as described below.

The AES-3 interface operates with bits whose widths are multiples of 1/128th of the sampling interval (the reciprocal of 128x the sample rate). At 48 kHz sample rate this works out to be 163 ns. This time period is called the Unit Interval (UI) since it defines the minimum interval on the interface. (It should be noted that this definition is different from that used in other branches of engineering such as telecom, where a Unit Interval refers to the width of a data bit and not the interface pulses.) Data ones are composed of two bits of opposite polarity, while data zeros consist of a single bit of double this width (326 ns). The synchronization patterns, called preambles, consist of these two pulse widths plus three unit interval wide pulses as illustrated in Fig. 22.

These short pulses require considerable bandwidth for proper transmission, typically 20 MHz or more. This can be seen by transforming the interface waveform into the frequency domain as shown in Fig. 23. The spectrum has components from under 1 MHz to more than 20 MHz. Discrete products are seen at 6.144 MHz and integer multiples of that frequency. If no data is sent on the interface, discrete products will also appear at ½ and ¼ of these frequencies. The spectral domain behavior is also driven by the rise times of the interface waveform as seen in the time domain plot shown earlier. Faster rise time signals will contain more energy at high frequencies than signals whose rise time is limited by interface bandwidth. More important than the effects of rise time on high frequency energy content of the interface is the interaction between rise time and interface jitter. This is illustrated in Fig. 24.

Jitter is the deviation of the interface waveform zero crossing times from the zero crossings of a perfectly stable clock whose period is one Unit Interval. It is not the deviation of the waveform pulse widths from the ideal Unit Interval width. Simply measuring the variations in pulse width by overlaying traces on a scope which are triggered by the edges of the incoming stream will only indicate edge to edge jitter. Variations in width which are correlated from pulse to pulse (all being larger than normal or all being smaller than normal) can accumulate into substantially larger deviations from the ideal edge locations.

The internal noise or instability of the clock oscillator in a device will create jitter in its digital output signal. With most devices, this inherent jitter is not large enough to cause problems with the proper reception of the digital signal. Some devices use digitally generated clocks created from a high frequency master clock which is divided down in a digital PLL to create an in-
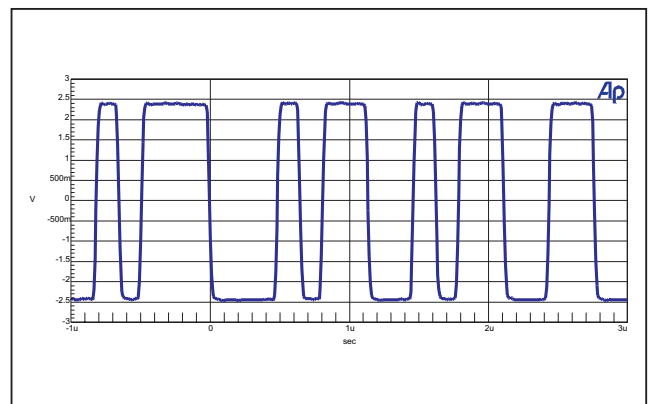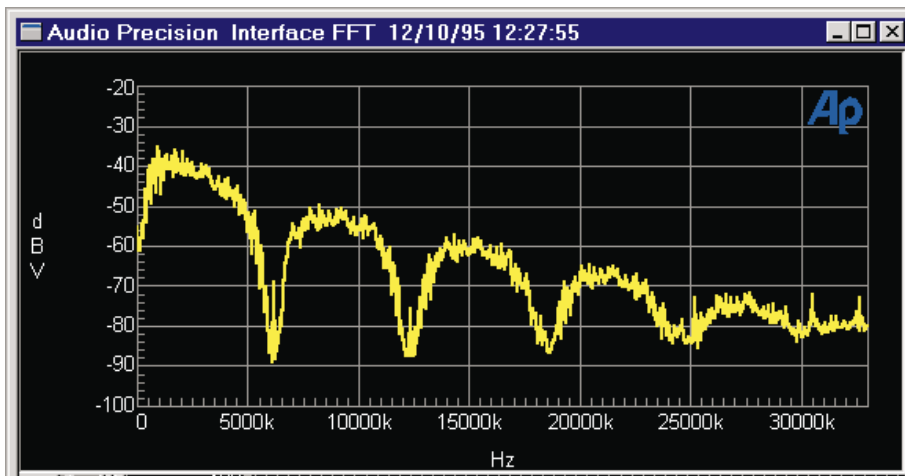

