Lecture 8: Distortion Metrics

Prof. Ali M. Niknejad

University of California, Berkeley

Copyright © 2005 by Ali M. Niknejad
In general, then, the output waveform is a Fourier series

\[ v_o = \hat{V}_{o1} \cos \omega_1 t + \hat{V}_{o2} \cos 2\omega_1 t + \hat{V}_{o3} \cos 3\omega_1 t + \ldots \]
The fractional second-harmonic distortion is a commonly cited metric

\[ HD_2 = \frac{\text{ampl of second harmonic}}{\text{ampl of fund}} \]

If we assume that the square power dominates the second-harmonic

\[ HD_2 = \frac{a_2 S_1^2}{a_1 S_1} \]

or

\[ HD_2 = \frac{1}{2} \frac{a_2}{a_1} S_1 \]
Third Harmonic Distortion

- The fractional third harmonic distortion is given by

\[ HD_3 = \frac{\text{ampl of third harmonic}}{\text{ampl of fund}} \]

- If we assume that the cubic power dominates the third harmonic

\[ HD_3 = \frac{a_3 S_1^2}{a_1 S_1} \]

or

\[ HD_3 = \frac{1}{4} \frac{a_3}{a_1} S_1^2 \]
Output Referred Harmonic Distortion

In terms of the output signal $S_{om}$, if we again neglect gain expansion/compression, we have $S_{om} = a_1 S_1$

$$HD_2 = \frac{1}{2} \frac{a_2}{a_1^2} S_{om}$$

$$HD_3 = \frac{1}{4} \frac{a_3}{a_1^3} S_{om}^2$$

On a dB scale, the second harmonic increases linearly with a slope of one in terms of the output power whereas the third harmonic increases with a slope of 2.
Signal Power

- Recall that a general memoryless non-linear system will produce an output that can be written in the following form

\[ v_o(t) = \hat{V}_{o1} \cos \omega_1 t + \hat{V}_{o2} \cos 2\omega_1 t + \hat{V}_{o3} \cos 3\omega_1 t + \ldots \]

- By Parseval’s theorem, we know the total power in the signal is related to the power in the harmonics

\[
\int_T v^2(t) dt = \int_T \sum_j \hat{V}_{o j} \cos(j\omega_1 t) \sum_k \hat{V}_{o k} \cos(k\omega_1 t) dt
\]

\[
= \sum_j \sum_k \int_T \hat{V}_{o j} \cos(j\omega_1 t) \hat{V}_{o k} \cos(k\omega_1 t) dt
\]
Power in Distortion

By the orthogonality of the harmonics, we obtain Parseval’s Theorem:

\[
\int_T v^2(t) dt = \sum_j \sum_k \frac{1}{2} \delta_{jk} \hat{V}_{oj} \hat{V}_{ok} = \frac{1}{2} \sum_j \hat{V}_{oj}^2
\]

The power in the distortion relative to the fundamental power is therefore given by

\[
\frac{\text{Power in Distortion}}{\text{Power in Fundamental}} = \frac{V_{o2}^2}{V_{o1}^2} + \frac{V_{o3}^2}{V_{o1}^2} + \cdots
\]

\[
= H D_2^2 + H D_3^2 + H D_4^2 + \cdots
\]
Total Harmonic Distortion

- We define the *Total Harmonic Distortion* \((THD)\) by the following expression

\[
THD = \sqrt{HD_2^2 + HD_3^2 + \cdots}
\]

- Based on the particular application, we specify the maximum tolerable \(THD\)

- Telephone audio can be pretty distorted \((THD < 10\%)\)

- High quality audio is very sensitive \((THD < 1\% \text{ to } THD < .001\%)\)

- Video is also pretty forgiving, \(THD < 5\% \text{ for most applications}\)

- Analog Repeaters \(< .001\%\). RF Amplifiers \(< 0.1\%\)
Intermodulation Distortion

So far we have characterized a non-linear system for a single tone. What if we apply two tones

\[ S_i = S_1 \cos \omega_1 t + S_2 \cos \omega_2 t \]

\[ S_o = a_1 S_i + a_2 S_i^2 + a_3 S_i^3 + \cdots \]

\[ = a_1 S_1 \cos \omega_1 t + a_1 S_2 \cos \omega_2 t + a_3 (S_i)^3 + \cdots \]

The second power term gives

\[ a_2 S_1^2 \cos^2 \omega_1 t + a_2 S_2^2 \cos^2 \omega_2 t + 2a_2 S_1 S_2 \cos \omega_1 t \cos \omega_2 t \]

\[ = a_2 \frac{S_1^2}{2} (\cos 2\omega_1 t + 1) + a_2 \frac{S_2^2}{2} (\cos 2\omega_2 t + 1) + a_2 S_1 S_2 (\cos(\omega_1 + \omega_2) t - \cos(\omega_1 - \omega_2) t) \]
Second Order Intermodulation

- The last term $\cos(\omega_1 \pm \omega_2)t$ is the second-order intermodulation term.

