A Distortion Primer—Part 1

LARRY KLEIN

In any compilation of the real (and imaginary) problems troubling audio reproduction, distortion would rank right up there near the top of the list. Almost everyone agrees that distortion is *not* a good thing. But beyond that basic point, arguments start. Exactly what's the problem? Simply this: Electronic audio distortion, while easy to measure in its various manifestations, is devilishly difficult to correlate with the perceptions of the human ear/brain apparatus. The situation is further complicated by some manufacturers of expensive audio equipment, accessories, and connecting wires, who are pleased to invent wonderfully esoteric distortion problems (with accompanying voodoo solutions) to satisfy the needs of the devout tweaks and techno-crazies.

Terminological confusion

My *Illustrated Encyclopedia of Electronics* tells me that "Distortion is any change in a signal that alters its basic waveform or the relationship between its various frequency components." Some of the misunderstanding about standard distortions and their audibility arises from ambiguities in terminology. For example, sometimes the technical name for a distortion describes the way the afflicted waveform looks on a scope (e.g., clipping distortion when the tops and/or bottoms of a waveform are decapitated); other times the name refers to the electronic flaw in the amplifier that produces the problem (e.g., crossover distortion).

The terms *harmonic distortion* (HD) and *intermodulation distortion* (IMD) in effect describe kinds of test procedures rather than specific flaws in the equipment under test. If an amplifier has a problem, the same condition should show up on both HD and IMD distortion tests—and provide entirely different measurement numbers. Keep in mind that the numbers provided by distortion-testing instruments are somewhat arbitrary; they depend as much on the type of test and the specifics of the test signal used as on the magnitude of the flaw in the amplifier. And for perhaps the same reason, none of the distortion-measurement numbers correlate directly with audible unpleasantness—or with each other. In other words, 2% distortion does not necessarily sound twice as bad as 1%, or even necessarily worse than 0.5%.

Harmonic distortion

In any discussion of harmonic distortion, keep in mind the distinction between the natural harmonics that are a part of all tones produced by musical instruments and the undesired spurious harmonics that result from flawed amplification. It is the natural harmonic content that causes the same musical note played on a clarinet, a piano, and a flute to sound different—and to look different on an oscilloscope. Any complex waveform can be "discussed" by a mathematical process known as a Fourier analysis and shown to be composed of a large number of odd and even harmonics. Figure 1 shows a violin note and its second, third, fourth, fifth, sixth, and eighth harmonic components. With the proper instrumentation, it is possible to detect harmonics as high as the twentieth.

HD comes about *not* through distortion of the harmonics of a signal, nor does it result from spurious harmonic frequencies generated by an oscillating amplifier. What happens is that the amplifier, because of some technical inadequacy, changes (distorts) the original shape of the signal waveform. That change can be quantified by analyzing it in terms of the spurious harmonics added to the fundamental test signal—which is, in general, the way the ear hears it.

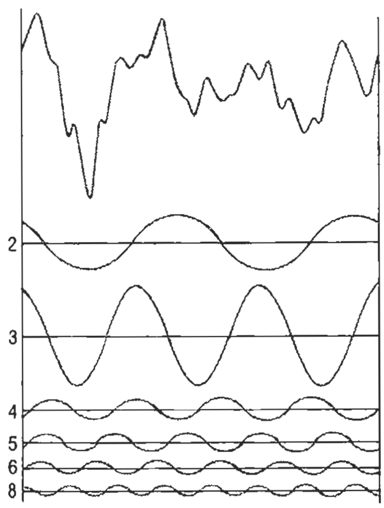

FIG. 1—FOURIER ANALYSIS of a violin note showing the relative strengths of the strongest natural harmonics.