Fig. 22. AES/EBU data stream format

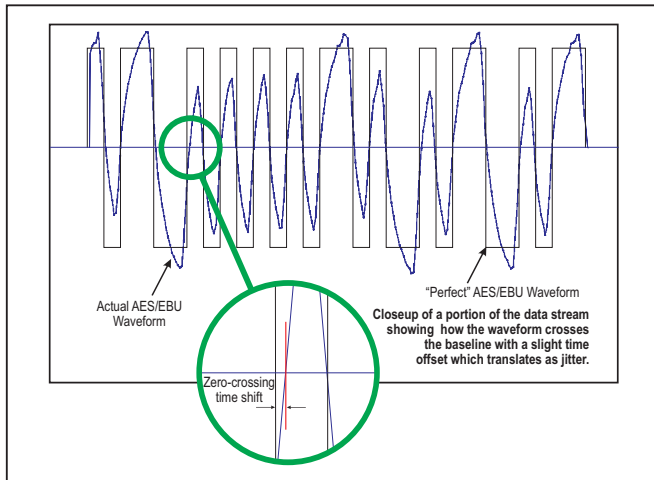**Fig. 24.** Spectrum of interface waveform.

Fig. 23. Effect of bandwidth reduction on jitter.

Actual AES/EBU Waveform

"Perfect" AES/EBU Waveform

Closeup of a portion of the data stream showing how the waveform crosses the baseline with a slight time offset which translates as jitter.
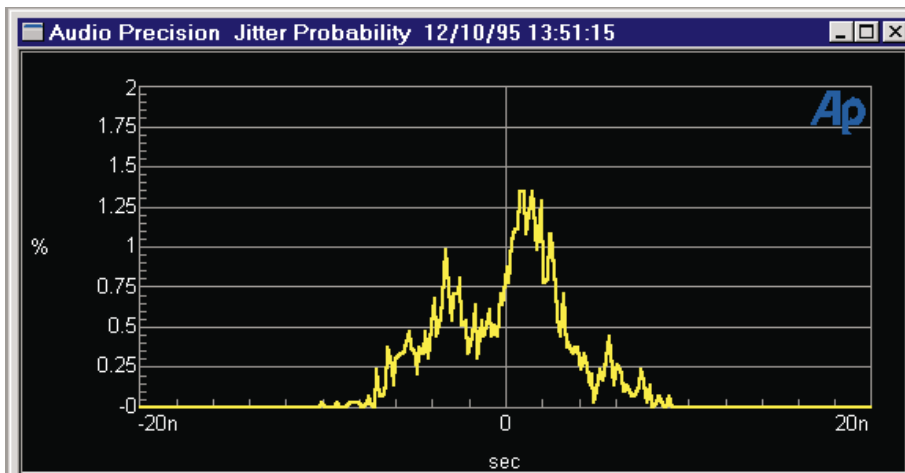
Zero-crossing time shift

Fig. 25. Jitter histogram of a noisy signal.

terface clock signal of the correct average frequency. This technique leads to notoriously jittery signals and has been responsible for generating signals which cannot be received by many receivers.

An excellent way to view the jitter behavior of a clock or interface signal, especially when it is particularly noisy, is via a histogram. Fig. 25 is an example of a jitter histogram measured from a noisy interface. The horizontal axis is the deviation of the interface signal zero crossings from their ideal positions. The vertical axis represents the likelihood of the zero crossing having that particular timing. A strongly bimodal histogram is indicative of squarewave jitter, while a bimodal distribution with a gradual transition between modes is a sign of sinewave jitter. Gaussian or skewed Gaussian shapes are indicative of random jitter.

