MICRO-461
Low-power Radio Design for the IoT

4. Basic Concepts in RF Design

Christian Enz

Integrated Circuits Lab (ICLAB), Institute of Microengineering (IMT), School of Engineering (STI)

Swiss Federal Institute of Technology, Lausanne (EPFL), Switzerland

EPFL

Outline

- **General considerations**

- Nonlinearity and time variance

- Effects of nonlinearity
  - Single tone input
    - Harmonic distortion
    - Gain compression
  - Two-tone input
    - Desensitization and blocking
    - Cross modulation
    - Intermodulation
    - Third-order Intercept Point

- Cascade of nonlinear stages

Units in RF Design

- The **power gain** is usually given in **decibel** and defined as

$$G_{dB} = 10\log\left(\frac{P_{out}}{P_{in}}\right) = 10\log\left(\frac{V_{out}^2}{R_L}\cdot\frac{R_S}{V_{in}^2}\right) = A_{dB} + 10\log\left(\frac{R_S}{R_L}\right)$$

- where A_{dB} is the **voltage gain** expressed in dB

$$A_{dB} = 20\log\left(\frac{V_{out}}{V_{in}}\right)$$

- Note that the power and voltage gain in dB are equal only if the input and output impedances are equal

- The absolute signal levels are often expressed in dBm rather than in watts or volts

- Used for power quantities, the unit dBm refers to "dB's above 1mW." To express the signal power P_{sig} in dBm, we write

$$L_{dBm} = 10\log\left(\frac{P_{sig}}{1mW}\right)$$

Voltage on a 50 Ω Resistor for a 0 dBm Signal Level

- An amplifier senses a sinusoidal signal and delivers a power of 0 dBm to a load resistance of 50 Ω. Determine the peak-to-peak voltage swing across the load

- Since 0 dBm is equivalent to 1mW, for a sinusoidal signal the rms amplitude is given by

$$\frac{V_{rms}^2}{R_L} = 1mW \quad \rightarrow \quad V_{rms} = \sqrt{R_L \cdot 1mW} = 223.6 mV_{rms}$$

- This corresponds to an amplitude of 316.2 mV or a peak-to-peak amplitude of 632.5 mV$_{pp}$

Example of Units in RF

Problem:

A GSM receiver senses a narrowband (modulated) signal having a level of −100dBm. If the front-end amplifier provides a voltage gain of 15 dB, calculate the peak-to-peak voltage swing at the output of the amplifier

Solution:

Since the amplifier output voltage swing is of interest, we first convert the received signal level to voltage

From the previous example, we note that −100dBm is 100 dB below 632 mV$_{pp}$

Also, 100 dB for voltage quantities is equivalent to 10^5. Thus, −100 dBm is equivalent to 6.32 μV$_{pp}$

This input level is amplified by 15 dB (≈ 5.62), resulting in an output swing of 35.5μV$_{pp}$

dBm Used at Interfaces Without Power Transfer
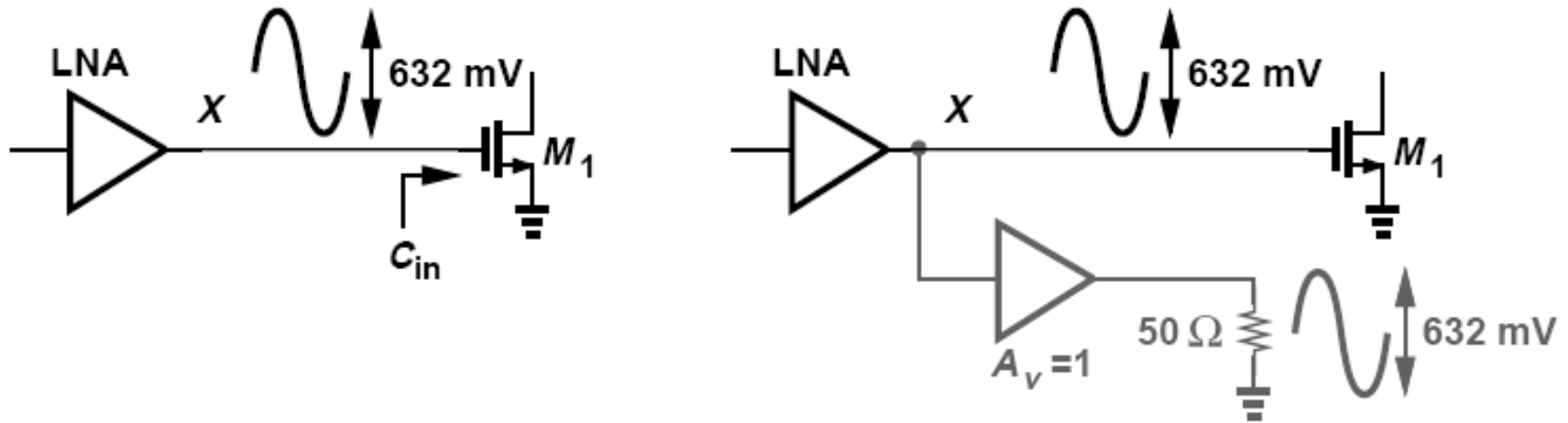
- dBm can be used at interfaces that do not necessarily entail power transfer

- We mentally attach an ideal voltage buffer to node X and drive a 50-Ω load

- We then say that the signal at node X has a level of 0 dBm, tacitly meaning that if this signal were applied to a 50-Ω load, then it would deliver 1 mW

Outline

- General considerations

- **Nonlinearity and time variance**

- Effects of nonlinearity
  - Single tone input
    - Harmonic distortion
    - Gain compression
  - Two-tone input
    - Desensitization and blocking
    - Cross modulation
    - Intermodulation
    - Third-order Intercept Point

- Cascade of nonlinear stages

Nonlinear and Time Variant Systems

- Linear system satisfy the **superposition principle**

$$\left.\begin{array}{c} x_1(t) \rightarrow y_1(t) \\ x_2(t) \rightarrow y_2(t) \end{array}\right\} \Rightarrow ax_1(t) + bx_2(t) \rightarrow ay_1(t) + by_2(t) \quad \text{for all values of } a \text{ and } b$$

- Any system that does not satisfy this condition is **nonlinear**

- A system is **time invariant** if

$$x(t) \rightarrow y(t) \quad \text{then} \quad x(t - \tau) \rightarrow y(t - \tau)$$

- A system is **time variant** if it does not satisfy this condition

- Although intuitively obvious concepts, linear and time invariance may sometimes be confused with each other as illustrated in the next slide

Confusion Between Nonlinear and Time Variant Systems

Nonlinear

time variant

Linear
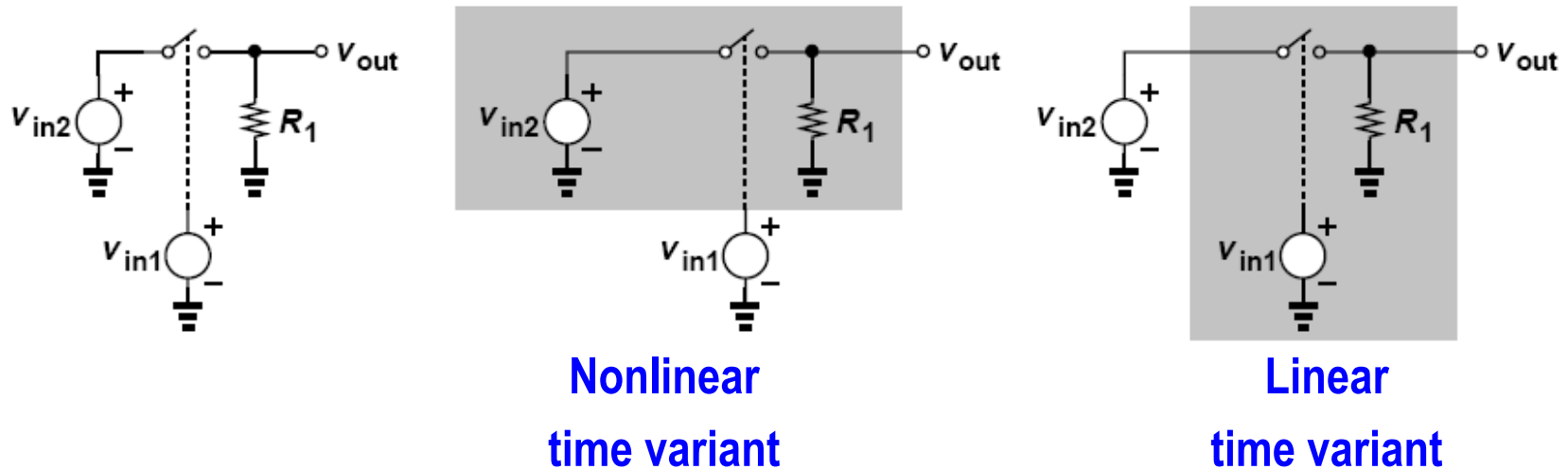
time variant

- Is the following system nonlinear or time variant?

