A Comparison of Nonlinear Distortion Measurement Methods

Conference Paper · May 1980

1 author:

Richard Cabot
XFRM, Inc.

COMPARISON OF NONLINEAR DISTORTION MEASUREMENT METHODS

Richard C. Cabot
Vice President
Audio Precision, Inc.
P.O. Box 2209
Beaverton, OR 97075

Several techniques are currently in use for measuring distortion of audio equipment. These include THD, SMPTE intermodulation, difference frequency intermodulation, DIM (sinewave-squarewave combination). A new technique is proposed which uses a relatively large number of sinewaves to effect complex intermodulation products across the entire audio band. This paper compares the various methods, both theoretically and practically. Examples of measurements on several test circuits are presented to illustrate the results.

INTRODUCTION

For many years there has been confusion surrounding distortion and its measurement. This paper will not discuss in detail the causes of "static" or "dynamic" distortion, or the many misconceptions surrounding their reduction or elimination. This has been done by others extremely well (Cordell 1979, Hofer 1980). Instead, a comparison of the distortion measurement techniques available today and their relationship to some common distortion mechanisms will be presented. A new distortion measurement technique using numerous simultaneous sinewaves will also be presented and compared with the more conventional techniques under the same test conditions.

Many people confuse the causes or forms of distortion with measures of distortion (see for example Frey 1979). This confusion occurs frequently in popular press articles. Harmonic distortion is often thought to be something different from intermodulation distortion. However, the device under consideration has an inherent nonlinearity which will generate harmonics, intermodulation products or both when stimulated by a signal. If the signal is a single tone, only harmonic products may be generated. Both harmonics and intermodulation components will be generated when the excitation is a multitone signal. Digital audio or other sampled data systems may shift the harmonic and intermodulation frequencies due to the folding action of the sampling process. This can, for example, result in harmonics appearing at frequencies which are not harmonic multiples of the input.

Since some of the test circuits in this paper demonstrate the measurement of "transient intermodulation distortion" or "TIM" some initial comments concerning it are in order. The word transient implies that the effect is not connected with steady state excitations of the device under test. The word transient carries a mystique which is helpful in marketing audio equipment but really does not apply to the behavior under consideration. The other half of the phrase TIM is also a misnomer. The distortion generated by the common TIM mechanisms will result in harmonics of the input as Jung 1977 and others have shown. These TIM mechanisms are really the early manifestations of slew rate limiting. The distortion mechanism commonly referred to as TIM is simply the result of nonlinear operation of a stage in the amplifier at high frequencies due to the large current required to charge and discharge the compensation capacitor.

The term Slewing Induced Distortion (SID) has been suggested by Jung as an alternative to TIM. This is certainly preferable and is correct if interpreted the way he intended. However some people are misled into thinking that the device under test must be slew rate limiting for the distortion to occur. A better term would be Slope Induced Distortion to emphasize the signal rate of change is responsible for exciting the nonlinearity. Also, some have mistakenly assumed that TIM and SID are different mechanisms, but they are merely different names for the same thing.

FORMS OF DISTORTION

Some of the many forms of nonlinear distortion are explained below, however this list is by no means complete. A thorough treatment of each of these is outside the scope of this paper. However an understanding of the variety of anomalies affecting audio circuits is helpful in comparing the many methods of measurement available and their advantages and disadvantages.

There are many sources of distortion which result from basic device nonlinearity. These are the so called "static" distortions, or transfer characteristic nonlinearities. Some examples are:

1. Input stage nonlinearity. Due to the logarithmic base-emitter characteristics of bipolar transistors, most input

stages of op-amps or power amps which use these devices follow a hyperbolic tangent voltage-to-current transfer characteristic.

2. Crossover distortion. In class B amplifiers where the output devices conduct for only half the cycle, discontinuities often arise in the region near zero output. As one device is turning off the total output impedance increases resulting in a region of reduced gain near zero output.

3. Mismatched output stages. This problem is most common in common emitter output stage amplifiers where the output stages have gain. If the gains of the halves are not matched exactly, the transfer characteristics will be asymmetrical.

4. Non-ohmic contacts in transistor sockets or solder joints can create a slight rectification action.

5. Input stage common mode distortion arises from non-ideal current sources in the input differential pair of most amplifiers. As the signal moves the input stage through its common mode range, the current through the differential pair will change. This modulates the gain, producing distortion.

These mechanisms have their parallels in devices other than power amplifiers; nonlinear cone suspensions in loudspeakers, tape saturation in magnetic recording, groove deformation in discs, etc. Digital-to-analog or analog-to-digital converters often exhibit a significant discontinuity at the zero crossing creating behavior similar to that of crossover distortion. These effects are not really frequency dependent, although their magnitude may change with the changes in feedback factor with frequency. Thus, the distortion varies with frequency in direct proportion to the change in feedback with frequency. This distinguishes such distortions from ones which inherently change magnitude with frequency - often called "dynamic" distortions.

Since dynamic distortions change with frequency, even without a change in feedback with frequency, the addition of frequency dependent feedback creates a high order dependence of distortion on frequency. For example, Jung 1977 shows that THD measurements on amplifiers with "soft TIM" rise as a cubic function of frequency. The term dynamic distortion has commonly referred to mechanisms which get worse with fast or high frequency signals. Some engineers restrict it further to mean the distortion generated by the input stage driving the amplifier compensation capacitance.

If we adopt the interpretation that dynamic distortion refers to all mechanisms which get worse with fast or high frequency signals the following might be considered "dynamic distortion."

1. Charge storage effects. Once turned on, bipolar transistors collect charge in the base-emitter region which tries to keep the transistor conducting even after it is supposed to turn off. This is usually a problem in output devices of a power amp, where large currents are involved and when the drive circuits lack the ability to remove this charge fast enough.

2. Nonlinear collector-base capacitance of transistors results in a nonlinear gain when the collector to base voltage changes with the signal. Since the offending element is a capacitance, the effect worsens with frequency.

3. Slewing induced distortion or "soft TIM" commonly refers to the rise in current drive required at high frequencies to charge the amplifier compensation capacitance and the subsequent nonlinearity of the stage supplying this current. The nonlinearity itself is normally a "static" one, such as the hyperbolic tangent nonlinearity of the input differential pair. This nonlinearity is merely excited more with increasing frequency because of the increased output required of it as frequency goes up.

4. Electrostatic coupling between very nonlinear subcircuits (such as power supply lines or meter driver rectifiers) and sensitive nodes like op amp summing junctions can produce a distortion which inherently increases with frequency.

5. Marginally stable circuits can often be driven into instability by high frequency high level excitation. This may appear as a distortion which gets worse as frequency increases.

6. Electromagnetic coupling of power supply leads to feedback leads (Hofer 1980) can drastically increase distortion. Most power supply lines conduct the majority of their current for only a portion of the input cycle. Since this current is coupled into the feedback path the amplifier perceives it as signal and amplifies it accordingly.

There are also many distortion mechanisms which increase with decreasing frequency. These can also be called dynamic distortion since their magnitudes change with the frequency of the input signal, even before the inclusion of frequency dependent feedback.