When several digital devices are cascaded without a system-wide master synchronization signal, each receives its clock from the previous device in the chain and provides the clock to the next one in the chain. The individual devices extract the clock from the incoming interface signal and create an output signal from this clock. Unfortunately, it is common for equipment to not only pass jitter received at its input to its output, but to amplify the jitter if it is within a particular frequency range. This is caused by the response of the internal clock recovery phase locked loop (PLL). The loop is designed to track the incoming sample rate and will therefore follow slow variations in clock frequency. As the frequency of sample rate variation is increased, the loop will (hopefully) attenuate the variations. The loop response is therefore a low pass filter, allowing low frequency jitter to pass unattenuated but reducing high frequency jitter. The loop response is obtained by plotting the amplitude of jitter on the device's output signal for a fixed amplitude, but variable frequency, jittered signal at the input. Ideally this response is a lowpass function with no peaking. However, in practice, many devices have several dB of jitter gain near the corner frequency of this lowpass function. Such
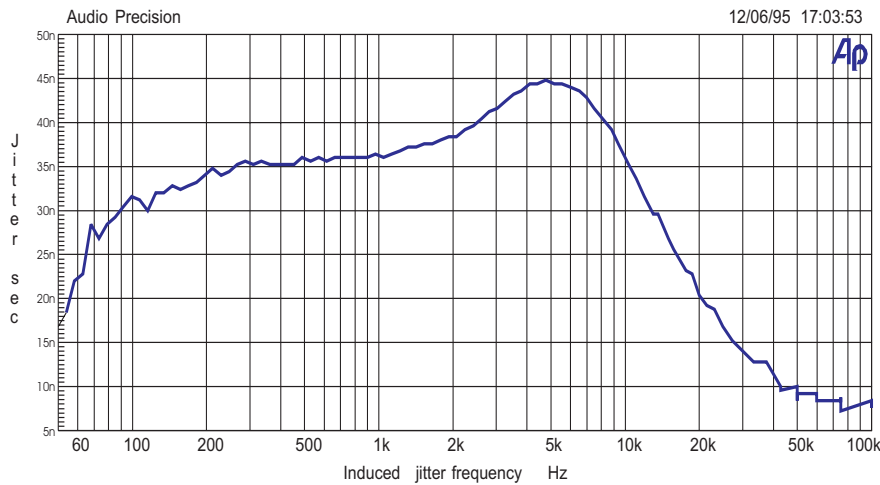
Fig. 26. Jitter transfer function of a typical device.

devices will amplify jitter occurring in that frequency range. If several such devices are cascaded the results can be disastrous for the later equipment in the chain (Dunn et al, 1993).

For equipment with an external sync reference input this jitter accumulation cannot occur, because each device extracts its output clock from the reference input and ignores the jitter on the signal input. However, the loop response from the reference input to the device output becomes the relevant parameter. Although jitter accumulation is no longer a concern, the jitter gain can produce excessive jitter in the output of an individual device.

Sinewave jitter is useful to determine the jitter transfer gain of a digital input / digital output device. It is also useful to isolate the effect of jitter as a function of frequency on a converter. By stimulating the device under test with a sinewave jittered AES-3 signal whose jitter frequency is adjustable, the jitter transfer function may be measured. An example of this for a typically medium priced digital processing device is shown in Fig. 26. The response has a broad peak in the 5 kHz region which reaches 2 dB of gain. When several of these devices are cascaded, jitter could rise to levels which would cause later devices in the chain to lose lock. Since the receiver design used in this device is a common commercially available chip used according to the manufacturer's recommendations, such a cascade is not unlikely.

The transient behavior of clock recovery circuits can be more easily assessed with squarewave jitter. The leading and trailing edge of a squarewave create a sudden shift in interface signal phase which must be followed by the device under test. This transient may cause the device to lose lock or to oscillate around the new interface phase. The loop dynamics are easy to view if the measurement equipment offers both a squarewave jitter source and a time domain display of jitter.

Any AES-3 receiver will take some finite time to acquire lock on the incoming signal. It must lock before it can recover data and before it may output a regenerated AES-3 signal.

This time from presentation of signal until locking is the receiver pll acquisition time. This can be assessed if the measurement system offers a time domain display of interface bit rate.

Eye patterns are a display of the envelope of possible interface signal waveshapes across one unit interval. By triggering an oscilloscope from a highly stable version of the recovered interface clock and setting the sweep speed to put one unit interval on screen, an eye pattern will result. Each successive sweep of the oscilloscope traces one trajectory of the interface waveform across a unit interval. As the traces curve up or down at the beginning and end of the unit interval, the display closes down to show a hole in the middle where the cell edges do not cross the horizontal axis.