- It actually depends on the input:

  ‣ If V_{in1} is the input and V_{in2} is part of the system then the system is **nonlinear** and **time variant**

  ‣ If V_{in2} is the input and V_{in1} is part of the system then the system is **linear** but **time variant**

Example of Time Variance

- Plot the output waveform of the circuit given in the previous slide if $v_{in1}(t) = A_1 cos(\omega_1 t)$ and $v_{in2}(t) = A_2 cos(1.25\omega_1 t)$
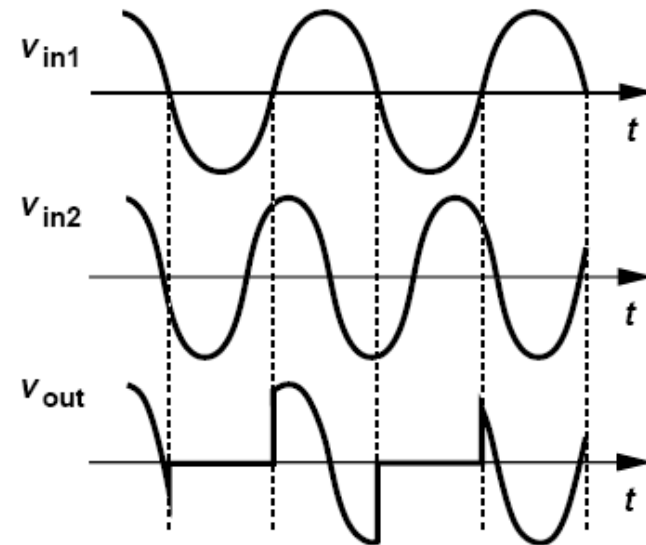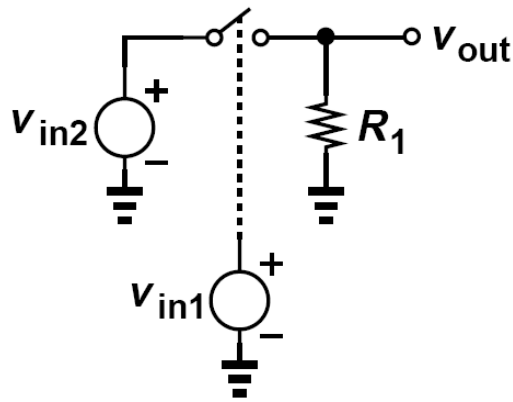
- As shown above, v_{out} tracks $v_{in2} > 0$ and is pulled down to zero by R_1 if $v_{in1} < 0$

- That is, v_{out} is equal to the product of v_{in2} and a square wave toggling between 0 and 1

Linear Time Variant Systems

- A linear system can generate frequency components that are not present in the input signal

- This is possible if the system is **time variant**

- An example is the modulation of a signal by a square wave as shown below
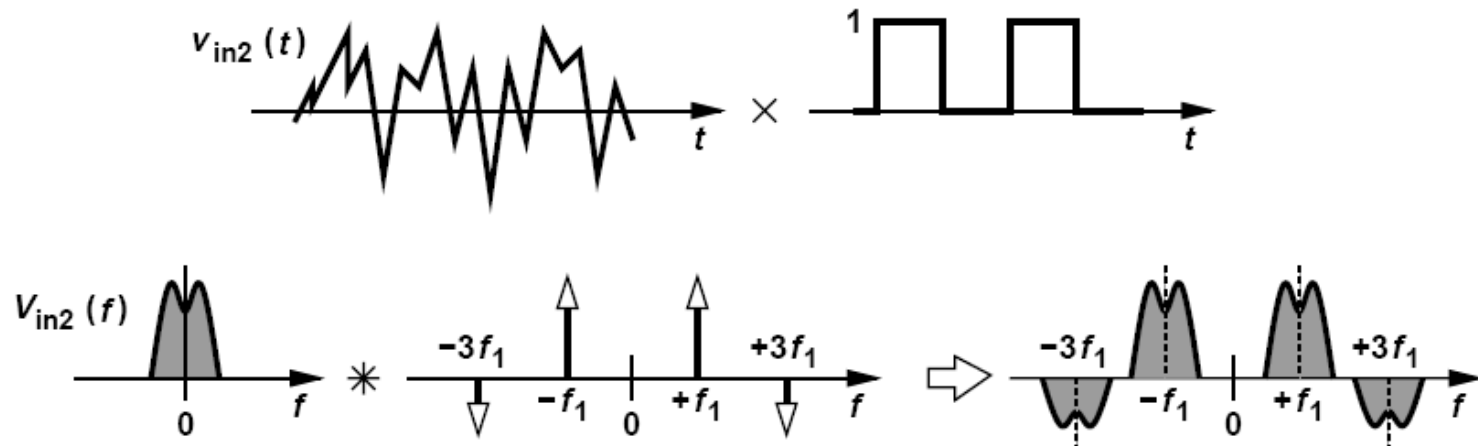
- The output consists of replicas of $V_{in2}(f)$ shifted to n/T_1

$$V_{out}(f) = V_{in2}(f) * \sum_{n=-\infty}^{+\infty} \frac{\sin(n\pi/2)}{n\pi} \delta\left(f - \frac{n}{T_1}\right) = \sum_{n=-\infty}^{+\infty} \frac{\sin(n\pi/2)}{n\pi} V_{in2}\left(f - \frac{n}{T_1}\right)$$

Memoryless System

- A system is called **memoryless** if its output does not depend on the past values of its input

- A **memoryless linear system** is characterized by

$$y(t) = \alpha \cdot x(t)$$

- If α depends on time, the system is **time variant**

- A memoryless **nonlinear** system can be described by a nonlinear input-output function f

$$y(t) = f\big(x(t)\big)$$

- which can often be approximated with a polynomial

$$y(t) = \alpha_0 + \alpha_1\, x(t) + \alpha_2\, x^2(t) + \alpha_3\, x^3(t) + \ldots$$

- where α_j are in general functions of time if the system is time variant

Outline

- General considerations

- Nonlinearity and time variance

- **Effects of nonlinearity**
  - Single tone input
    - Harmonic distortion
    - Gain compression
  - Two-tone input
    - Desensitization and blocking
    - Cross modulation
    - Intermodulation
    - Third-order Intercept Point

- Cascade of nonlinear stages

Effects of Nonlinearity

- Distinction between **single tone** input and **multiple (two) tone** input

- **Single tone** input
  - ▸ Harmonic distortion
  - ▸ Gain compression

- **Two-tone** input
  - ▸ Desensitization and blocking
  - ▸ Cross modulation
  - ▸ Intermodulation

# Single Tone Input – Harmonic Distortion

- If a sinusoid is applied to a nonlinear system, the output exhibits frequency components that are integer multiples of the input frequency

- In the case of a 3rd-order memoryless and time invariant system given by $y(t) = \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t)$ where α_1 is the gain of the linear system, with $x(t) = A cos(\omega t)$, we get

$$y(t) = \alpha_1 A \cos(\omega t) + \alpha_2 A^2 \cos^2(\omega t) + \alpha_3 A^3 \cos^3(\omega t)$$

$$= \alpha_1 A \cos(\omega t) + \frac{\alpha_2 A^2}{2}(1 + \cos(2\omega t)) + \frac{\alpha_3 A^3}{4}(3\cos(\omega t) + \cos(3\omega t))$$

$$= \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4}\right)\cos(\omega t) + \frac{\alpha_2 A^2}{2}\cos(2\omega t) + \frac{\alpha_3 A^3}{4}\cos(3\omega t)$$

- Even order harmonics result from α_j with even j and vanish if the system has odd symmetry (fully differential)

- Amplitude of the nth harmonic consists of terms proportional to A^n and other terms proportional to higher powers of A. Neglecting the latter for small A the nth harmonic can be assumed proportional to A^n