1. Fuse distortion (Greiner 1979) has been known for some time to be a major source of low frequency distortion. It arises from thermal changes in the resistance of the fuse due to heating by a large signal current.

2. Thermal distortion may be a factor in the active devices themselves, if the relative time constants are short enough to allow junction temperature to change with the low frequency signal voltage.

3. Power supply rejection, or lack thereof, can cause distortion due to the signal currents modulating the power supply voltages as large amounts of power are delivered to a load. This is usually complicated by the fact that the filter capacitors are only being recharged at a rate of twice the mains frequency. This makes the distortion an interesting function of the signal frequency and the mains frequency as well as the power supply regulation factor.

4. Output terminator distortion can arise when high level, low frequency, currents begin to saturate the output inductor of the RLC output compensation network (Stanley and McLaughlin 1977).

5. Capacitors can cause distortion due to voltage coefficients or lack of polarizing voltage. This distortion will get worse with decreasing frequency, again making a difference between midband distortion and very low frequency distor-

tion. For an excellent review of the theory and practice concerning this see Jung and Marsh 1980.

Digital encoding and decoding has introduced some new sources of nonlinearity, both static and dynamic. For a discussion of the effects of MSB errors in converters and its interaction with dither see Cabot 1991. There are undoubtedly other distortion mechanisms which should be included in the lists above. These are presented since they will be valuable later in discussing the efficiency and appropriateness of various measurement methods.

DISTORTION MEASUREMENT METHODS

Numerous techniques have been proposed for measuring static and dynamic distortion. These range from conventional techniques such as THD and twin-tone IMD to highly unusual methods such as the three-tone test of Cordell and the comb-spectrum test of Jensen and Sokolich 1988. Each has been proposed by its proponents to be the best for a particular, or for all, distortion measurement applications. The ones which have survived and are in common use today are:

Harmonic distortion
SMPTE intermodulation distortion
CCIF intermodulation distortion
Sine-square intermodulation distortion (DIM)

These measurement methods will be examined in more detail below. The major factors to consider when examining any audio test procedure are as follows:

1. Signal spectral characteristics
2. Signal time domain characteristics
3. Information delivered
4. Sensitivity
5. Residual or noise floor limitations
6. Instrumentation complexity and cost
7. Ease of use
8. Adaptability to use with all audio devices
9. Correlation with audibility

This list is not in order of importance, since clearly item 9 is the most relevant to the listening experience. However, the experiments necessary to provide such data have not been done in sufficient detail to permit drawing reliable conclusions. The concept of information delivered (item #3 above) has conflicting requirements. From the consumers viewpoint there should be one test which exposes all forms of distortion, weighted as by the ear. This would allow simple comparison of equipment when making a purchase. The equipment designer would like to have one test for each source of distortion. This would enable pinpointing the area requiring attention in a design, and to evaluate the improvement resulting from a modification. This touches on a distinct advantage of tests which are a function of frequency. Insight about the cause of the distortion can often be obtained by varying the frequency of the test.

This paper presents some experimental data on circuits designed to stimulate some of the various causes of distortion. The circuits are the same as used in Cabot 1980 which were based on those used a similar study by Leinonen and Otala 1978.

Harmonic Distortion

Harmonic distortion is probably the oldest and most universally accepted method of measuring linearity. 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. Performing a spectral analysis on this signal will show that in addition to the original input sinewave, there will be components at harmonics (multiples) of the input frequency. Total harmonic distortion (THD) is then defined as the ratio of the RMS voltage of the harmonics to that of the fundamental (input frequency) component. 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). Alternately 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 difference in the measured distortion. Therefore measurements with this system are called THD+N to emphasize the noise contribution's presence.

The use of a sine wave test signal has the distinct advantage of simplicity, both in instrumentation and in use. This simplicity has an additional benefit as 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 slope 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 show 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 is 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.

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 19kHz or 38kHz is related to subcarrier problems in FM receivers. Jung, Stephens and Todd 1977 have shown a rise in harmonic distortion proportional to frequency cubed when slope distortions (soft TIM) are encountered. This will be discussed in more detail in the discussion on measurement results.

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. A tape recorder with a cutoff frequency of 22kHz (typical for a good reel to reel) will only allow measurement of third harmonic of a 7kHz input. Total harmonic distortion measurements on a 20kHz input are impossible because all 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 analyzer used.

Some questions have been raised about the audibility of harmonic distortion. For low order distortions, the masking of the ear will help to suppress them. Also, the low order harmonics are usually dominant in musical instruments, further helping to mask them. This led Shorter (1950) to propose weighting the higher order harmonics more heavily. It has also been suggested that high frequency harmonic distortion measurements aren't relevant since the components would not be audible. This argument would appear to be quite valid except for the following. When dealing with a measurement technique it is not necessary for the quantity of interest to be the quantity actually measured. It is certainly acceptable to measure the rise time of a system and infer the bandwidth. Except in the cases of a grossly underdamped or heavily overdamped system, the bandwidth is quite well predictable from the rise time. Why then shouldn't it be acceptable to measure 20kHz harmonic distortion to assess the high frequency linearity of a system? As long as the bandwidth is available to measure the distortion components, the test will still have value. It may not have a direct correspondence with the audible quality, but it is doubtful that any other test procedure does.

SMPTE Intermodulation

Intermodulation measurements using the SMPTE method have been around since the 1930s (Hilliard 1941). The test signal is a low frequency (usually 60Hz) and a high frequency (usually 7kHz) tone, summed together in a 4 to 1 amplitude ratio. 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 root mean 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 sets 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 tone. 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. This simplifies design of the remaining circuitry.

Considering the SMPTE test in the time domain it becomes quite easy to 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 (Stanley and McLaughlin 1977). 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 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.

It is often claimed that because the distortion components in a SMPTE test are not harmonically related to either input, they will be more noticeable to the ear. Musical instruments are rich in harmonics, but contain few if any components which are inharmonic. With the typical 60Hz low frequency tone used in SMPTE measurements the sidebands will be within the masking range of the ear. As with 20kHz THD measurements, it is quite possible for the test to be indicative of the audible performance of the device even if the test signal distortion is not audible.

One advantage in sensitivity that the SMPTE test has in detecting low frequency distortion mechanisms is that the 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.

The inherent insensitivity to wow and flutter has fostered widespread use of the SMPTE test in applications which involve recording of the signal. Much use is made of SMPTE IM in the disc recording and film industries (Read and Scoville 1948, Roys 1947, Roys 1953, Stephani and Bluthgen 1979). 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.

Several papers have been written (Callendar and Matthews 1951, Maxwell 1953, Waddington 1964, Warren and Hewlett 1948) which compare SMPTE intermodulation readings to harmonic distortion reading. For most classic transfer characteristic nonlinearities the SMPTE test is approximately 12dB more sensitive. However when heavy feedback is used or when dynamic effects are present the difference becomes considerably less predictable. An excellent introduction to the theory and application of SMPTE IM measurements may be found in Scott 1945.