- The intermodulation distortion $IM_2$ is defined when the two input signals have equal amplitude $S_i = S_1 = S_2$

$$IM_2 = \frac{\text{Amp of Intermod}}{\text{Amp of Fund}} = \frac{a_2}{a_1} S_i$$

- Note the relation between $IM_2$ and $HD_2$

$$IM_2 = 2HD_2 = HD_2 + 6\text{dB}$$
Practical Effects of $IM_2$

- This term produces distortion at a lower frequency $\omega_1 - \omega_2$ and at a higher frequency $\omega_1 + \omega_2$.

- Example: Say the receiver bandwidth is from 800MHz – 2.4GHz and two unwanted interfering signals appear at 800MHz and 900MHz.

- Then we see that the second-order distortion will produce distortion at 100MHz and 1.7GHz. Since 1.7GHz is in the receiver band, signals at this frequency will be corrupted by the distortion.

- A weak signal in this band can be “swamped” by the distortion.

- Apparently, a “narrowband” system does not suffer from $IM_2$? Or does it?
Low-IF Receiver

- In a low-IF or direct conversion receiver, the signal is down-converted to a low intermediate frequency $f_{IF}$.

- Since $\omega_1 - \omega_2$ can potentially produce distortion at low frequency, $IM_2$ is very important in such systems.

- Example: A narrowband system has a receiver bandwidth of 1.9GHz - 2.0GHz. A sharp input filter eliminates any interference outside of this band. The IF frequency is 1MHz.

- Imagine two interfering signals appear at $f_1 = 1.910\,\text{GHz}$ and $f_2 = 1.911\,\text{GHz}$. Notice that $f_2 - f_1 = f_{IF}$.

- Thus the output of the amplifier/mixer generates distortion at the IF frequency, potentially disrupting the communication.
Cubic IM

Now let’s consider the output of the cubic term

\[ a_3 s_i^3 = a_3 (S_1 \cos \omega_1 t + S_2 \cos \omega_2 t)^3 \]

Let’s first notice that the first and last term in the expansion are the same as the cubic distortion with a single input

\[ \frac{a_3 S_1^3}{4} (\cos 3\omega_1 t + 3 \cos \omega_1 t) \]

The cross terms look like

\[ \binom{3}{2} a_3 S_1 S_2^2 \cos \omega_1 t \cos^2 \omega_2 t \]
Third Order IM

- Which can be simplified to

\[
3 \cos \omega_1 t \cos^2 \omega_2 t = \frac{3}{2} \cos \omega_1 t (1 + \cos 2\omega_2 t) =
\]

\[
= \frac{3}{2} \cos \omega_1 t + \frac{3}{4} \cos (2\omega_2 \pm \omega_1)
\]

- The interesting term is the intermodulation at \(2\omega_2 \pm \omega_1\)

- By symmetry, then, we also generate a term like

\[
a_3 S_1^2 S_2 \frac{3}{4} \cos (2\omega_1 \pm \omega_2)
\]

- Notice that if \(\omega_1 \approx \omega_2\), then the intermodulation

\[
2\omega_2 - \omega_1 \approx \omega_1
\]
Inband IM3 Distortion

Now we see that even if the system is narrowband, the output of an amplifier can contain in band intermodulation due to $IM_3$.

This is in contrast to $IM_2$ where the frequency of the intermodulation was at a lower and higher frequency. The $IM_3$ frequency can fall in-band for two in-band interferer
Definition of $IM_3$

- We define $IM_3$ in a similar manner for $S_i = S_1 = S_2$

$$IM_3 = \frac{\text{Amp of Third Intermod}}{\text{Amp of Fund}} = \frac{3}{4} \frac{a_3}{a_1} S_i^2$$

- Note the relation between $IM_3$ and $HD_3$

$$IM_3 = 3HD_3 = HD_3 + 10\text{dB}$$
Complete Two-Tone Response

We have so far identified the harmonics and $IM_2$ and $IM_3$ products.

A more detailed analysis shows that an order $n$ non-linearity can produce intermodulation at frequencies $j\omega_1 \pm k\omega_2$ where $j + k = n$.