The outer extremes of the eye pattern represent the maximum excursion of the waveform during the interval and essentially display the maximum peak-to-peak signal level. This is of limited utility. The inner extremes of the eye pattern (Fig. 27) represent the minimum excursion of the interface waveform during the unit interval and represent the difficulty a receiver would have decoding the signal. The AES-3 standard specifies the minimum eye-pattern size, or "opening", with which a correctly functioning receiver must operate. There is no specification in the AES standard for how the minimum eye opening is to be ob-
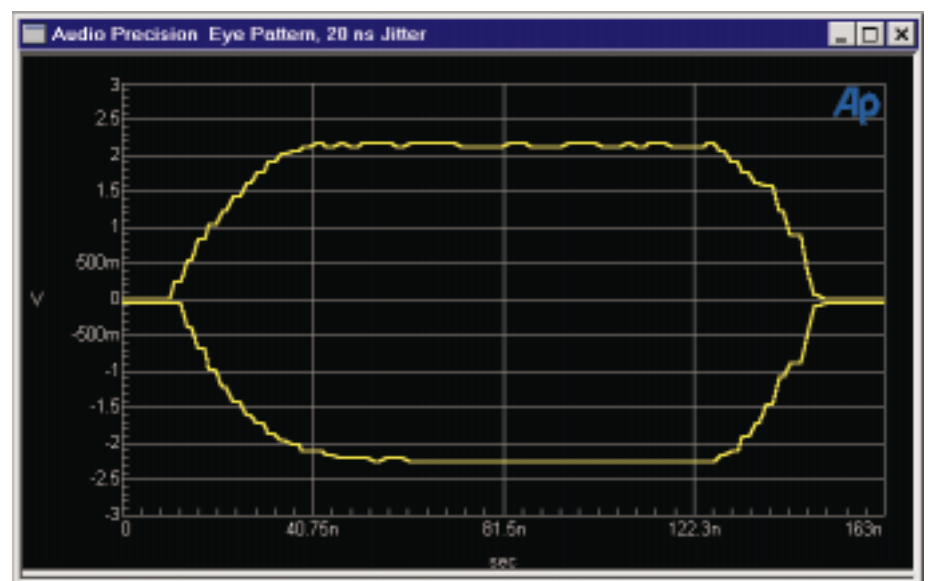

Fig. 27. Eye pattern inner traces.

tained. The height reduction may come from low signal level, cable roll-off, interfering noise, or any combination of the three. The width reduction may come from cable roll-off, jitter or a combination of the two.

Long cables create high frequency roll-off because of distributed capacitance working against the wire resistance. This high frequency attenuation progressively increases the interface signal rise time with increasing cable length. The AES standard includes a suggested equalizer to be inserted at the receiver to compensate for the cable roll-off. To see what effect this will have on signal recoverability, it is helpful to be able to switch such an equalizer in line with the signal before viewing the signal on a scope. One test equipment manufacturer has instead provided an automatically adjusting equalizer which introduces a variable degree of high frequency boost and displays the amount of boost introduced. This number indicates the degree of cable roll-off, giving a rough measure of cable quality and length as well as an indication of the fixed equalizer which should be permanently installed in the line.

The AES-3 standard specifies that receiving devices should tolerate 7 Volts peak of 20 kHz common-mode signal. The interface was originally planned to carry an analog version of the digital audio as a common-mode signal. This was never exploited in practical systems. However, in large installations there may be considerable pickup of common-mode interference from nearby power lines, video lines, data lines or ground potential differences. If the receiving device has inadequate common-mode rejection there will be leakage of this interference into the digital signal path. If the interference is low frequency it will shift the slicing point of the data comparator, creating jitter and reducing data integrity.

Interfering noise of sufficient magnitude will cause errors in a receiver of an AES-3 signal. When viewed on an eye diagram, the noise reduces the eye height, often without significantly affecting the eye width. (If the AES-3 signal suffers from limited bandwidth,
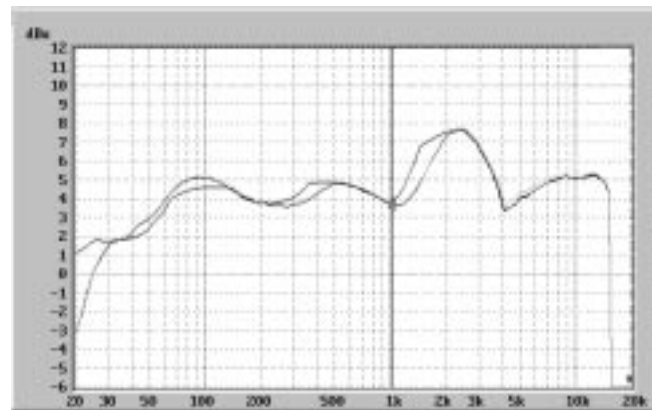
added noise will shift the zero crossings due to the finite rise times.) When the eye closes, errors are unavoidable. If the noise level is sufficient to close the eye to 200 mV the AES-3 specification requires the receiver to still correctly receive the data. By adding noise of selectable amplitude to the transmitted signal, the test equipment can degrade the signal seen by the receiver to the threshold of error and determine the receiver margin.