CCIF Intermodulation

The CCIF IM distortion test differs from the SMPTE test in that a pair of signals closely spaced in frequency are applied to the device under test. The nonlinearity in the device under test causes intermodulation products between the two signals which are subsequently measured. 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 done with only a multi-pole low pass filter. Measurement of the odd order component requires spectrum analysis. Aagard 1958 and Maxwell 1953 studied the CCIF test in the 1950's and concluded that it had several advantages over either harmonic or SMPTE IM testing. The signals and distortion components may almost always be arranged to be in a passband of a nonlinear system. At low frequencies the required spacing becomes proportionately smaller, requiring a higher resolution in the spectrum analysis. At such frequencies a THD measurement may be more convenient.

The distortion products generated in this test are usually very far removed from the input signal. This positions them outside the range of the auditory systems masking effects. If a test which measures what the ear might hear is desired, the CCIF test is a good candidate. The question of correlation with audibility for this test is covered very well by Stanley and McLaughlin 1977. However, whether one component of the test is more audible than another (second order more than third or vice versa) is not covered. For any particular set of test frequencies audibility must be determined by applying the masking characteristics of the ear.

Moller 1979 discusses the use of swept frequency two tone IM tests to study "TIM" or slope induced distortion. This approach has numerous advantages over the specialized tests which have been designed for the same purpose. The ability to adjust the test frequency enables qualitative study of the distortion mechanism. Factors such as the steepness of the change in distortion with frequency and the frequency at which the change begins are useful in separating static from dynamic distortion. The dominant order of distortion is useful in pinpointing the nonlinearity.

The totally in band character of the test is one of its most attractive attributes. Unlike the sine-square test, the CCIF test may be used to severely band limited system. Distortion both before and after the band limiting point will be tested. Methods which lose their sensitivity with extreme band limiting can only test the circuitry before the bandlimiting.

Some insight into the performance of the CCIF test as an indicator of SID can be obtained by examining the normalized derivative of the signal. The peak amplitude of the CCIF signal is twice that of a single sine wave. If $\omega 1$ and $\omega 2$ are the two input frequencies, normalizing the peak derivative by the peak amplitude we find that the normalized peak rate of change is ($\omega 1$ & $\omega 2$)/2. This is equal to that of a sine wave at the average input frequency. If the sine waves are very close in frequency compared to their mean frequency we find that the signal derivative will have the same general shape as the signal itself. Thus, the CCIF test should be similar to a THD test at the mean frequency in terms of its sensitivity to SID, but it will allow all of the resulting high order components to appear in band. The exact sensitivity depends on the type of nonlinearity and the distortion order measured, as well as the signal frequency and level.

Sine-Square Test

The sine-square test was originally proposed by Schrock 1975 and later modified by Leinonen, Otala and Curl 1976. This consists of summing a sine wave at 15kHz and a square wave at 3.18kHz and applying them to the device under test. The signal is low pass filtered at either 30kHz or 100kHz, and is named the DIM (for dynamic intermodulation) 30 or DIM 100 respectively. As the amplifier slews on the corners of the square wave it becomes nonlinear and distorts the sinewave. The resulting intermodulation components are measured with a spectrum analyzer, RMS summed, and expressed as a percent of the 15kHz sinewave amplitude.

Leinonen, Otala and Curl 1976 specify nine components in the audio band which need to be summed. However this is only a portion of the components generated. In testing circuit 4 (described later) at 10V output there were 25 components within the 20kHz audio band that were greater than .1% of the 15kHz amplitude. Making this measurement is not trivial. Hofer 1986 and Skritek independently developed a technique to simplify the instrumentation and extend its dynamic range. It involves summing only two of the components, but they are selected to reflect both even and odd order nonlinearities. Extensive testing described in his paper validates the correlation of the modified technique with the values obtained from the original Leinonen et al technique.

This procedure is a fixed frequency test, which gives no insight into the mechanisms responsible for the distortion observed. The cross products which give rise to the various components are so complex that it is impossible to work backwards from the measured values to infer the distortion characteristics in the time domain. All one obtains from the measurement is a single number.

The method measures both static and dynamic distortions, some very well — others very poorly. Leinonen et all suggest using a triangle wave of the same peak amplitude as the square wave, summed with the 15kHz sine wave to measure the static components only. Their premise is that the slow rise of the triangle will make it about as sensitive to static distortion as the SMPTE IM test (which it closely resembles in the time domain). Unfortunately, the level of the fundamental in the two tests is lower by 4dB in the triangle case. This changes the sensitivity of the test to static distortions and makes direct comparison between the two impossible in the general case. Although the higher order components are still present they are reduced in amplitude. It is therefore impossible to compare the transfer function nonlinearity in the two cases. Even if the static components could be measured separately in this manner, the way the static and dynamic components in the sine-square test add is dependent on their phase and is therefore unknown. For example, two equal level components at the same frequency may add in phase to get a resultant at +6dB or they may subtract to get no resultant. A 90 phase difference between them will give a +3dB resultant.

Since the sensitivity of both tests comes from their slew rate, it is instructive to calculate it. The peak slew rate for a DIM30 signal is 0.32 V/µsec/V. The DIM100 signal peak slew rate is 1.0 V/µsec/V. The inherent slew rate of the DIM100 test is quite excessive and not very representative of real life signals. If we examine the spectrum of the DIM100 test, we find that 24% of the power in the signal is above 20kHz (outside the audio band). Eighteen percent is greater than 30kHz, while a 50kHz band contains all but 13%.

The DIM30 is probably not too bad as an audio frequency test, though the broadcast version called DIM-B is probably the most realistic. Most microphones have steep rolloffs above 25kHz or so and digital audio equipment will not pass anything above 22 kHz. The DIM30 signal was used for all of the measurements reported in this paper.

FASTtest^{tm} Total Distortion

The FASTtest total distortion measurement was developed and described by Cabot 1991. The distortion measurement is part of a technique which allows very fast measurement of linear errors such as amplitude and phase response vs frequency, interchannel crosstalk and noise. Originally developed to allow very fast measurements of broadcast links, the technique has also found wide application in production test, due to its high speed, and in tape recorder testing, since it does not need synchronization between source and receiver.

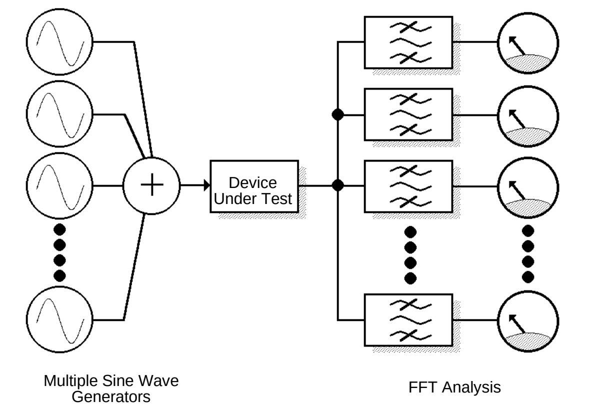

Figure 1 Basic FASTtest Measurement Technique

The operation of the FASTtest measurement technique is illustrated in Figure 1. 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.

Most of the distortion products will fall between the original stimulus frequencies and will include both harmonics and intermodulation products of these frequencies. The FASTtest total distortion measure 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 as illustrated in Figure 2. 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.