When testing an amplifier's HD performance, you feed in as distortionless a sine wave as can be generated. The HD analyzer, which is connected across the amplifier's output test load, operates by nulling out the input test signal and reading (as a percentage of it) whatever harmonics and noise are introduced by the amplifier. If, say, a 3-kHz test signal was used, amplifier nonlinearity might produce spurious harmonics at 6 kHz, 9 kHz, and so forth. The term *THD* indicates that the lumped *total* of all the harmonic components is included in the measurement. A more sophisticated instrument, called a spectrum analyzer, is capable of indicating the relative strengths of *each* of the spurious harmonics. It is recommended by the EIA Amplifier standard (RS-490) and is used by many test labs.

To illustrate the mechanisms involved, an exaggerated example of distortion is shown in Fig. 2. Let us say that a malfunction of the amplifier causes third-harmonic distortion of waveform (*a*), a 1000-Hz sine-wave input signal. The distorted output signal (*c*) would look as though a 3000-Hz tone (*b*) were combined with the 1000-Hz tone. Keep in mind that a distorting amplifier does not actually generate spurious harmonic waveforms and mix them with the original wave; what it does is distort the original waveform in such a way that the output waveform looks as it would if specific spurious harmonics were added. Of course, in real life we would have not only third-harmonic distortion but also an assortment of various odd and even harmonics of various strengths.

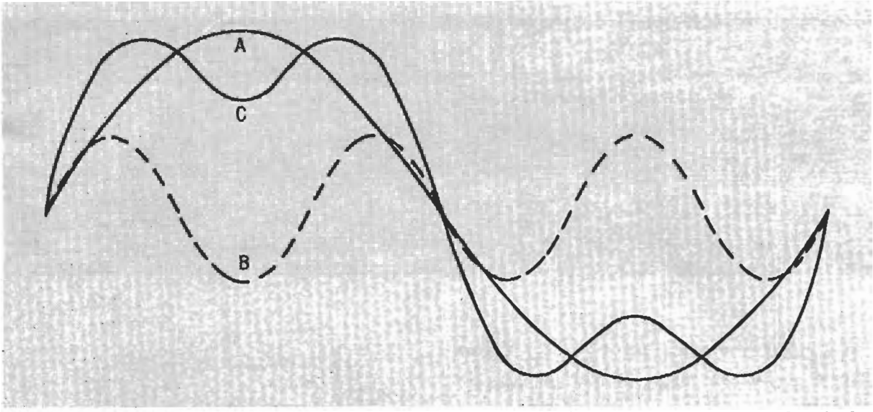

FIG. 2—SIMPLIFIED ILLUSTRATIONS of how a spurious third harmonic (*b*) combined with the input signal (*a*) produces a distorted signal (*c*).

There's some evidence that the specific HD content of a distorted signal (meaning the relative strengths of the distortion components extending up to the tenth harmonic or higher) is more audibly significant than the absolute THD figure. In other words, depending upon the ways that two amplifiers are distorting a piece of music, a measured 3% THD from one might sound a lot worse than 3% from the other.

Intermodulation distortion

The same amplifier nonlinearities that produce THD also produce intermodulation distortion (IMD), but through a somewhat different mechanism. When two (or more) signals are fed through a nonlinear amplifier, the signals tend to *intermodulate*, meaning that they interact in a specific and undesirable way. If, for example, a low-frequency signal of 40 Hz was traveling through a nonlinear amplifier along with a higher frequency of, say, 2 kHz, spurious sum and difference frequencies that are known as IM products would be produced at 1920, 1960, 2040, 2080 Hz, and so on and so forth.

There are two different IMD test techniques in current use, both employing a pair of test tones applied simultaneously. The older SMPTE (Society of Motion Picture and Television Engineers) IMD test uses a composite 60- and 7000-Hz test signal in a 4:1 ratio, while the IHF-IM test uses two equal-amplitude high-frequency tones. The description of the IHF test incorporated in the current EIA Standard (formerly IHF-A-201 1966) reads as follows: "The percentage of IHF intermodulation distortion (IHF-IM) of a composite signal composed primarily of two relatively high-frequency sinusoidal signals, one having a frequency of f_1 and the other having a frequency of f_2, of equal amplitude, is numerically equal to 100 times the square root of the sum of the squares of the second- through fifth-order distortion components divided by the square root of the sums of the squares of the amplitudes of the sums of the components at frequencies f_1 and f_2." All of which, I think, helps explain the relative popularity of the SMPTE method over the IHF-IM method.