Single Tone Input – Harmonic Distortion

- Output still periodic and can therefore be expressed as a Fourier series
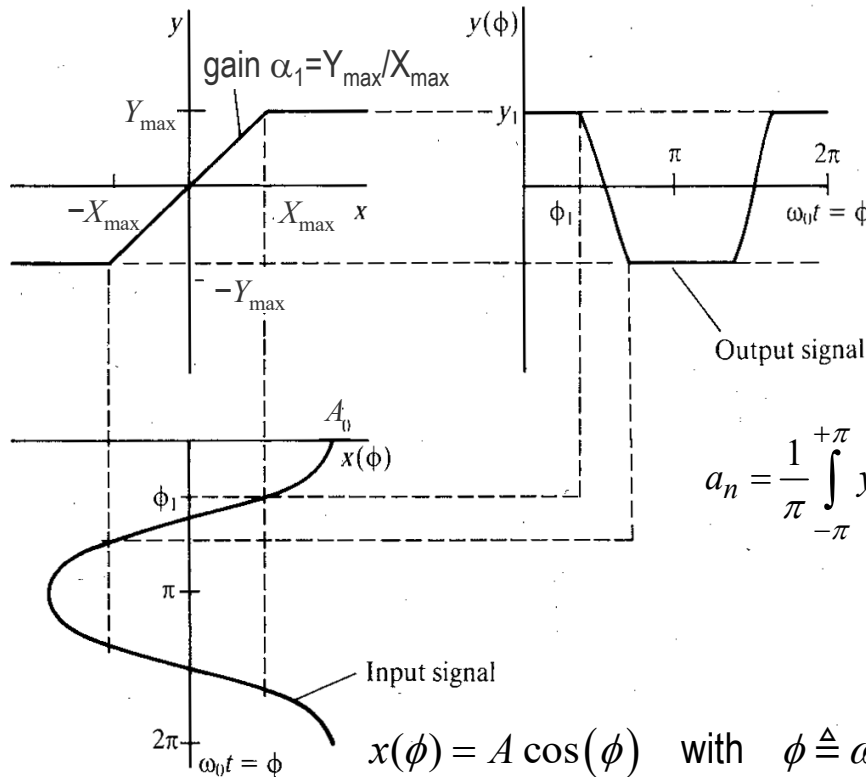
$$y(\phi) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos(n\phi) + b_n \sin(n\phi)$$

$$= \frac{a_0}{2} + \sum_{n=1}^{+\infty} r_n \cos(n\phi - \theta_n)$$

where $\phi \triangleq \omega_0 t$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{+\pi} y(\varphi)\cos(n\varphi)\,d\varphi \quad \text{and} \quad b_n = \frac{1}{\pi} \int_{-\pi}^{+\pi} y(\varphi)\sin(n\varphi)\,d\varphi$$
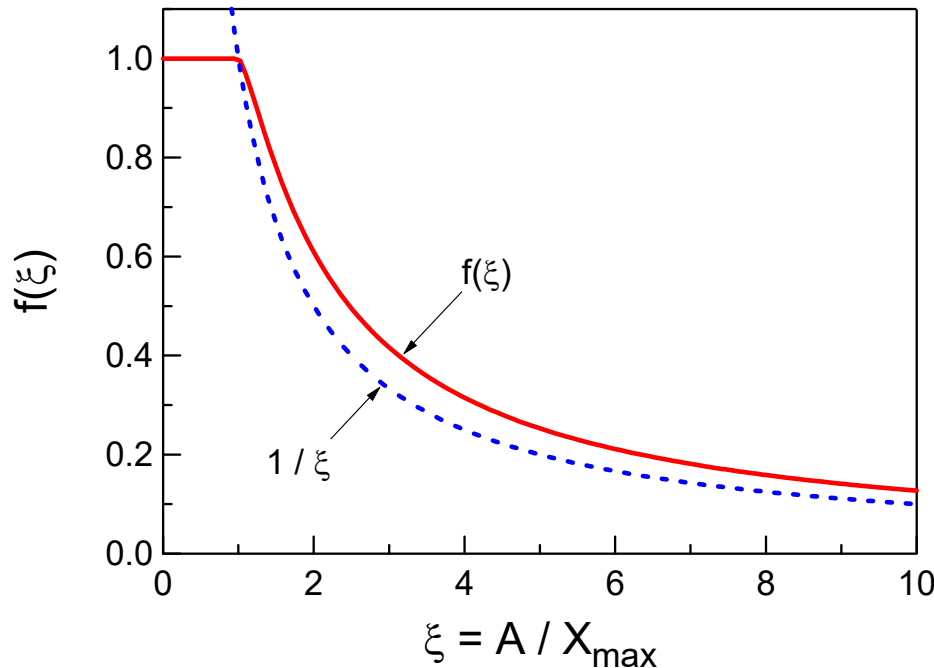
$$r_n = a_n^2 + b_n^2 \quad \text{and} \quad \theta_n = \arctan\left(\frac{b_n}{a_n}\right)$$

$$x(\phi) = A\cos(\phi) \quad \text{with} \quad \phi \triangleq \omega_0 t$$

Output signal

Input signal

gain $\alpha_1 = Y_{max}/X_{max}$

- $a_0 = 0$ since $y(\phi)$ signal has no dc component

- $b_n = 0$ since $y(\phi)$ signal is an even function (i.e. $y(-\phi) = y(\phi)$)

- Only odd-order harmonics since $y(\phi)$ is an odd function (i.e. $y(-\phi) = -y(\phi)$)

Single Tone Input – Gain Compression

- The coefficient a_1 corresponding to the fundamental is given by

$$a_1 = \alpha_1\, A\, f(\xi) \quad \text{with} \quad \xi = \frac{A}{X_{\max}} \quad \text{and} \quad f(\xi) = \begin{cases} 1 & \xi < 1 \\[2mm] \dfrac{2}{\pi}\left[\arcsin\left(1/\xi\right) + \dfrac{\sqrt{1-\left(1/\xi\right)^2}}{\xi} \right] & \xi \geq 1 \end{cases}$$

The compression rate is defined as

$$CR = \frac{\text{actual fundamental output}}{\text{fundamental output if linear}}$$

$$= \frac{a_1}{\alpha_1 A} = f(\xi)$$

$f(\xi)$

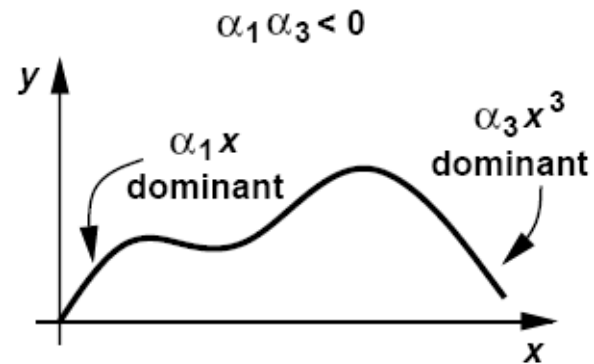$1 / \xi$

$\xi = A / X_{\max}$

Single Tone Input – Sign of α_1 and α_3

$$y(t) = \alpha_1\, x(t) + \alpha_2\, x^2(t) + \alpha_3\, x^3(t) \quad \text{where } \alpha_1 \text{ is the gain of the linear system}$$
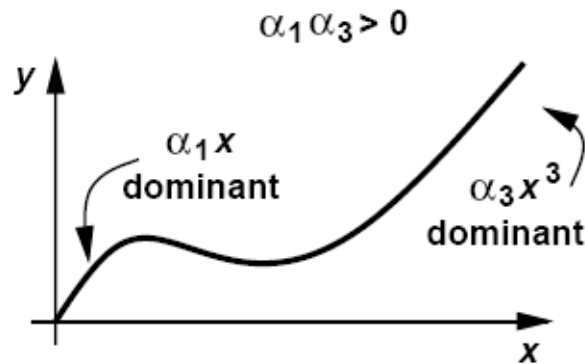
Expansive
(for example MOSFET current)

Compressive
(for example differential pair)

- The nonlinear characteristic can either be **expansive** $(\alpha_1 \cdot \alpha_3 > 0)$ or **compressive** $(\alpha_1 \cdot \alpha_3 < 0)$

- Most voltage RF circuits are compressive whereas current mode circuits may be expansive