All tones are spaced by the difference $\omega_2 - \omega_1$. 
Distortion of AM Signals

- Consider a simple AM signal (modulated by a single tone)

\[ s(t) = S_2 (1 + m \cos \omega_m t) \cos \omega_2 t \]

- where the modulation index \( m \leq 1 \). This can be written as

\[ s(t) = S_2 \cos \omega_2 t + \frac{m}{2} \cos(\omega_2 - \omega_m) t + \frac{m}{2} \cos(\omega_2 + \omega_m) t \]

- The first term is the RF carrier and the last terms are the modulation sidebands
Cross Modulation

- Cross modulation occurs in AM systems (e.g. video cable tuners)
- The modulation of a large AM signal transfers to another carrier going thru the same amp

\[ S_i = S_1 \cos \omega_1 t + S_2 (1 + m \cos \omega_m t) \cos \omega_2 t \]

- CM occurs when the output contains a term like

\[ K (1 + \delta \cos \omega_m t) \cos \omega_1 t \]

- Where \( \delta \) is called the transferred modulation index
Cross Modulation (cont)

For \( S_o = a_1 S_i + a_2 S_i^2 + a_3 S_i^3 + \cdots \), the term \( a_2 S_i^2 \) does not produce any CM.

The term
\[
a_3 S_i^3 = \cdots + 3a_3 S_1 \cos \omega_1 t \left( S_2 (1 + m \cos \omega_m t) \cos \omega_2 t \right)^2
\]
is expanded to
\[
= \cdots + 3a_3 S_1 S_2^2 \cos \omega_1 t (1 + 2m \cos \omega_m t + m^2 \cos^2 \omega_m t) \times \frac{1}{2} (1 + \cos 2\omega_2 t)
\]

Grouping terms we have in the output
\[
S_o = \cdots + a_1 S_1 (1 + \frac{3a_3}{a_1} S_2^2 m \cos \omega_m t) \cos \omega_1 t
\]
CM Definition

$CM = \frac{\text{Transferred Modulation Index}}{\text{Incoming Modulation Index}}$

$CM = 3 \frac{a_3}{a_1} S_2^2 = 4 IM_3$

$= IM_3(\text{dB}) + 12 \text{dB}$

$= 12 HD_3 = HD_3(\text{dB}) + 22 \text{dB}$
Consider the CE BJT amplifier shown. The biasing is omitted for clarity.

The output voltage is simply

\[ V_o = V_{CC} - I_C R_C \]

Therefore the distortion is generated by \( I_C \) alone. Recall that

\[ I_C = I_S e^{qV_{BE}/kT} \]
Now assume the input $V_{BE} = v_i + V_Q$, where $V_Q$ is the bias point. The current is therefore given by

$$I_C = I_S e^{\frac{V_Q}{V_T}} e^{\frac{v_i}{V_T}} I_Q$$

Using a Taylor expansion for the exponential

$$e^x = 1 + x + \frac{1}{2!} x^2 + \frac{1}{3!} x^3 + \cdots$$

$$I_C = I_Q \left(1 + \frac{v_i}{V_T} + \frac{1}{2} \left( \frac{v_i}{V_T} \right)^2 + \frac{1}{6} \left( \frac{v_i}{V_T} \right)^3 + \cdots \right)$$
BJT CE Distortion (cont)

- Define the output signal $i_c = I_C - I_Q$

$$i_c = \frac{I_Q}{V_T} v_i + \frac{1}{2} \left( \frac{q}{kT} \right)^2 I_Q v_i^2 + \frac{1}{6} \left( \frac{q}{kT} \right)^3 I_Q v_i^3 + \cdots$$

- Compare to $S_o = a_1 S_i + a_2 S_i^2 + a_3 S_i^3 + \cdots$

$$a_1 = \frac{qI_Q}{kT} = g_m$$

$$a_2 = \frac{1}{2} \left( \frac{q}{kT} \right)^2 I_Q$$

$$a_3 = \frac{1}{6} \left( \frac{q}{kT} \right)^3 I_Q$$
Example: BJT HD2

For any BJT (Si, SiGe, Ge, GaAs), we have the following result

\[ HD_2 = \frac{1}{4} \frac{q \hat{v}_i}{kT} \]

where \( \hat{v}_i \) is the peak value of the input sine voltage

For \( \hat{v}_i = 10 \text{mV} \), \( HD_2 = 0.1 = 10\% \)

We can also express the distortion as a function of the output current swing \( \hat{i}_c \)

\[ HD_2 = \frac{1}{4} \frac{a_2}{a_1^2} S_{om} = \frac{1}{4} \frac{\hat{i}_c}{I_Q} \]

For \( \frac{\hat{i}_c}{I_Q} = 0.4 \), \( HD_2 = 10\% \)
Example: BJT IM3

Let’s see the maximum allowed signal for $IM_3 \leq 1\%$

$$IM_3 = \frac{3}{4} \frac{a_3}{a_1} S_1^2 = \frac{1}{8} \left( \frac{q \hat{v}_i}{kT} \right)^2$$

Solve $\hat{v}_i = 7.3\text{mV}$. That’s a pretty small voltage. For practical applications we’d like to improve the linearity of this amplifier.