Unlike THD, IMD distortion components do not have a harmonic relationship with the music and, therefore, can't be heard as part of the music. For that reason, IMD is generally thought of as more audibly unpleasant. However, I would say that, given the very low distortion figures of *all* of today's better standard-brand audio amplifiers, neither THD nor IMD are likely to be audible, *assuming* that the amplifier is working properly and is never driven into overload. And, even under overload clipping conditions, with complex program material such as a loud symphonic work, it is well documented that distortion (of any flavor) has to reach approximately 6% before it becomes audible. That is true because the spurious distortion frequencies are overwhelmed (technically, "psychoacoustically masked") by music occurring at the same frequencies. However, when the test signal is a pure tone, distortion as low as 0.15% can be heard. Probably for all of the above reasons, it seems that few professional testers will bother with IM measurements.

In next month's wrap up on our distortion discussion, we will look at some of the popular "new" distortions and try to place the entire topic in a real-world context. **R-E**

AUDIO UPDATE

A Distortion Primer—Part 2

LARRY KLEIN

The most well-known of the "non-standard" distortions—transient intermodulation (TIM)—achieved prominence in the early 1970's, mostly through the work of Matti Otala, a Finnish engineering researcher. He rediscovered that under certain conditions, amplifiers using high levels of overall output-to-input (or *global*) negative feedback would experience input overload even though the applied signal was theoretically too small to cause such a problem. I say *re*discovered, because the effect had been documented, discussed, and solutions outlined in the early 1950's.

In a nutshell, the TIM story is this: A rapidly changing audio signal—meaning one with high-frequency components—would overdrive a feedback amplifier's input stage, while a signal of the same amplitude, but without high frequencies, would go through without problems. The overload occurred because the amp's input stage parameters assumed operation with the gain reduction of negative feedback, but the feedback signal did not get back to the input fast enough to prevent overload. A basic solution to TIM is to design for sufficient amplifier bandwidth—*before* feedback—to ensure that high-frequency signals can *slew* (travel fast enough) through the amplifier to avoid problems. Today, competent engineers can easily achieve adequate slew rate—and thereby forestall TIM—without really straining very much of their design talents.

Marketing esoteric distortions

If, as I (and others) claim, the distortion problem is essentially trivial in today's amplifiers, why all the technical papers and amplifier advertisements touting recently invented/discovered varieties of distortions and their cures? As far as I can tell, TIM and other similar obscure amplifier problems appeal to two groups: engineering academics seeking scholarly publishing credits and, especially, manufacturers and their advertising agencies.

Audio manufacturers, in an annual effort to differentiate their new products from those of their competitors, regularly discover—and eliminate—previously unrecognized sources of distortion. That reflects the need to ascribe special audible virtues to products that are really very good but, in truth, are no better than those of their competitors. If sales of a manufacturer's latest models can be enhanced by incorporating a newly developed lateral-feedback circuit to eliminate recently discovered problems of side-slip distortion, why not go for it!

In regard to those dedicated listeners who continue to hear problems or special desirable qualities in certain amplifiers, I would be willing to bet an extremely expensive set of diamond-encrusted tweeter cones that what they are hearing, for better or worse, has far more to do with minor frequency-response deviations than with any kind of old, new, or yet-to-be-discovered amplifier distortion mechanism.

Loudspeaker distortion

No one argues that loudspeakers are anywhere as distortion-free as amplifiers. In essence, a loudspeaker system is required to convert the electrical audio waveform supplied by an amplifier into an analogous three-dimensional acoustic waveform. Considering that loudspeakers do their jobs using an assortment of driven diaphragms vibrating in special boxes, the wonder (as someone once said of a chess-playing dog) is not that it does it well, but that it can do it at all!