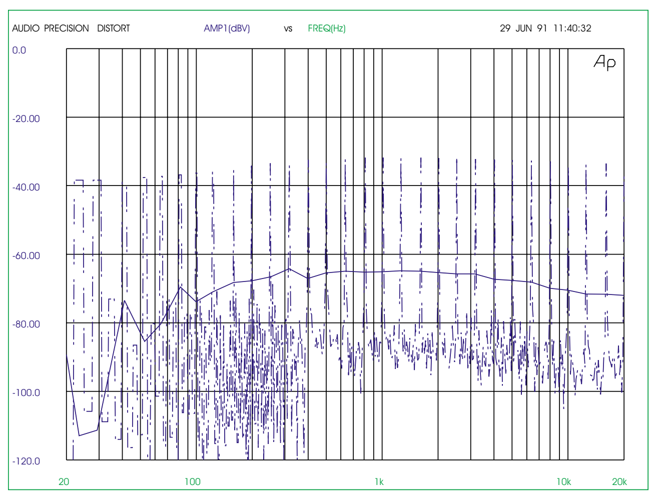

Figure 2 FASTtest Distortion Measurement Technique

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 number of individual sinewaves in the FASTtest 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 this results in a 5.96Hz frequency resolution. This capability 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. 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. The resulting crest factor will be proportional to the square root of the number of tones and is the maximum possible for the particular 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. However, for some signal spectra it is possible to produce a crest factor equal to a single sinewave. For the measurements reported in this paper a 59 tone test signal was used. The frequencies were spaced every 1/6th octave from 18 Hz to 22 kHz. The crest factor was 3.54, approximately 2.5 times that of a single sinewave.

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.

TEST CIRCUITS

Five circuits, similar to those of Leinonen and Otala 1978 were constructed for evaluating the various tests. All circuits were built around an NE5533 dual op amp, one half simulating the input stage and one half simulating the output stage. The dominant pole compensation capacitance is placed between the two stages, with a corner frequency of 2.1kHz. Both stages are wired as non-inverting amplifiers. Nonlinearities are introduced into one stage or the other and the relative gains are adjusted to provide a good model of the defect under study. The input impedance of each circuit was set at 600 Ohms, with no attenuation in the network. All circuits were run from 12 volt power supplies, adequately decoupled to prevent any instabilities. The circuits were tested with a high impedance (100 kOhm) load.

Circuit #1

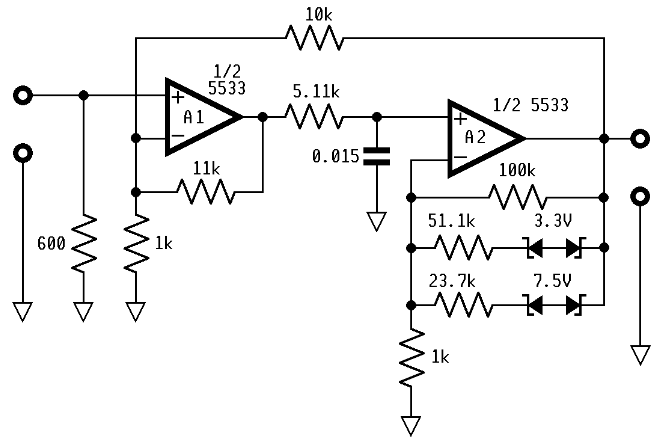

Figure 3 Output Stage Non-linearity

Symmetrically nonlinear output stage distortion is simulated by the circuit shown in Figure 3. The four zener diodes in the feedback loop of A2 provide a smooth nonlinearity. It was discovered later that the impedances used in the output stage feedback loop were too high, creating a high frequency rolloff.

Circuit #2

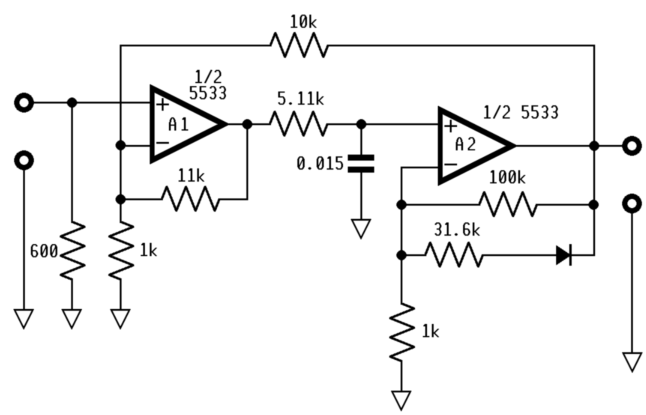

Figure 4 Output Stage Asymmetry

The circuit in Figure 4 simulates an asymmetrical output stage via the diode in the feedback loop of the second op amp. Again, the impedances are sufficiently high that rolloff in the output stage can be noticed.

Circuit #3

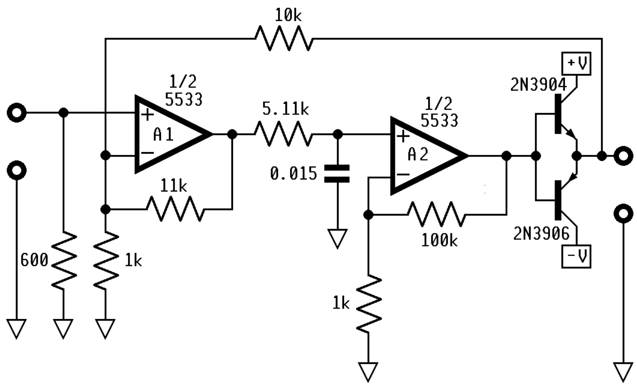

Figure 5 Crossover Distortion

The output stage crossover distortion simulation circuit in Figure 5 uses a pair of unbiased transistors within the overall feedback loop. The two 11k resistors used in Leinonen and Otala are superfluous and have been removed. One additional anomaly which was noticed early on is the lack of a path for feedback around the entire circuit. Since the transistors are totally unbiased, the output amplifier A2 must turn on one of the two transistors before feedback can get to A1 through the 10k feedback resistor. This creates an asymmetric crossover distortion, not a common problem in amplifiers and not what was expected.

Circuit #4

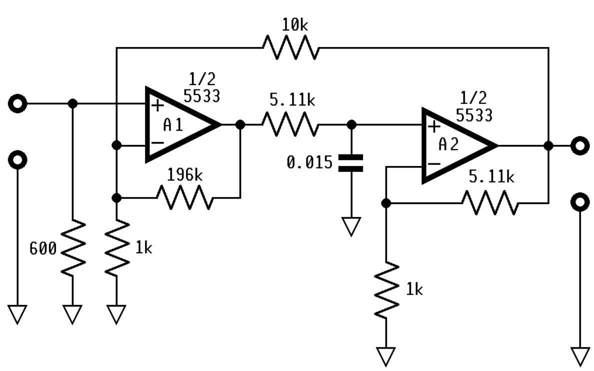

Figure 6 Hard Slew Rate Limiting

The case of hard slew rate limiting is modeled by having very high gain in the first stage and correspondingly lower gain in the output stage as shown in Figure 6. This creates a fairly linear system until the voltage swings required of the input stage to drive the compensation capacitance cause clipping and subsequent distortion