This paper has described several impairments which can affect AES-3 signals. In the real world these can (and do) occur simultaneously. The AES-3 standard only specifies that a receiver must correctly receive a signal whose eye height and width have shrunk to the values specified in the standard. What combination of impairments are included is not specified. The eye height may be reduced by the effects of additive noise, inadequate common-mode rejection, low signal amplitude and cable roll-off. By applying these in various combinations the robustness of a device under test may be examined. Similarly, the standard specifies that a receiver must correctly function down to an eye of one half nominal width. This narrowing of the eye might be caused by jitter, cable roll-off or shifting of the slicing point by inadequate common-mode rejection or additive noise.

Measurement stability or repeatability

The issue of measurement repeatability was discussed with regard to amplitude and frequency measurements. However, this is an issue with any measure-

ment due to noise or sampling uncertainty issues. The tradeoff is usually one of accuracy, speed, or (in the case of sampled signals and quantized signals) sample density and converter resolution. The concepts of repeatability and accuracy are quite distinct. It is entirely possible to have highly inaccurate readings which are very repeatable. Conversely, it is possible to have readings which, on average, are quite accurate but have variability in excess of their accuracy.

One aspect to getting repeatable and accurate readings is that of settling. Some measurement equipment depends on fixed delays between a setting change and the taking of a reading to obtain stabilized results. Some of these delay-based schemes are more intelligent than a single delay value for all conditions. The more advanced ones make the delay a function of frequency and also depend on the measurement being performed, taking longer for THD+N measurements than for amplitude measurements for example.

This is quite acceptable if the equipment is only used for measuring itself or for measuring some well behaved devices such as power amplifiers or gently sloping equalizers. Devices with large response variations such as high-Q notch filters, devices with dynamic characteristics such as compressors or limiters, devices with dynamic delays such as reverberators, etc. can cause serious problems when measured with fixed delay schemes. The concept of device under test settling being more stringent than the test equipment's own settling is often ig-

nored by instrument manufacturers (especially when demonstrating their products).

To deal with real-world devices Audio Precision devised a comprehensive settling algorithm which starts from minimum delay values based on the variables mentioned above. It then performs an additional step of comparing new data values to previous data to see when the readings have stabilized. Although this carries a slight time penalty for collecting the additional data readings, in a well designed implementation it is generally faster overall than the longer delays that must be used to insure adequate settling for the device under test. Fig. 28 shows the response of a typical DUT with delay-based settling and data-based settling overlaid. The errors become significant when the device's response deviates from flat.

It is common to make a direct link between generator operation and analyzer operation. This makes use of the knowledge about the generator waveform and frequency to set the analysis averaging time, measurement filters and any delays required for settling. This avoids the time required for measuring the input signal frequency in the analyzer before setting these measurement parameters. Although this works very well for situations where the generator and analyzer are co-located, it falls apart when the equipment is separated by any substantial distance or when making measurements from record/replay devices such as tape recorders.

Jitter measurements on converters

Jitter exists in all digital signals, it is only a question of magnitude. Interface jitter is jitter on the signal between two digital devices. Sampling jitter is jitter on the clock of an A/D or D/A converter. The degradation introduced by jitter on an interface depends on the design of the interface receiver. The degradation introduced by jitter on a sampling clock depends on the design of the converter. On better designed equipment, there is a stage between the interface clock recovery and the converter clock generation which filters out jitter above some

Fig. 29. Spectrum of a 5kHz signal with 3kHz sinewave induced jitter.
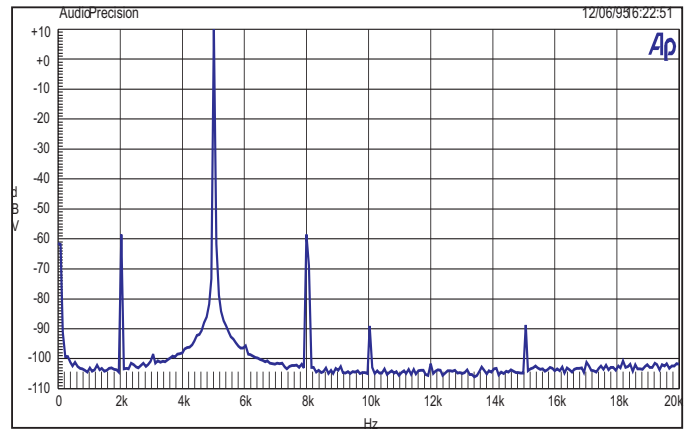
cut-off frequency, improving the audio performance of the device. It is incorrect to assume that an interface with more jitter will perform worse than one with less jitter. Similarly, it is incorrect to assume that a converter running from a clock with higher jitter will perform worse than one operating from a clock with lower jitter.