Single Tone Input – 1-dB Compression Point

- In most circuits of interest, the output is a compressive (or saturating) function of the input, i.e. the gain approaches zero for sufficiently high input levels. In this case coefficient $\alpha_3 < 0$. The gain for the fundamental is therefore a decreasing function of A

$$G = \frac{Y_1}{A} = \alpha_1 + \frac{3\alpha_3 A^2}{4}$$

- The 1dB compression point is defined as the amplitude for which the gain drops by 1dB compared to the linear asymptote and can be approximated by

$$20\log\left|\alpha_1 + \frac{3}{4}\alpha_3 A_{-1dB}^2\right| = 20\log\left|\alpha_1\right| - 1dB$$
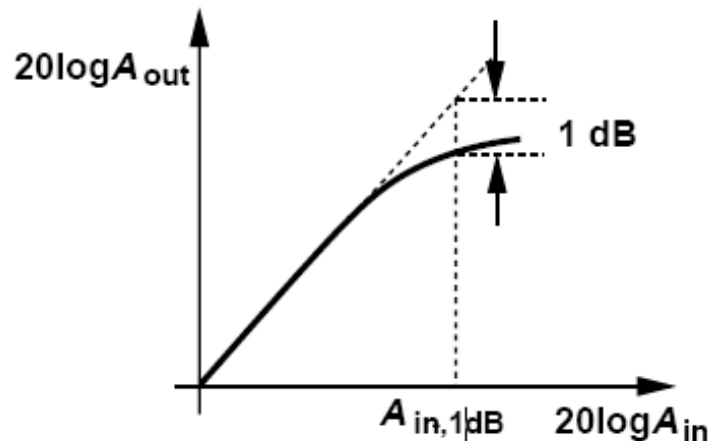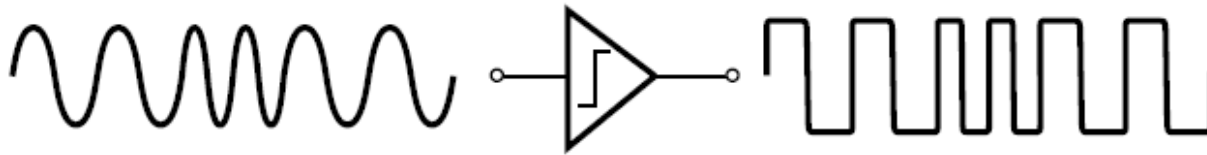

$$A_{-1dB} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}\left(1 - 10^{-\frac{1}{20}}\right)} = \sqrt{0.145\left|\frac{\alpha_1}{\alpha_3}\right|}$$

$$P_{-1dB} = 20\log\left(A_{-1dB}\right) = 10\log\left(\frac{|\alpha_1|}{|\alpha_3|}\right) - 8.4$$
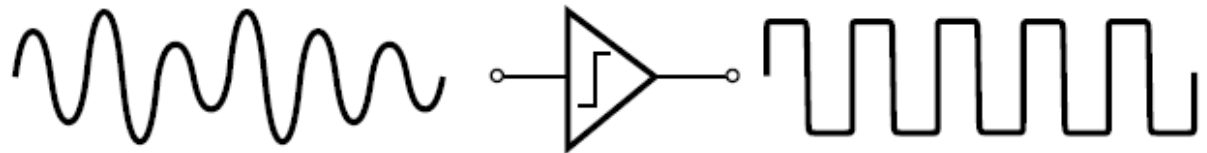
Gain Compression: Effect on FM and AM Waveforms

Frequency Modulation

(a)

Amplitude Modulation

- FM signal carries no information in its amplitude and hence tolerates compression.

- AM contains information in its amplitude, hence distorted by compression

Two-tone Input – Intermodulation Products

- Input signal with two sinusoidal components at two different frequencies ω_1 and ω_2

$$x(t) = A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t) = A_1 \cos(\varphi_1) + A_2 \cos(\varphi_2)$$

- The output of a memoryless and time invariant **nonlinear system** defined by

$$y(t) = \alpha_1 \, x(t) + \alpha_2 \, x^2(t) + \alpha_3 \, x^3(t) + \ldots$$

- The linear term is simply equal to the input multiplied by the gain α_1

$$\alpha_1 \, x(t) = \alpha_1 \, A_1 \cos(\varphi_1) + \alpha_1 \, A_2 \cos(\varphi_2)$$

- The 2$^{\text{nd}}$-order term is given by

$$\alpha_2 \, x^2(t) = \alpha_2 \left[A_1 \cos(\varphi_1) + A_2 \cos(\varphi_2) \right]^2 = \alpha_2 \left[A_1^2 \cos^2(\varphi_1) + 2 A_1 \, A_2 \cos(\varphi_1)\cos(\varphi_2) + A_2^2 \cos^2(\varphi_2) \right]$$

$$= \frac{\alpha_2 A_1^2}{2} + \frac{\alpha_2 A_2^2}{2} \left.\right\} \text{constant dc terms}$$

$$+ \frac{\alpha_2 A_1^2}{2} \cos(2\varphi_1) + \frac{\alpha_2 A_2^2}{2} \cos(2\varphi_2) \left.\right\} 2^{\text{nd}} - \text{harmonics}$$

$$+ \alpha_2 A_1 A_2 \left[\cos(\varphi_1 + \varphi_2) + \cos(\varphi_1 - \varphi_2) \right] \left.\right\} 2^{\text{nd}} - \text{order IM products}$$

Two-tones Input – Intermodulation Products

- The 3rd-order term is given by

$$\alpha_3 x^3(t) = \alpha_3 \left[A_1 \cos(\varphi_1) + A_2 \cos(\varphi_2) \right]^3 = \alpha_3 \begin{bmatrix} A_1^3 \cos^3(\varphi_1) + 3A_1^2 A_2 \cos^2(\varphi_1)\cos(\varphi_2) + \\ +3A_1 A_2^2 \cos(\varphi_1)\cos^2(\varphi_2) + A_2^3 \cos^3(\varphi_2) \end{bmatrix}$$

$$= \left. \frac{3\alpha_3}{2}\left(A_1 A_2^2 + \frac{A_1^3}{2} \right)\cos(\varphi_1) + \frac{3\alpha_3}{2}\left(A_1^2 A_2 + \frac{A_2^3}{2} \right)\cos(\varphi_2) \right\} \text{fundamentals}$$

$$+ \left. \frac{\alpha_3 A_1^3}{4}\cos(3\varphi_1) + \frac{\alpha_3 A_2^3}{4}\cos(3\varphi_2) \right\} 3^{rd} - \text{harmonics}$$

$$\left. \begin{matrix} + \dfrac{3\alpha_3 A_1^2 A_2}{4}\left[\cos(2\varphi_1 + \varphi_2) + \cos(2\varphi_1 - \varphi_2) \right] \\[2mm] + \dfrac{3\alpha_3 A_1 A_2^2}{4}\left[\cos(2\varphi_2 + \varphi_1) + \cos(2\varphi_2 - \varphi_1) \right] \end{matrix} \right\} 3^{rd} - \text{order IM products}$$

- The 3rd-order term produces fundamentals, 3rd-order harmonics and 3rd-order **intermodulation** (IM) products at $2\omega_1 + \omega_2$, $2\omega_1 - \omega_2$, $2\omega_2 + \omega_1$ and $2\omega_2 - \omega_1$

- If ω_1 and ω_2 are close, IM products at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ can be troublesome because they are close to ω_1 and ω_2 and might fall in the passband

Two-tones Input – Desensitization and Blocking

- Assuming that $x_1(t) = A_1 \cos(\omega_1 t)$ is the desired signal and $x_2(t) = A_2 \cos(\omega_2 t)$ is an out-of-band interferer

- Looking only at the fundamental component at ω_1 of the output signal

$$y(t) = \left(\alpha_1 A_1 + \frac{3}{4} \alpha_3 A_1^3 + \frac{3}{2} \alpha_3 A_1 A_2^2 \right) \cos(\varphi_1)$$

- The 3rd term shows that the output signal at the desired frequency is affected by the interferer amplitude A_2

- In the case the interferer is much stronger than the desired signal $(A_1 \ll A_2)$

$$y(t) = \left(\alpha_1 + \frac{3}{2} \alpha_3 A_2^2 \right) A_1 \cos(\varphi_1) = \alpha_1 \left(1 + \frac{3}{2} \frac{\alpha_3}{\alpha_1} A_2^2 \right) A_1 \cos(\varphi_1)$$

- If $\alpha_3 < 0$ (compressive nonlinear characteristic), the amplitude of the desired signal can be reduced by the interferer causing **desensitization**

- A sufficiently large interferer can even reduce the amplitude to zero causing **blocking** (in RF the term blocking signal refers to interferers that can desensitize a circuit causing blocking)