Circuit #5

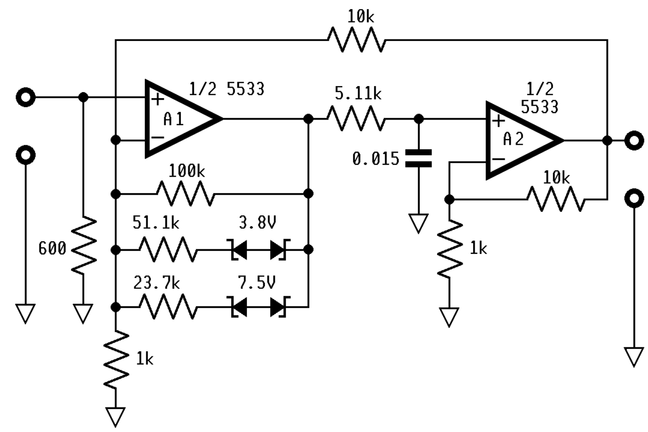

Figure 7 Input Stage Non-linearity

In order to closely simulate the real life problem of slope induced distortion, the circuit of Figure 7's input stage op amp was made nonlinear via zener diode feedback. The use of high impedances is again going to reduce the distortion of the input stage at very high frequencies. This should not be of great consequence except at low input levels.

TEST RESULTS and DISCUSSION

Each of the distortion measurement methods described above was used to characterize the five test circuits. The distortion percentages were measured and graphed as a function of input level. Since the circuits are all constant gain devices (no gain compression or expansion) the output level may be obtained by multiplying the horizontal axis by 11. To maintain legibility but allow easy comparison, the eight measurements for each circuit are grouped on two graphs. The 1 kHz THD+N, 20 kHz THD+N, DIM30 and FASTtest total distortion measurements are grouped together and the CCIF 2nd order, CCIF 3rd order, SMPTE 4:1 amplitude ratio and SMPTE 1:1 amplitude ratio measurements are together.

Output Nonlinearity

Figures 8 and 9 show the results for the output nonlinearity test circuit. Below input voltages in the 50 mV to 100 mV range the measurements appear noise limited. Above this region the distortion rises smoothly, reaching a maximum around 500 mV. The correlation between tests is generally good. Most of the traces lie within a narrow range of one another, only the CCIF 2nd order and 20 kHz THD+N deviate significantly. However this deviation is only a scaling difference, the general shape of the curves are the same. The increase in harmonic distortion at 20 kHz over the 1 kHz case is due to the reduced feedback at higher frequencies. The underlying nonlinearity is not frequency dependent. The 2nd order CCIF data is substantially lower than the other because the nonlinearity is symmetrical. If the nonlinearity was totally symmetric there would be no 2nd order distortion

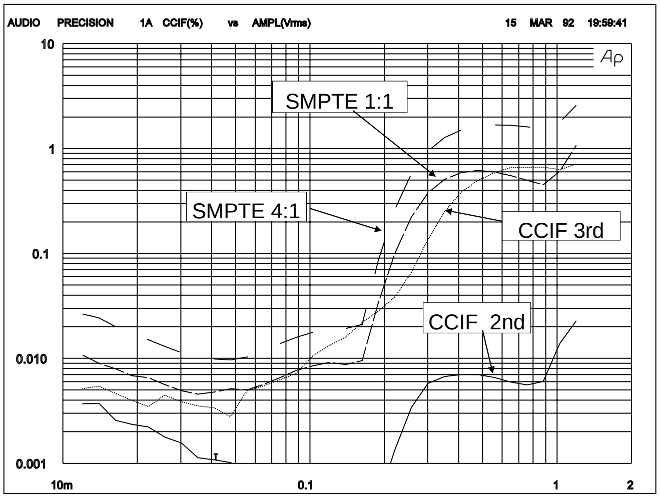

Figure 8 CCIF 2nd order, CCIF 3rd order, SMPTE 4:1, SMPTE 1:1. Output Nonlinearity test circuit

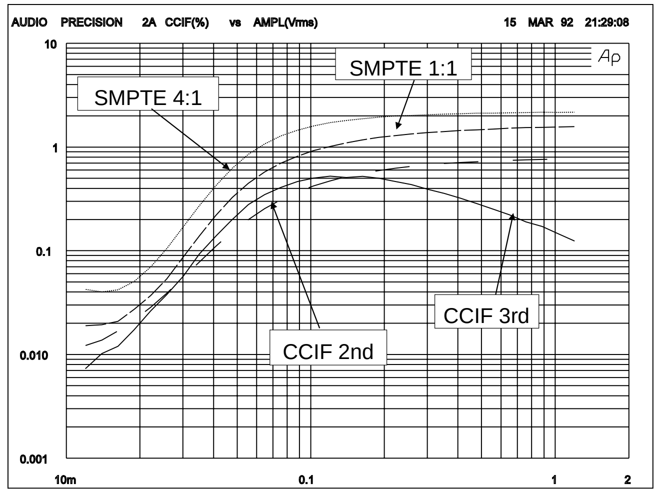

Figure 10 CCIF 2nd Order, CCIF 3rd order, SMPTE 4:1, SMPTE 1:1. Output Stage Assymmetry test circuit

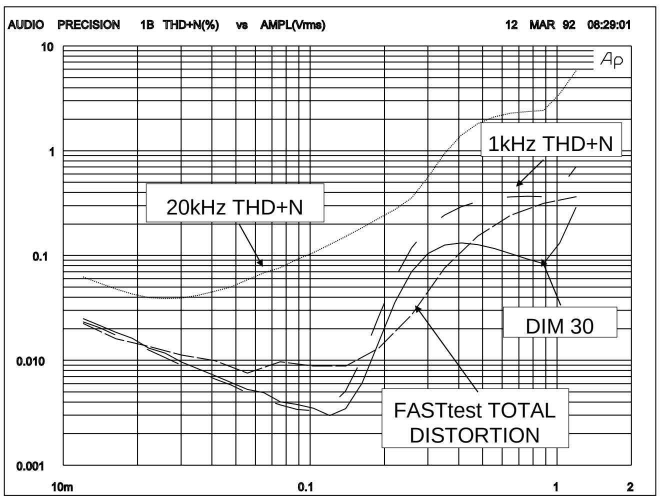

Figure 9 1kHz THD+N, 20 kHz THD+N, DIM30, FASTtest total distortion output nonlinearity test circuit

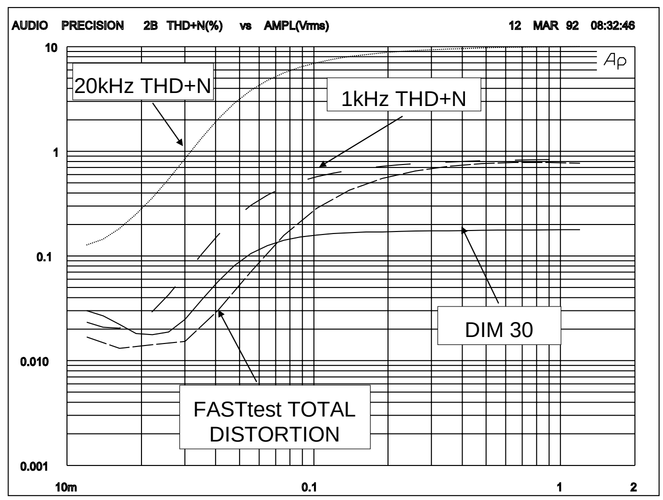

Figure 11 1 kHz THD+N, 20 kHz THD+N, DIM30, FASTtest total distortion. Output Assymmetry test circuit

products at all. The slight asymmetry is due to bias current and offset voltage effects in the op-amps skewing the position of nonlinearity on the transfer curve. The CCIF second order distortion is also low because the distortion components occur at a low frequency where the feedback is high. This enables the amplifier to reject more of the distortion, reducing the sensitivity of the test. The 12dB increase in sensitivity for the SMPTE intermodulation predicted by Warren and Hewlett 1948 does not occur. The DIM test shows a considerable sensitivity to this distortion mechanism as was shown in Leinonen and Otala, indicating that the DIM test is not only a test of slewing or dynamic nonlinearities.