Jitter on an interface or on a D/A converter clock may appear to some degree on the reconstructed signal. This may be tested by introducing jitter on the interface and measuring the degradation of the reconstructed signal output. Jitter is a phase or time modulation effect, producing modulation sidebands on an audio signal. For sinewave jitter, the sidebands will generally also be sinusoidal and their amplitude proportional to the jitter magnitude as illustrated in Fig. 29. The transfer function of the jitter amplitude to the sideband amplitude as a function of jitter frequency may be characterized. For random jitter, the sidebands will also be random, creating an elevated noise floor.

Jitter on an A/D converter clock may appear to some degree on the sampled signal. This may be tested by introducing jitter on the clock and measuring the degradation of the digital signal output. Jitter will produce modulation sidebands on the digital signal. As with the D/A case, sinewave jitter will produce sinusoidal sidebands which may be measured to quantify the jitter. The transfer function of the jitter to the sideband level and its effect on the noise floor may be measured as described previously.

Varying the frequency of the jitter signal may have a significant effect on

the device under test behavior. The jitter is normally reduced by the filtering action of the receiver circuits in the device under test. These will filter out high frequency jitter components by virtue of the limited bandwidth of the phase lock loop. This behavior may be non-linear, depending on jitter amplitude, since many phase detectors have dead band behavior for small phase deviations. This is sometimes intentional to facilitate locking to noisy signals.

Jitter on a reference input can affect A/D conversion performance if, as is normal, the device extracts its sample clock from the reference input. If the device has no reference input, as with inexpensive processing equipment, jitter on the digital audio input will be the relevant parameter to test. To properly characterize an A/D converter, it is therefore necessary to stimulate it with a low distortion analog signal while simultaneously driving its reference input or digital audio input with a jittered digital signal. This apparatus is diagrammed in Fig. 30. Measuring converter performance with only the analog signal input or only the jittered digital input would not indicate the audio degradation. A medium-priced combined analog and digital signal processor was used for the DUT and was stimulated with 35 ns of jitter on its digital input when the analog input was driven with a 997 Hz sinewave. Fig. 31 shows the THD+N as a function of jitter frequency. The distortion is seen to rise for jitter frequencies between 5 kHz and 20 kHz.

Jitter on the digital signal input or reference input similarly affects D/A performance. To properly characterize
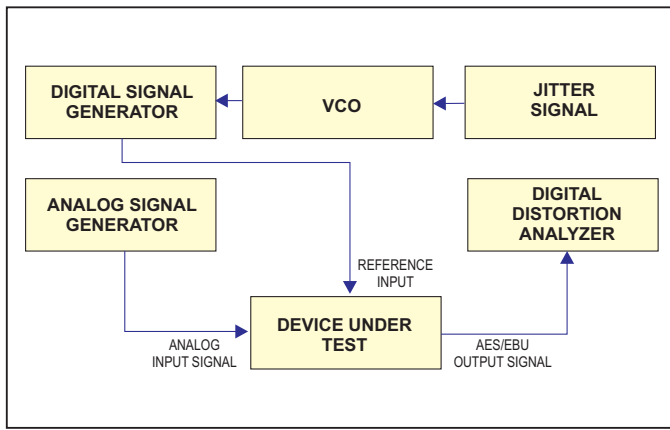
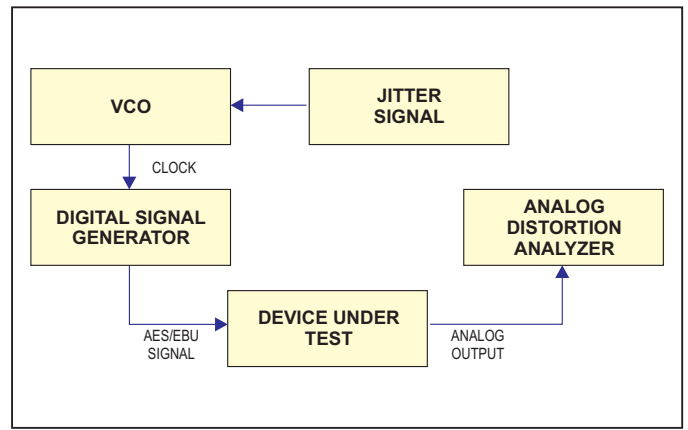**Fig. 30.** Evaluating a device for jitter susceptibility.