Desensitization and Blocking – Example

- Consider an odd-order memoryless nonlinear system defined by
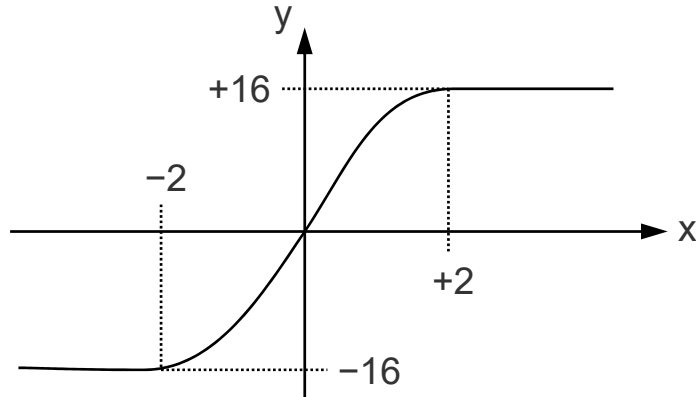
$$y = \begin{cases} \alpha_1 x + \alpha_3 x^3 = 12x - x^3 & |x| \le 2 \\ 16 & |x| > 2 \end{cases}$$

$$x(t) = A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t)$$
$$= A_1 \cos(\varphi_1) + A_2 \cos(\varphi_2)$$

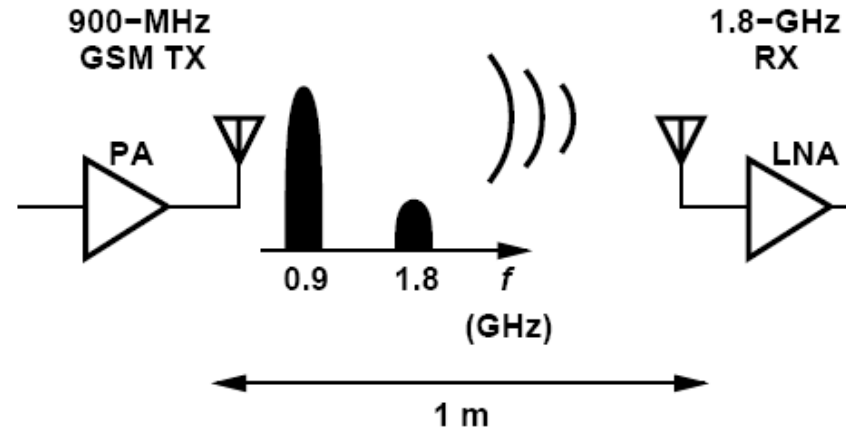- The output signal y_1 at ω_1 for $A_1 < 2$ (no saturation) is given by

$$y_1 = 12 A_1 \left(1 - \frac{A_1^2}{16} - \frac{A_2^2}{8} \right) \cos(\varphi_1) \cong 12 A_1 \left(1 - \frac{A_1^2}{16} \right) \cos(\varphi_1) \quad \text{for } A_2 \ll A_1$$

- In the case the interferer is much stronger than the desired signal ($A_1 \ll A_2$)

$$y_1 = 12 A_1 \left(1 - \frac{A_2^2}{8} \right) \cos(\varphi_1) \quad \text{with } A_1, A_2 < 2 \quad \text{(no saturation)}$$

- Blocking occurs for $A_2 = 2.83$ (not consistent with assumption that A_1 and $A_2 < 2$)

# Example of Gain Compression

Problem:

A 900-MHz GSM transmitter delivers a power of 1 W to the antenna. By how much must the second harmonic of the signal be suppressed (filtered) so that it does not desensitize a 1.8-GHz receiver having P1dB = -25 dBm? Assume the receiver is 1 m away and the 1.8-GHz signal is attenuated by 10 dB as it propagates across this distance.

Solution:

The output power at 900 MHz is equal to +30 dBm. With an attenuation of 10 dB, the second harmonic must not exceed -15 dBm at the transmitter antenna so that it is below P1dB of the receiver. Thus, the second harmonic must remain at least 45 dB below the fundamental at the TX output. In practice, this interference must be another several dB lower to ensure the RX does not compress.

Two-tone Input – Cross Modulation
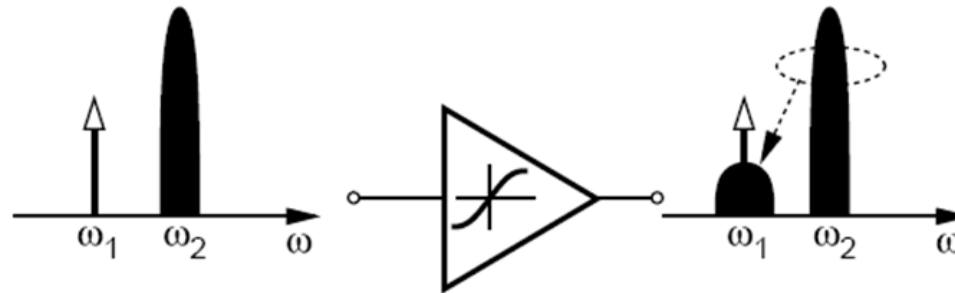
- Another phenomenon that occurs when a weak signal and a strong interferer pass through a nonlinear system is the transfer of modulation (or noise) on the amplitude of the interferer to the amplitude of the weak signal

- For example if the amplitude of the interferer is modulated by a sinusoid $A_2(1 + m \cdot \cos(\omega_m t))$ where m is the modulation index ($m < 1$), assuming $A_1 \ll A_2$, the output is given by

$$y(t) \cong \left[\alpha_1 A_1 + \frac{3}{2} \alpha_3 A_1 A_2^2 \left(1 + m\cos(\omega_m t)\right)^2 \right] \cos(\omega_1 t) + \dots$$

$$= \left[\alpha_1 A_1 + \frac{3}{2} \alpha_3 A_1 A_2^2 \left(1 + \frac{m^2}{2} + \frac{m^2}{2}\cos(2\omega_m t) + 2m\cos(\omega_m t)\right) \right] \cos(\omega_1 t) + \dots$$

Two-tone Input – Cross Modulation

- The desired signal at the output contains amplitude modulation at ω_m and $2\omega_m$

- A common case of cross modulation arises in amplifiers that must simultaneously process many independent signal channels (e.g. in cable television transmitters). Nonlinearities in the amplifier corrupt each signal with the amplitude variations in other channels

Example of Cross Modulation

Problem:

Suppose an interferer contains phase modulation but not amplitude modulation. Does cross modulation occur in this case?

Solution:

Expressing the input as $x(t) = A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t + \phi)$, where the second term represents the interferer (A_2 is constant but ϕ varies with time)

$$y(t) = \alpha_1 \left[A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t + \phi) \right] + \alpha_2 \left[A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t + \phi) \right]^2$$

$$+ \alpha_3 \left[A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t + \phi) \right]^3$$

We now note that (1) the second-order term yields components at $\omega_1 \pm \omega_2$ but not at ω_1; (2) the third-order term expansion gives $3\alpha_3 A_1 \cos(\omega_1 t) A_2^2 cos^2(\omega_2 t + \phi)$, which results in a component at ω_1. Thus,

$$y(t) = \left(\alpha_1 + \frac{3}{2} \alpha_3 A_2^2 \right) A_1 \cos(\omega_1 t) + \dots$$