Output Asymmetry

The graphs in Figures 10 and 11 show several interesting points. The similarity of all of the measurements is again apparent. The 20 kHz curve is again significantly above the others since the feedback effect is lower at the higher test frequency. Again, the underlying nonlinearity is not frequency dependent. The even order CCIF test results follow the other results well now since the nonlinearity is inherently asymmetric. The odd order CCIF curves roll over around 100 mV input when the diode in the feedback loop turns on. This is an anomaly of the test circuit and could be reduced or eliminated with a more accurate simulation circuit. The nonlinearity is not purely second order since bias currents and offset voltages in the op amps will slightly turn on the diode. At small outputs the diode appears almost linear, resulting in little distortion. As the output voltage approaches the diode turn on voltage the distortion increases. After the diode is turned on hard, the distortion becomes predominantly second order and third order components drop. The DIM30 test results are significantly lower than the others, although they again show the same shape with level variation.

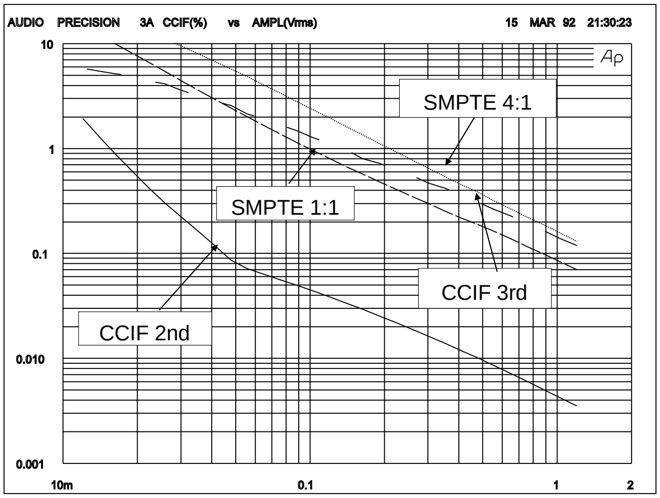

Figure 12 CCIF 2nd order, CCIF 3rd order, SMPTE 4:1, SMPTE 1:1. Crossover Nonlinearity test circuit

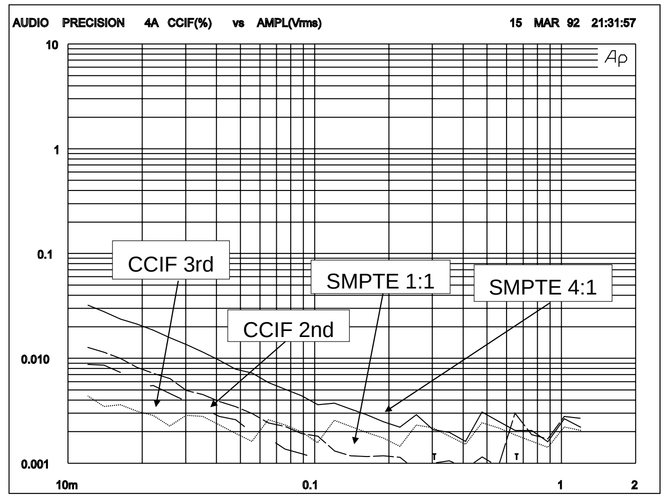

Figure 14 CCIF 2nd order, CCIF 3rd order, SMPTE 4:1, SMPTE 1:1. Hard Slew-Rate Limiting test circuit

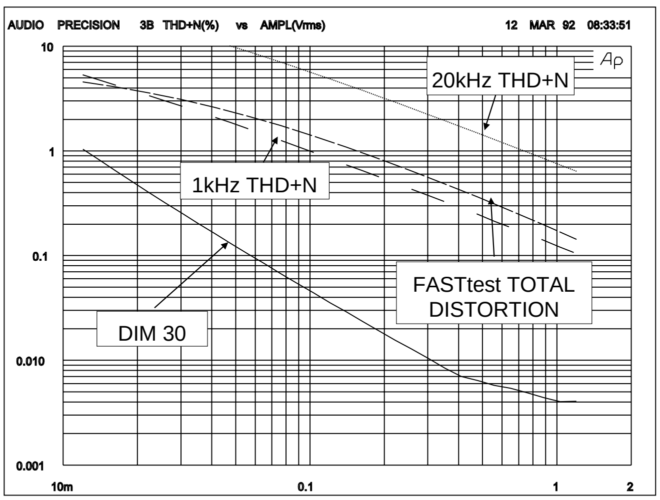

Figure 13 1kHz THD+N, 20 kHz THD+N, DIM30, FASTtest total distortion crossover nonlinearity test circuit

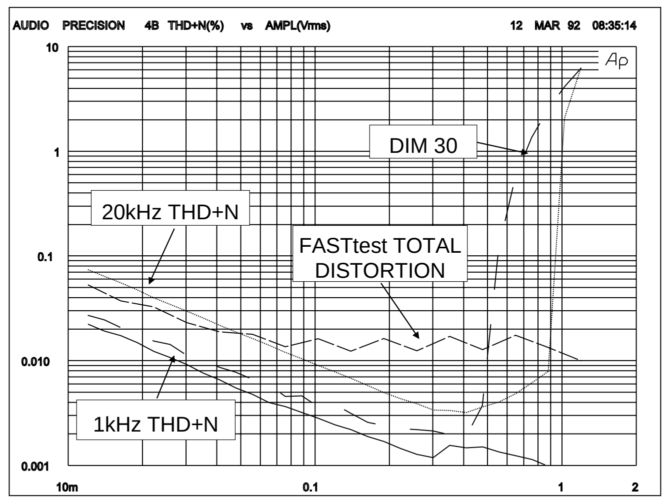

Figure 15 1kHz THD+N, 20 kHz THD+N, DIM30, FASTtest Total Distortion. Hard Slew-Rate Limiting test circuit

Crossover Distortion

Figures 12 and 13 are the results for the case of crossover distortion. Unlike the previous measurements, the distortion decreases with increasing input voltage. This is because the nonlinearity is near the zero transition of the transfer function and is a decreasing percentage of the signal amplitude as the signal amplitude increases. Once again the results have the same shape curve with a scale factor difference between them. As expected from the difference in feedback, the 20 kHz THD+N is substantially larger than the other results. The DIM30 and CCIF second order results are significantly less sensitive than the others, approximately a factor of 30. As seen earlier, the CCIF 2nd order test gives little indication of this distortion mechanism because it measures even order, not odd order nonlinearities. The reduced sensitivity of the DIM30 test is probably because of the small amount of time the signal spends in the crossover region. This is one example where the DIM30 measurement offers a significant insight into the type of nonlinearity, but it is one which can also be obtained by the downward trend of the curves.

