Total Harmonic Distortion – Auditory masking

It is common to assess the quality of power amplifiers by measuring *total harmonic distortion* (THD). In the following lines, we will see what this metric is and what it can tell us about an audio device. The relationship between input and output of an audio device should be ideally described by a straight line, figure 1. Audiophiles use the expression "straight *wire with gain*" to describe a faultless amplifier. In real devices however, the characteristic V_{out} - V_{in} is a curve that deviates from the straight line.

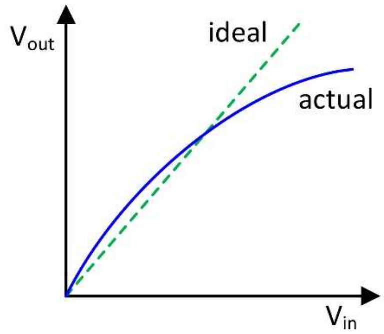

Figure 1. Input-output characteristic

A device with a non-linear input-output characteristic if fed with a tone of frequency f_o will produce at its output the harmonics $2f_o$, $3f_o$, ... in addition to the tone f_o. If H_1 the rms value of the fundamental (rms is approximately 70% of the amplitude) and H_2, H_3, H_4, … the rms value of the harmonics the THD expressed in percentage is calculated by the formula

$$THD = \frac{\sqrt{H_2^2 + H_3^2 + H_4^2 + \cdots}}{H_1} \cdot 100 \ \ (\%)$$

In the numerator we have the sum of all the harmonics representing the distortion products and in the denominator the fundamental which is identical to the input (it only differs in magnitude). Let's see how this works out for an amplifier with a gain of 20 that only produces 2nd and 3rd harmonic at its output. Suppose we feed the amplifier with a 1 kHz tone that has an rms value 0.5 V. Ideally, the output should be a 1 kHz tone with an rms value 0.5·20 = 10 V. But as have assumed the amplifier produces harmonics as well. Suppose that H_2 = 10 mV and H_3 = 5 mV. What is the THD? Using a calculator we do the following calculation

$$THD = \frac{\sqrt{0.01^2 + 0.008^2}}{10} \cdot 100 = 0{,}128\%$$

The above calculation is simple, but it becomes tedious when the amplifier produces many harmonics.

There is another metric related to harmonic distortion called the *Intermodulation Distortion* (IMD). IMD differs from THD in that two frequencies f_1 and f_2 are used as the input. Due to the device non-linearity, in addition to the harmonics $2f_1$, $3f_1$, …, $2f_2$, $3f_2$, …, the output will also contain the sum and difference frequencies. Suppose $f_2 > f_1$, the sum and difference frequencies are

$$f_1 + f_2 \quad f_2 - f_1$$

$$2f_1 + f_2 \quad 2f_1 - f_2 \quad f_1 + 2f_2 \quad f_1 - 2f_2 \quad 2f_1 + 2f_2 \quad 2f_2 - 2f_1$$

and so on. The number of distortion products is theoretically infinite, however in practice only a few components will stand above noise. The IMD is the ratio of the distortion components to the sum of the fundamentals f_1 and f_2. In most cases IMD provides no extra information as the reason it is generated is the same with harmonic distortion, amplifier nonlinearity.

Enough with the definitions and the technicalities. Here comes the important question. How low should the distortion of a power amplifier should be kept so that it is inaudible? Unfortunately, this question does not have a simple answer. In psychoacoustic experiments it has been found that for simple tones we can hear distortion of the order of 0.3%. However, when the reference signal is music the threshold of audibility rises to 1% and in some cases to 3%. Why? Because in the presence of more than one tone a phenomenon called *masking* occurs. In figure 2 a number of curves is shown, each one corresponding to a certain reproduction level L_{CB} of the fundamental, which has a frequency of 1 kHz. The y-axis is the level of the test tone, a second tone played at the same time with the fundamental. Any test tone with a level below the curve is masked, meaning that it is completely inaudible. The dashed line represents the threshold of audibility (see out note about the equal-loudness contours).

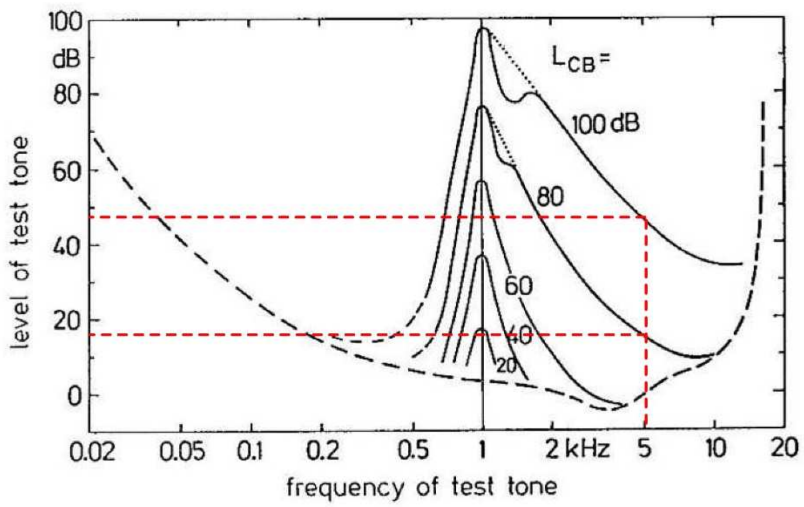

Figure 2. The masking curves for a masker that has a frequency 1 kHz

To understand how masking works take for example the L_{CB} = 100 dB curve and assume that the test tone has a frequency of 5 kHz (5th harmonic). It can be seen that the 5 kHz test tone needs to have a level greater than 47 dB in order to be audible. In other words, the distortion associated with the 5th harmonic needs to be above 0.22%. Next, we move to the 80 dB curve (the fundamental is reproduced at 80 dB). In this case, the limit for the fifth harmonic in order to be audible is 16 dB, or equivalently the distortion must be above 0.06%.

Masking is greater for harmonics closer to the fundamental, meaning that distant harmonics can be heard more easily. This fact explains why we are able to tolerate relatively high amounts of distortion in loudspeakers. The drivers produce high distortion with few harmonics close to the fundamental; these have a high chance to be masked by the fundamental. The masking characteristics of the ear may also explain why class B output stages don't sound so good. This is because a Class B stage, or an underbiased Class AB stage, produces a multitude of harmonics that extent beyond the audio spectrum. Distant harmonics that are high in amplitude are not masked. Opamp manufactures provide a lot of information in the datasheets. There we can see that opamp distortion increases as the output decreases.

From what we have discussed so far this is not good; masking becomes ineffective when the level of the fundamental goes down.

In this short note we have discussed harmonic distortion. We have only scratched the surface though, the way we hear is completely different from the way instruments work. Don't get me wrong, measurements are useful but designing audio equipment is as much an art as it is engineering.

For Echo Diastasis

Spyros Loutridis