The basic loudspeaker problem is linearity of transduction. That means that the speaker cone (and the voice coil that drives it) must move in exact accord with the audio input signal. Voice-coil motion constrained by its suspension or operating in a magnetically nonlinear portion of the voice-coil gap will generate large amounts of second- and third-order harmonic distortion. And any voice-coil movement not coupled accurately to the cone, and any cone movement not directly controlled by the voice coil, will distort the sound in some way. It's clear from test data and careful listening that in the last 15 to 20 years driver design and performance have improved dramatically.

Assessing distortion

About 15 years ago, I found myself on a business trip in England visiting the Rank HiFi speaker research and manufacturing facility. An unexpected bonus of my visit was a day-long meeting with Dr. Peter Fryer, who was then deep into an investigation of speaker distortion and all its ramifications. One of his primary objectives was to assess the audibility of the types and levels of distortion commonly produced by loudspeakers.

Since IM distortion, as mentioned last month, is generally thought of as one of the worst culprits in making things sound bad, Fryer tackled it first. A distortion generator was built that could be set to inject a calibrated amount of IM ranging from 0.1% to 10%. A virtually distortionless amplifier and speaker system were designed and built and served as the "test bed" for all of the subsequent experiments.

Much to everyone's surprise, it was necessary to crank up the IM level to 5% or 6% before it became audible on typical complex musical material, rock or classical. With simpler music, such as a solo piano, 2% IM was clearly audible. And, when sine waves were used as the test signals, IM levels of 0.1 percent could be detected under carefully controlled conditions. None of that surprised me, since the findings neatly replicated the results of some amplifier IM tests that I had been involved in four years earlier.

Fryer's next series of tests involved kinds of distortion *not* found in amplifiers. *Delayed resonance* was something that had concerned the British for years, but to my knowledge was never an issue among U.S. speaker designers. Simply described, it is the tendency of parts of a driver's cone assembly to store energy and continue to release it for some milliseconds after the original signal has ceased. It was described to me (you'll have to imagine the British accent) as "the speaker carrying on broadcasting long after the program has finished."

Unlike IM tests, the delayed-resonance research results were not easily summarized in numerical form. A basic finding was that broad low-Q resonances were far more audible than sharper peaks covering narrower bands, probably because the low-Q resonances were activated for a greater proportion of the signal. When the peak is very broad and low (a Q of less than 1), the audible effect is simply an increase in level over the affected portion of the frequency band.

The primary finding of the research was that it behooves the speaker-driver (and speaker-system) designer to eliminate resonances whenever possible, a task that has been significantly facilitated in the past decade by laser analysis of cone movement and modal vibration analysis of speaker-cabinet walls.

Doppler distortion

A distortion that excites partisan bickering among speaker-system designers, *Doppler distortion* occurs when a cone is undergoing large low-frequency excursions while simultaneously reproducing high frequencies. The theory is that the high frequencies reaching the listener's ears will be alternately compressed and stretched by the low-frequency movements of the cone. There is no question that speakers do produce Doppler distortion. The real question is: How audible is it under normal playing conditions?

A Doppler distortion generator was developed using a delay line that could be varied at a rate determined by a low-frequency signal. A total voice-coil movement of more than two inches could be simulated, certainly more than enough to simulate any real-world condition. Fryer summarized the results of days of experiments by saying that, with the possible exception of small full-range speakers (such as are found in portable radios), normal loudspeaker systems used in the home will never produce enough Doppler distortion for it to be audible.

The bottom line

As with amplifiers, most of what we hear going wrong in hi-fi speakers is in the frequency domain. However, straightening out an amplifier's frequency response is duck soup compared to the task facing a speaker designer. If an amplifier's response has a bump or a dip, a few resistors and capacitors will usually flatten it nicely. Speaker frequency response, on the other hand, involves manipulating magnetic, mechanical, and acoustic variables, in addition to the electronics of the crossover network. But improved materials and know-how continue to make the task immeasurably easier. And anyone who has done any comparison listening in the past dozen years knows that speakers continue to improve. **R-E**