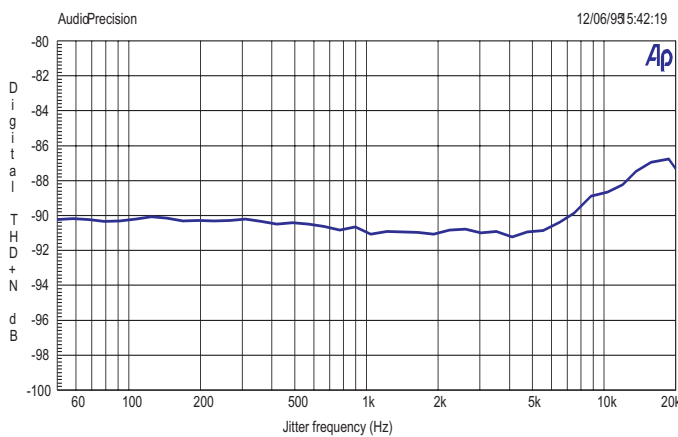**Fig. 32.** D/A jitter susceptibility characterization

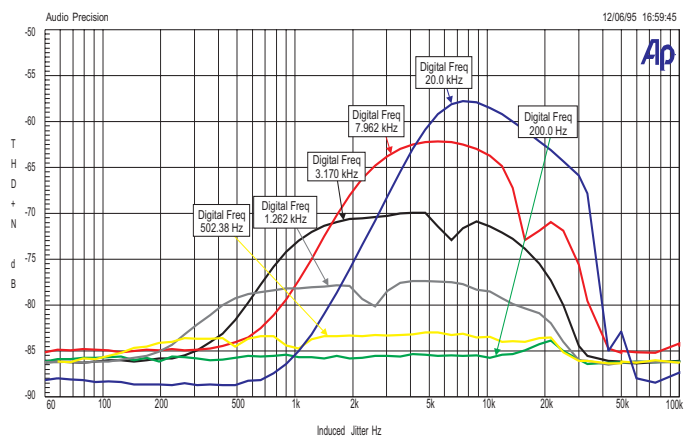**Fig. 31.** THD+N as a function of jitter frequency.

Fig. 33. THD+N as a function of jitter frequency for various signal frequencies.

a D/A, it is necessary to stimulate it with a jittered digital audio signal which carries audio information as shown in Fig. 32. Normally, jitter would pass from the digital input to the D/A clock circuits, but the converter might be more affected by the reference input if the signal is internally reclocked with the reference clock. Using the same device under test as the last example, the D/A distortion was measured as a function of jitter frequency. Again, the digital input was jittered with 35 ns of sinewave jitter and the THD+N was measured as a function of jitter frequency. The performance, shown in Fig. 33, rapidly degrades for jitter above 500 Hz.

Jitter on a digital input can affect performance in an analog to analog system through its effect on the internally recovered sampling clocks. Ultimately, it comes down to how well the clock recovery circuits can extract a stable clock from the interface. Audio performance of digital to digital processing equipment may also be affected if the equipment contains any sample rate conversion stages. They may be tested in much the same way as D/A converters, although the distortion measurements are made in the digital domain. If there is no sample rate conversion, the digital input jitter will pass through to the output with some gain or loss. This may cause interfacing problems if the jitter gain is excessive. Excessive jitter gain may cause the performance of the final digital to analog conversion stage to suffer but it will not introduce any audio degradation within the device itself.

Summary

Various analog and digital audio measurements were described. The architectures typically used in audio test equipment were reviewed. The strong need for simultaneity of digital and analog generation was presented and its application to converter measurements was explained. The advantages of simultaneous measurement in multiple domains was similarly detailed.

Novel techniques employing multitone signals for fast audio measurements were examined and applications of sampling frequency correction technology to this and conventional FFT measurement were covered. Synchronous averaging of FFT data was presented and the subsequent noise reduction demonstrated.

made many of the comparative measurements and graciously supplied the material excerpted from his book. My wife Liane patiently waited while I occasionally gave this paper more attention than her.

The *FAST*TEST frequency correction technology, *FAST*TEST total distortion technique, *FAST*TEST triggering technology, synchronous FFT averaging, and other technologies described in this paper are the subject of US patents issued, pending or applied for.