The desired signal at ω_1 does not experience cross modulation

Two-tones Input – Intermodulation Products

- Discarding the DC and harmonics we obtain the following IM products

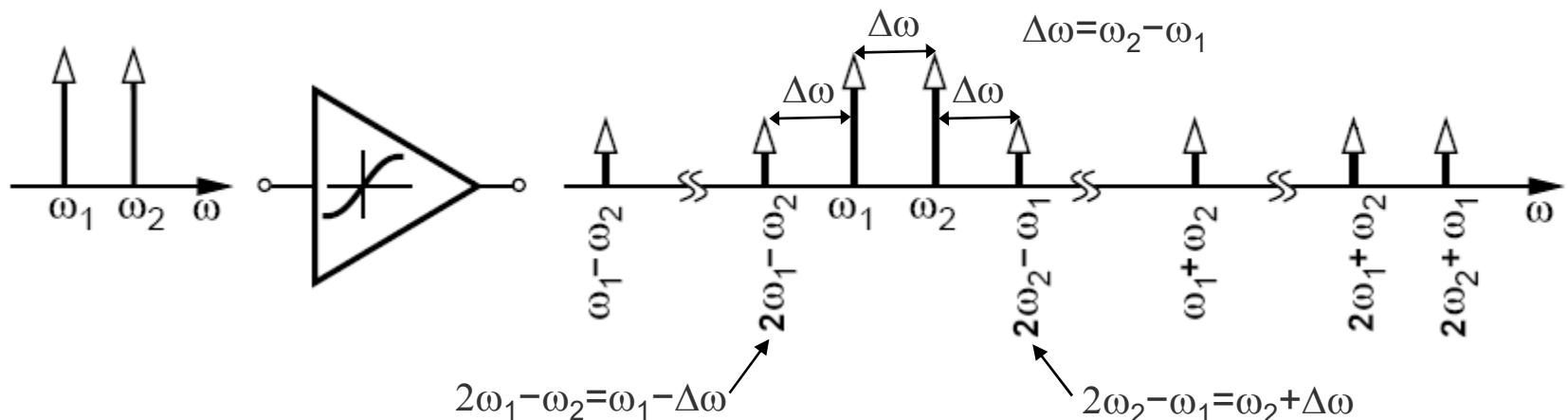$$\varphi_1 \pm \varphi_2 : \quad \alpha_2 A_1 A_2 \big[\cos(\varphi_1 + \varphi_2) + \cos(\varphi_1 - \varphi_2) \big]$$

$$2\varphi_1 \pm \varphi_2 : \quad \frac{3\alpha_3 A_1^2 A_2}{4} \big[\cos(2\varphi_1 + \varphi_2) + \cos(2\varphi_1 - \varphi_2) \big]$$

$$2\varphi_2 \pm \varphi_1 : \quad \frac{3\alpha_3 A_1 A_2^2}{4} \big[\cos(2\varphi_2 + \varphi_1) + \cos(2\varphi_2 - \varphi_1) \big]$$

- And these fundamental components

$$\varphi_1 : \quad \left(\alpha_1 A_1 + \frac{3}{4}\alpha_3 A_1^3 + \frac{3}{2}\alpha_3 A_1 A_2^2 \right) \cos(\varphi_1)$$

$$\varphi_2 : \quad \left(\alpha_1 A_2 + \frac{3}{4}\alpha_3 A_2^3 + \frac{3}{2}\alpha_3 A_2 A_1^2 \right) \cos(\varphi_2)$$

Two-tone Input – Intermodulation

- The 3rd-order IM product of two strong out-of-band interferers may fall in the band of interest corrupting the desired weak signal

$$2\omega_1 - \omega_2 = \omega_0$$
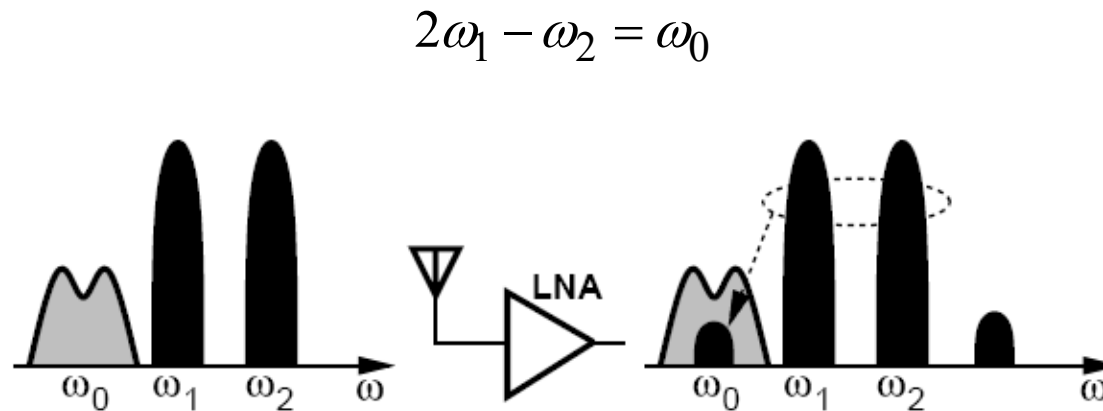
- While operating on the amplitude, this effect degrades the performance even if the modulation is on the phase (zero-crossing points are still affected)

- This phenomenon cannot be quantized by harmonic distortion

- It requires another characterization called **third-order intercept point (IP3)**

Example of Intermodulation
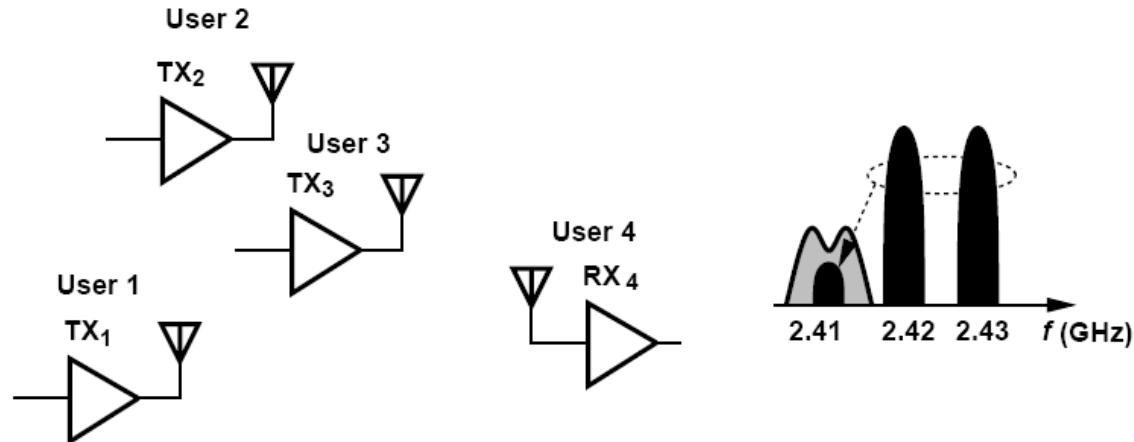
Problem:

Suppose four Bluetooth users operate in a room as shown in figure below. User 4 is in the receive mode and attempts to sense a weak signal transmitted by User 1 at 2.410 GHz. At the same time, Users 2 and 3 transmit at 2.420 GHz and 2.430 GHz, respectively. Explain what happens.

Solution:

Since the frequencies transmitted by Users 1, 2, and 3 happen to be equally spaced, the intermodulation in the LNA of RX4 corrupts the desired signal at 2.410 GHz

Intermodulation: Tones and Modulated Interferers
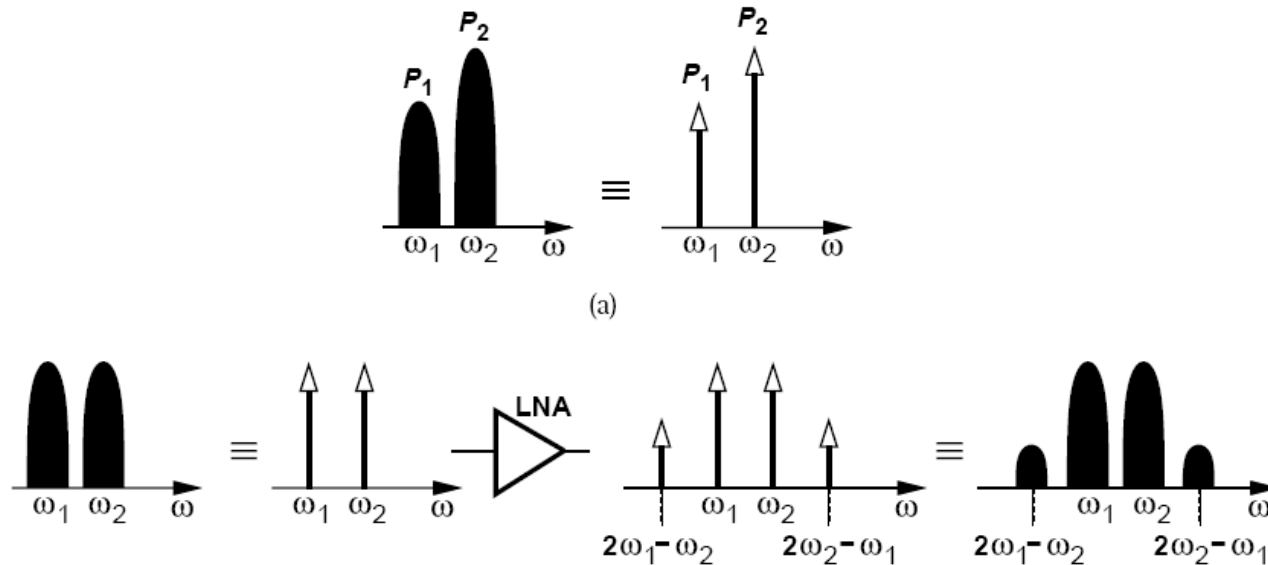
- In intermodulation analysis:
  - ▸ approximate the interferers with tones
  - ▸ calculate the level of intermodulation products at the output
  - ▸ mentally convert the intermodulation tones back to modulated components so as to see the corruption.