The CCIF 2nd and 3rd order measurements show a nonlinear relationship between output level and the distortion decrease. The bias current and offset effects make one transistor bias on sooner than the other. As the level increases the nonlinearity becomes more symmetrical as the second transistor is turned on. As the output voltage rises the distortion also becomes a smaller percentage of the output voltage, resulting in a doubly fast drop in distortion. This did not occur with the other methods since they measure both even and odd order nonlinearities simultaneously.

Hard Slew Limiting

Figures 14 and 15 show the results under hard slew-rate limiting. When the slew rate of the input sinewave reaches the limit of the circuit, the THD rises almost vertically. It is interesting to note that the frequency used for the THD test in Leinonen and Otala was 10kHz. This is the highest frequency which will not excite this effect at any of the voltage levels tested. Hence they were able to conclude that "THD 10 only vaguely indicates the presence of distortion at high output levels." The 20 kHz THD+N test used here clearly illustrates the effect and uses less bandwidth than the DIM30 test signal.

SMPTE intermodulation shows little sensitivity to hard slew limiting. This is due to two factors. With one of the tones at low frequencies there is not as much energy at high frequencies to excite the slew rate limit phenomenon, especially in the 4:1 amplitude ratio version. The second is that the change in slew rate of the composite signal due to the lower frequency tone is very small. The SMPTE test looks for modulation of the upper tone by the lower one. For this to occur in a slew-limiting situation, the low frequency tone must change the slew rate of the upper tone both upwards and downwards as it cycles. This requires as high a lower frequency as possible.

The explanation by Leinonen and Otala for the lack of measurable effect with CCIF testing is wrong. They state that the maximum rate of change of the CCIF method is higher than the DIM method but lasts a shorter time. Actually, the maximum slew rate of the DIM test is considerably higher (see Cabot 1980).

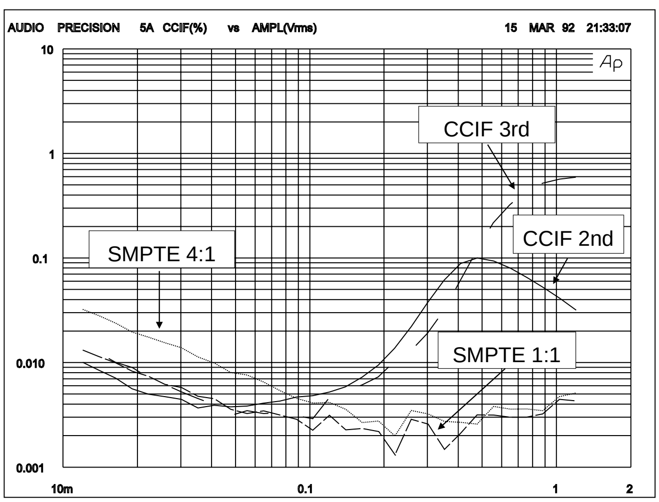

Figure 16 CCIF 2nd order, CCIF 3rd order, SMPTE 4:1, SMPTE 1:1. Input Stage Nonlinearity test circuit

Most practical circuits will slew rate limit in one direction a little sooner than the other. This will cause the second-order component to rise first, then the third-order will rise.

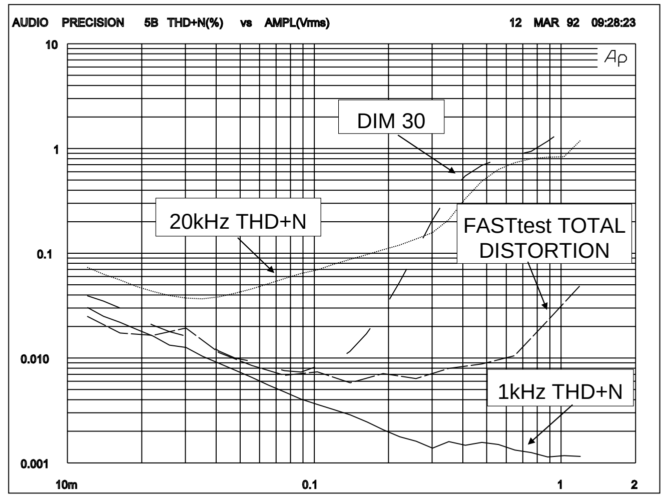

Figure 17 1kHz THD+N, 20 kHz THD+N, DIM30, FASTtest Total Distortion. Input Stage Nonlinearity test circuit

Output Nonlinearity

In the real world, hard slew limiting is rare. The onset of slew or slope-induced distortion is gradual, due to the gradual nonlinearity of the input stage. Figures 16 and 17 show the results on the simulation circuit for this form of distortion.

The 1 kHz THD+N does not show any effect but the 20 kHz case clearly does. The distortion rises when a given slew rate is exceeded but the 1 kHz signal can never exceed this rate at a level within the circuit's range. This rise of distortion with frequency will follow a cubic function as predicted by Jung, Stephens and Todd 1977. The sensitivity of the SMPTE test to this is very low as was the case for hard slew-rate limiting. The second-order CCIF test yields a good sensitivity to input stage nonlinearity, especially at the higher output levels. Since the nonlinearity is symmetric, the third-order CCIF distortion should be quite significant for high level signals. The FASTtest signal shows the distortion effect but the sensitivity is significantly lower. This is because the signal energy is spread across the audio band, reducing the energy available to drive the circuit into slew rate limiting. By reducing the number of low frequency tones in the FASTtest signal or by increasing the amplitude of the high frequency tones the sensitivity may be significantly improved.

CONCLUSIONS

A brief review of some of the complex sources of distortion in audio amplifiers and other audio devices has been presented. This was used as background for a discussion of various techniques for measuring distortion. Some of the advantages of and problems with these techniques were considered.

A new distortion measurement technique which uses a signal more closely resembling program material was presented. It is considerably faster than other distortion measurement techniques and allows simultaneous measurement of frequency response, phase, crosstalk, and noise. It has been shown to correlate well with more established test methods except when hard slew rate limiting must be investigated. For this condition, the energy in the test signal must be shifted toward higher frequencies so as to induce the slew rate limiting behavior. The new technique may easily be adapted to do this by appropriate choice of signal frequencies.

For detailed engineering measurement applications THD is the most flexible and useful technique. Measuring as a function of frequency and level a complete profile of device performance may be determined. For engineering applications where severely bandlimited systems must be characterized, other techniques must be employed. If static nonlinearities are to be measured, or are of special interest, the SMPTE method is quite useful. It is uniquely useful in some applications such as power amplifier output compensation network testing. For testing dynamic nonlinearities in bandlimited systems, the CCIF method (using both 2nd order and 3rd order components) is the most flexible. The DIM sine-square test was shown to be effective as a measurement tool but it is sensitive to more than just "transient" or dynamic distortions.