References

Data sheet for the AD536A rms to DC converter, Analog Devices, Norwood MA, 1992

Audio Engineering Society, AES standard method for digital audio engineering–Measurement of digital audio equipment, AES17-1992

Audio Engineering Society, AES standard method for digital audio engineering–Serial transmission format for two-channel linearly represented digital audio data, AES3-1996

Owners manual for the Model 5500 Audio Measurement System, Amber Electro Design, Montreal Canada, 1986

R. C. Cabot, "Testing Digital Audio Devices in the Digital Domain," presented at the 86th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 37, p. 399 (1989 May), preprint 2800.

R. C. Cabot, "Performance Limitations of Digital Filter Architectures," presented at the 89th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 38, p. 871 (1990 Nov.), preprint 2964.

R. C. Cabot, "Measuring AES/EBU Digital Interfaces," *J. Audio Eng. Soc.*, vol. 38, pp. 459-468 (1990 June).

R. C. Cabot, "Noise Modulation in Digital Audio Equipment," presented at the 90th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.*, vol. 39, p. 382 (1991 May), preprint 3021.

R. C. Cabot, "Comparison of Nonlinear Distortion Measurement Methods," in *Proc. AES 11th Int. Conf. On Audio Test & Measurement*, pp. 53-65 (1992).

R. C. Cabot, "Measurements on Low Bit Rate Coders," *Proc. AES 11th Int. Conf. On Audio Test & Measurement*, pp. 216-222 (1992).

R. C. Cabot, "Measuring the Effects of Sampling Jitter," presented at the 99th of the Audio Engineering Society, Workshop *Jitter: Is it a Problem?* (1995 Oct.).

R. C. Cabot, "Synchronous FFTs; Its importance With Multitone Testing," *AUDIO.TST*, vol. 11 #2 (1996).

J. Dunn, B. McKibbon, R. Taylor and C. Travis, "Towards Common Specifications for Digital Audio Interface Jitter," presented at the 95th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 41, p. 1051 (1993 Dec), preprint 3705.

R. A. Finger, "On the Use of Computer-Generated Dithered Test Signals," *J. Audio Eng. Soc.*, vol. 35, pp. 434-445 (1987 June).

R. A. Finger, "Review of Frequencies and Levels for Digital Audio Performance Measurements. *J. Audio Eng. Soc.*, vol. 34, pp. 36-48 (1986 Jan./Feb.).

M. Mahoney, *DSP-Based Testing of Analog and Mixed Signal Circuits*, IEEE Press (1987).

R. E. Metzler, *Audio Measurement Handbook*, Audio Precision Inc., Beaverton, OR, USA (1993).

R. E. Metzler, "Test and Calibration Applications of Multitone Signals," in *Proc. AES 11th Int. Conf. On Audio Test & Measurement*, pp. 29-36 (1992).

E. Rosenfeld, "Accuracy and Repeatability with DSP Test Methods," IEEE 1986 International Test Conference, paper 21.5, pp. 788-795.

E. Rosenfeld, "DSP Measurement of Frequency," IEEE 1986 International Test Conference, paper 26.2, pp. 981-986.

J. Tierney, C. Rader, and B. Gold, "A Digital Frequency Synthesizer," *IEEE Trans. on Audio and Electroacoustics*, vol. AU-19, No. 1, (1971 Mar.).

J. Vanderkooy and S. P. Lipshitz, "Dither in Digital Audio," J. *Audio Eng. Soc.*, vol. 35, pp. 966–975 (1987 Dec.).

THE AUTHOR

Richard C. Cabot received B.S., Masters and Ph.D. degrees from Rensselaer Polytechnic Institute, Troy, N.Y. After teaching at Rensselaer, he joined Tektronix, and later left to help found Audio Precision Inc., a firm specializing in audio test and measurement equipment for the professional audio industry. He holds the position of Chief Technology Officer and is responsible for all engineering and product development.

Dr. Cabot is active in the AES, having served as President, Vice President and Governor, and is currently chairman of the AES working group on digital audio measurements. Dr. Cabot is a Fellow of the AES, a Senior Member of the IEEE and a member of numerous related technical and professional organizations. He holds 10 patents and has written extensively for industry publications. He is a registered professional engineer in both Electrical Engineering and Acoustics. He was also appointed to a term on the Oregon Central Business Registration Coordinating Council, advising state government on streamlining the registration and regulation of businesses in Oregon.