Example of Gain Compression and Intermodulation

Problem:

A Bluetooth receiver employs a low-noise amplifier having a gain of 10 and an input impedance of 50 Ω. The LNA senses a desired signal level of -80 dBm at 2.410 GHz and two interferers of equal levels at 2.420 GHz and 2.430 GHz. For simplicity, assume the LNA drives a 50-Ω load.

(a) Determine the value of α_3 that yields a P_{1dB} of -30 dBm.

(b) If each interferer is 10 dB below P1dB, determine the corruption experienced by the desired signal at the LNA output.

Solution:

(a) From previous equation, $\alpha_3 = 14.5\ V^2$

(b) Each interferer has a level of -40 dBm (= 6.32 mV$_{pp}$), we determine the amplitude of the IM product at 2.410 GHz as:

$$\frac{3\alpha_3 A_1^2 A_2}{4} = 0.343 mV_p = -59.3 dBm$$

Two-tone Input – Third-order Intercept Point

▪ Allows to characterize nonlinearity by a two-tone test signal with each tone having equal amplitude $A_1 = A_2 = A$

▪ The amplitude of the fundamental and of the 3rd-order IM products at $2\omega_1 \pm \omega_2$ and at $2\omega_2 \pm \omega_1$ are given by
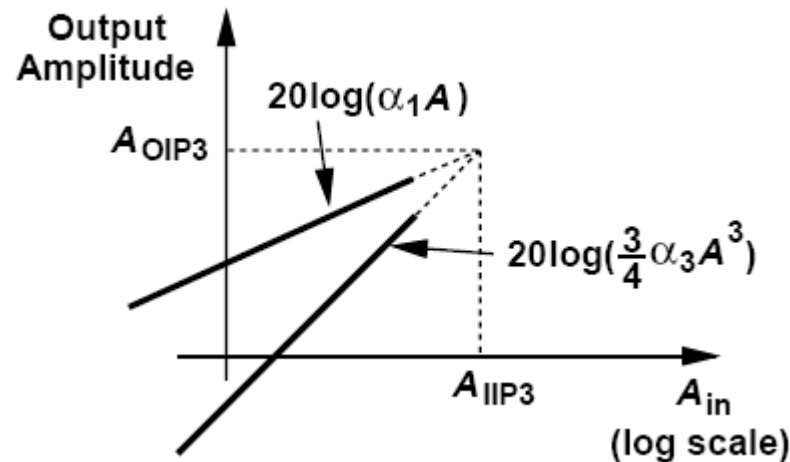
$$IM_1 = \left(\alpha_1 + \frac{9}{4}\alpha_3 A^2 \right) A$$

and

$$IM_3 = \frac{3}{4}\alpha_3 A^3$$

▪ If $\alpha_1 \gg 9\alpha_3 A^2/4$, the input level for which the output components at ω_1 and ω_2 have the same amplitude as those at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ is given by

$$|\alpha_1| A_{IIP3} = \frac{3}{4}|\alpha_3| A_{IIP3}^3 \quad \text{and thus} \quad A_{IIP3} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}}$$

Two-tone Input – Third-order Intercept Point

- IP3 can be measured at a single input level chosen well below the -1dB compression point, considering that

$$\frac{IM_1}{IM_3} \cong \frac{|\alpha_1| A_{in}}{3|\alpha_3| A_{in}^3/4} = \frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}\frac{1}{A_{in}^2} = \frac{A_{IIP3}^2}{A_{in}^2}$$

$$IIP_3\big|_{dBm} = 20\log A_{IIP3} = \frac{1}{2}\underbrace{\left(20\log IM_1 - 20\log IM_3\right)}_{=\Delta P\big|_{dB}} + \underbrace{20\log A_{in}}_{=P_{in}\big|_{dBm}} = \frac{\Delta P\big|_{dB}}{2} + P_{in}\big|_{dBm}$$

where A_{in} is the input amplitude, IM_1 the amplitude of the fundamental at ω_1 and ω_2 and IM_3 the amplitude of the 3rd-order intermodulation products
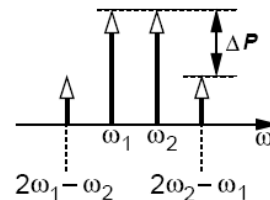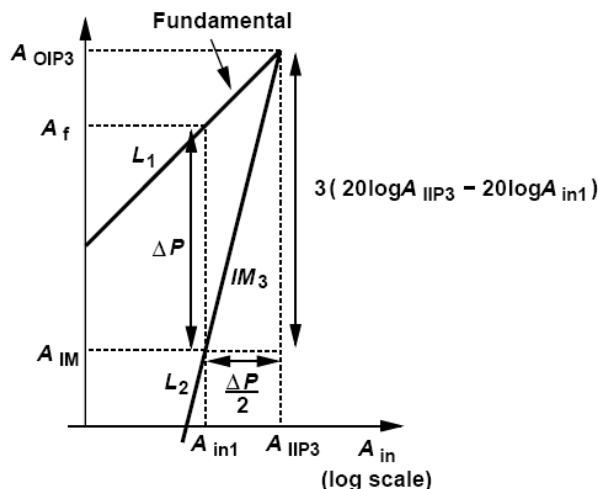

If all signals levels are expressed in dBm, the input 3rd-order intercept point is equal to half the difference between the magnitudes of the fundamentals and the IM3 product at the output plus the corresponding input level

$$IIP_3\big|_{dBm} = \frac{\Delta P\big|_{dB}}{2} + P_{in}\big|_{dBm}$$

Two-tone Input – Third-order Intercept Point
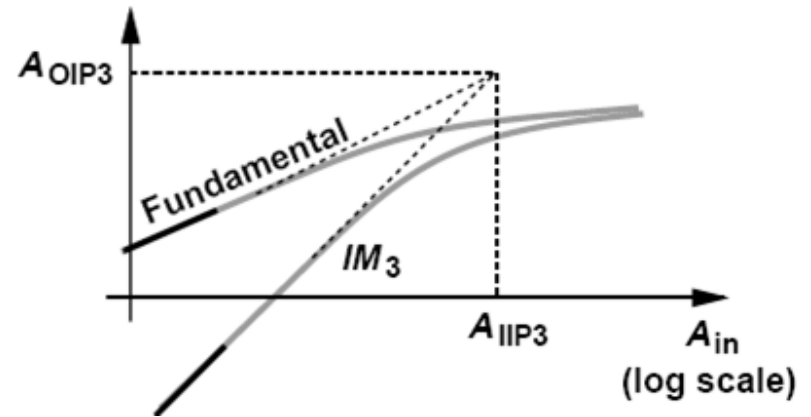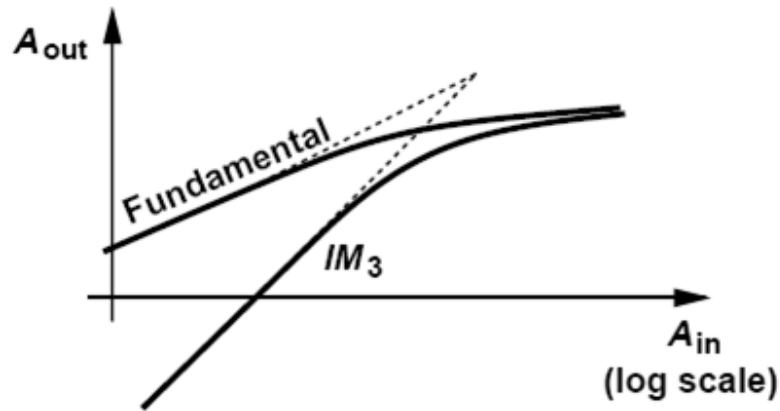
- Because fundamental and IM3 components are both subject to compression, the IP3 is not a directly measureable quantity, but a point obtained by extrapolation of measurements made at low amplitude (as shown on the right)