Anyone seriously interested in designing high performance audio frequency equipment should certainly be equipped to perform a variety of distortion tests. For investigative work each has its applications and can give insight into the device under test. For proof of performance testing or quick checks of proper operation the FASTtest technique is very effective and provides results which are comparable to the other techniques for most devices

.

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to the many people who helped him with this paper. Rush Hood, Barb Cabot, Ron Olisar, Richard Dunipace and Gordon Starosciak helped immensely with the original version of the paper. Bob Wright developed some useful software routines for the total distortion calculation. Special thanks go to Bruce Hofer for the many discussions about the ideas presented in this paper. Wayne Jones did an excellent job in formatting the graphs, text and drawings for the paper and integrating them together.

REFERENCES and BIBLIOGRAPHY

1. Aagaard, James S. - "An Improved Method for the Measurement of Nonlinear Audio Distortion" - IRE Transactions on Audio, November/December 1958.

2. Belcher, R.A. - "Audio Nonlinearity: a comb-filter method for measuring distortion" - BBC Research Department Report #1976/12.

3. Bloch, Alan - "Measurement of Nonlinear Distortion" - Journal of the Audio Engineering Society, January 1953.

4. Borenius, J.H.L. - "Distortion Measurements with Noise as a Test Signal" - AES Preprint 1203, March 1977.

5. Callendar, M.A., and S. Matthews - "Relations Between Amplitudes of Harmonics and Intermodulation Frequencies" - Electronic Engineering, June 1951.

6. Cabot, Richard C. - "A Comparison of Nonlinear Distortion Measurement Methods." - AES Preprint 1638, May 1980.

7. Cabot, Richard C. - "Testing Digital Audio Devices in the Digital Domain." - Proceedings of the AES 7th International Conference, May 1989.

8. Cabot, Richard C. - "Performance Limitations of Digital Filter Architectures" - AES Preprint 2964, September 1990.

9. Cabot, Richard C. - "Fast Response and Distortion Testing" - AES Preprint 3045, February 1991.

10. Cordell, Robert R. - "A Fully In-Band Multitone Test for Transient Intermodulation Distortion." - AES Preprint 1519, November 1979.

11. Crown, IM Analyzer Owners Manual. Elkhart, IN.

12. Feldman, Len - "All About TIM Distortion" - Radio Electronics, June, 1979.

13. Frey, Douglas R. - "A Uniform Approach to Distortion Analysis" - AES Preprint 1521, Nov. 1979.

14. Greiner, Richard A. - "Amplifier Loudspeaker Interfacing" - AES Preprint 1462, May, 1979.

15. Gunthersen, Carsten - "TIM Reconsidered" - Unpublished paper, The Acoustics Laboratory, Technical University of Denmark, 1979.

16. Hilliard, John K. - "Distortion Tests by the Intermodulation Method" - Proceedings of the IRE, December 1941.

17. Hofer, Bruce - "The Fine Art of Audio Power Amplifier Design," presented at the Portland Chapter of the AES, February 1980.

18. Hofer, Bruce - "Practical Extended Range DIM Measurements" - AES Preprint 2334, March 1986.

19. Jensen, Deane and Gary Sokolich - "Spectral Contamination Measurement" - AES Preprint 2725, November 1988.

20. Jung, Walt - "Slew rate tests for audio distortion" - Broadcast Engineering, August 1977.

21. Jung, Walt, Mark L. Stephens and Craig C. Todd - "Slewing Induced Distortion and Its Effect on Audio Am-

plifier Performance—With Results" - AES Preprint 1252, May 1977.

22. Jung, Walter G. and Richard Marsh - "Picking Capacitors" - Audio, February 1980.

23. Leach, W. Marshall Jr. - "Build a Double Barreled Amplifier" - Audio, April, 1980.

24. Le Bel, C.J. - "Measuring and Analyzing Intermodulation" - Audio Engineering, July, 1951.

25. Le Bel, C.J. - "An Experimental Study of Distortion" - Journal of the Audio Engineering Society, October 1954.

26. Leinonen, Eero and Matti Otala - "Correlation Audio Distortion Measurements" - Journal of the Audio Engineering Society, January/February 1978.

27. Leinonen, Eero, Matti Otala and John Curl - "Method for Measuring Transient Intermodulation Distortion (TIM)" - AES Preprint 1185, October 1976.

28. Maxwell, Donald E. - "Comparative Study of Methods for Measuring Nonlinear Distortion in Broadcasting Audio Facilities" - Journal of the Audio Engineering Society January 1953.

29. Moller, Henning - "Multidimensional Audio: Part 2" - Journal of the Audio Engineering Society, June, 1979.

30. Muraoka, Teruo and Kikuji Wagatsuma and Masami Yamazaki - "Distortion Measurements by Multiple-Tone Method" - AES Preprint 1396, November 1978.

31. Otala, Matti and Eero Leinonen - "The Theory of Transient Intermodulation Distortion" - IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-25, #1, February 1977.

32. Read, G.W. and R. R. Scoville - "An Improved Intermodulation Measuring System" - Journal of the SMPTE, February 1948.

33. Roys, H.E. - "Intermodulation Distortion Analysis as Applied to Disk Recording and Reproducing Equipment" - Proceedings of the IRE, October 1947.

34. Roys, H.E. - "Distortion in Phonograph Reproduction" - Journal of the Audio Engineering Society January 1953.

35. Scott, H.H. - "Audible Audio Distortion" - Electronics, January 1945.

36. Shorter, D.E.L. - "The Influence of High Order Products on Nonlinear Distortion" - Electronic Engineering 1950 (vol. 22, #4).

37. Skritek, P. and H. Pichler - "Extended Application of TIM Test Procedures" - AES Preprint 1557, November 1979.

38. Skritek, P. and H. Pichler - "A Simple Method for Computation and Detection of Static and Transient Intermodulation Distortions" AES Preprint 1425, November 1978.

39. Stanley, Gerald and Dave McLaughlin - "Transient Intermodulation Distortion and Measurement" - AES Preprint 1308, November 1977.

40. Stephani, Otfried and Bjorn Bluthgen - "An Investigation into the Increase of Nonlinear Distortion Products from Virgin Tape to Disc Playback" - AES Preprint 1453, March 1979.

41. Schrock, Clifford - "The Tektronix Cookbook of Standard Audio Tests" - Tektronix, Inc., 1975.

42. Schrock, Clifford - "AM Broadcast Measurements Using the Spectrum Analyzer" - Tektronix, Inc. 1976.

43. Takahaski, Susumu and Susumu Tanaka - "A New Method of Measuring Transient Intermodulation Distortion: A Comparison with the Conventional Method" - AES Preprint 1539, November 1979.

44. Thomsen, Carsten and Henning Moller - "Swept Measurements of harmonic, difference-frequency and intermodulation distortion" - B & K application notes #15-098.

45. Waddington, D.E.O'N - "Intermodulation Distortion Measurement" Journal of the Audio Engineering Society July, 1964.

46. Warren, W.J. and W. R. Hewlett - "An Analysis of the Intermodulation Method of Distortion Measurement" - Proceedings of the IRE, April 1948.