Comparison Between Single Tone and Two-tone
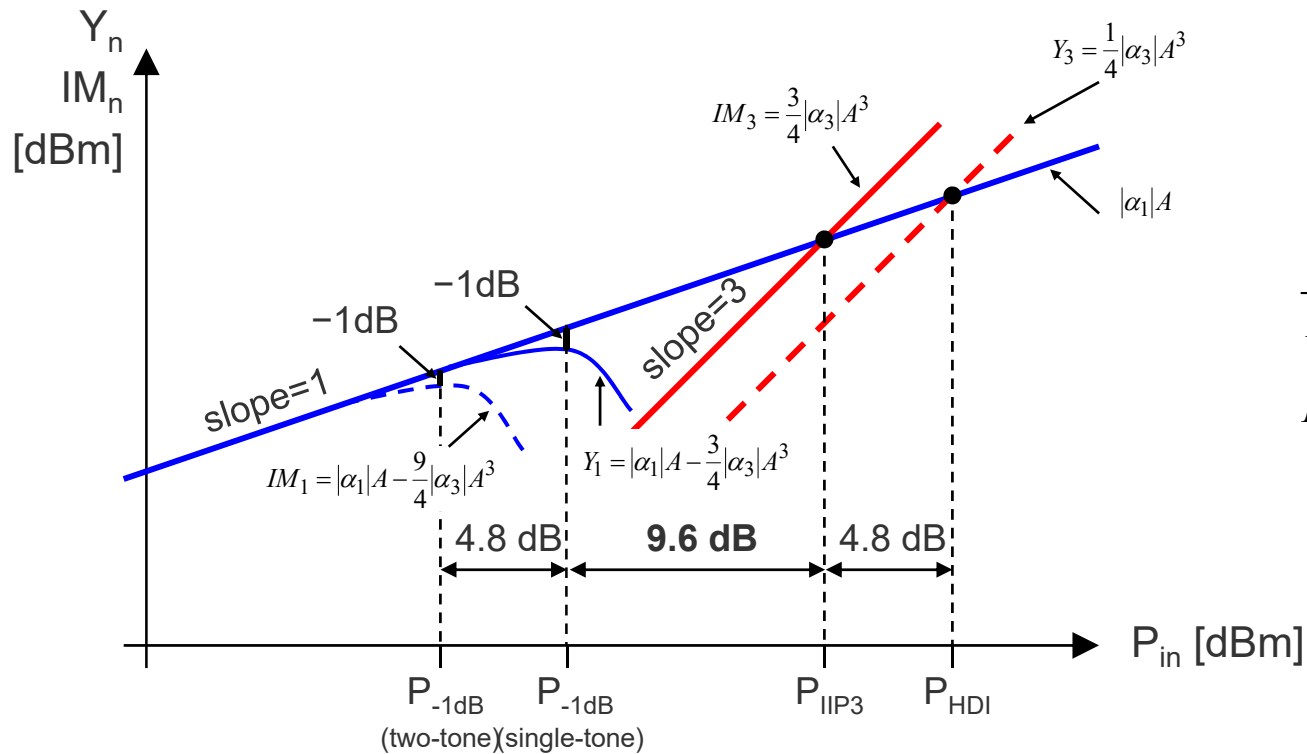
$$\frac{A_{IIP3}}{A_{-1dB}} = \frac{1}{\sqrt{1 - 10^{-\frac{1}{20}}}} \cong 3$$

$$P_{IIP3} - P_{-1dB} \cong 9.6\,dB$$

Single tone −1 dB compression point:

$$A_{-1dB} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}\left(1 - 10^{-\frac{1}{20}}\right)}$$

$$P_{-1dB} = 20\log\left(A_{-1dB}\right) = 10\log\left(\frac{|\alpha_1|}{|\alpha_3|}\right) - 8.4$$

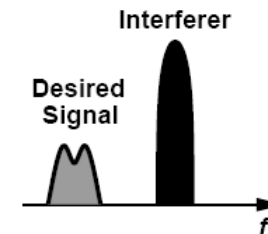Two tones IIP3:

$$A_{IIP3} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}}$$

$$P_{IIP3} = 20\log\left(A_{IIP3}\right) = 10\log\left(\frac{|\alpha_1|}{|\alpha_3|}\right) + 1.25$$

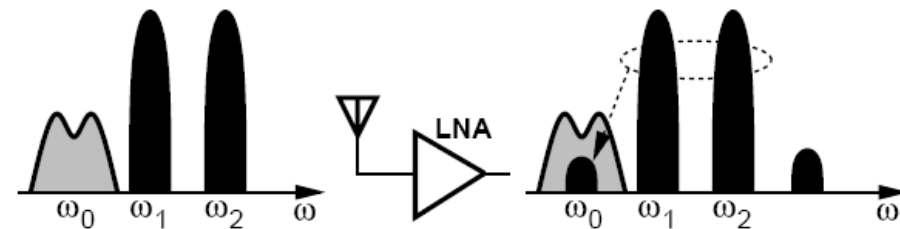# Effects of nonlinearity – Summary

- Single signal leading to **harmonic distortion**
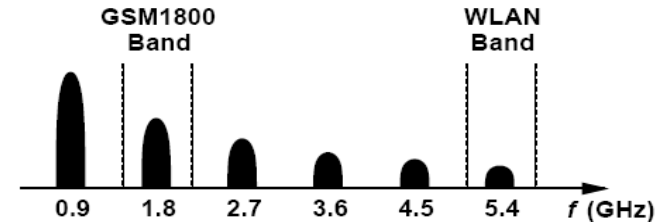
- Single + one large interferer leading to **desensitization**
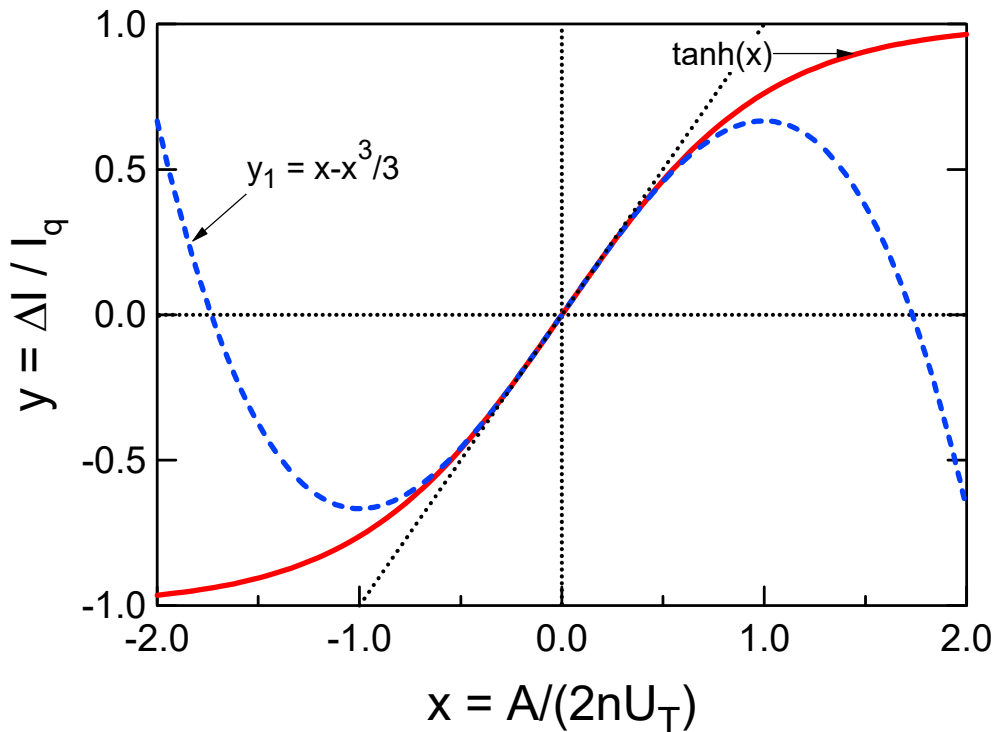

- Single + two large interferers leading to **intermodulation**

Examples: Differential Pair in Weak Inversion

■ The differential output current ΔI_{out} of a MOS differential pair biased in weak inversion (or with $n = 1$ for bipolar) is given by

$$y = \frac{\Delta I_{out}}{I_q} = \tanh\left(\frac{A}{2nU_T}\right) = \tanh(x) \quad \text{with} \quad x = \frac{A}{2nU_T}, U_T = \frac{kT}{q}, n = 1.2\ldots1.6$$
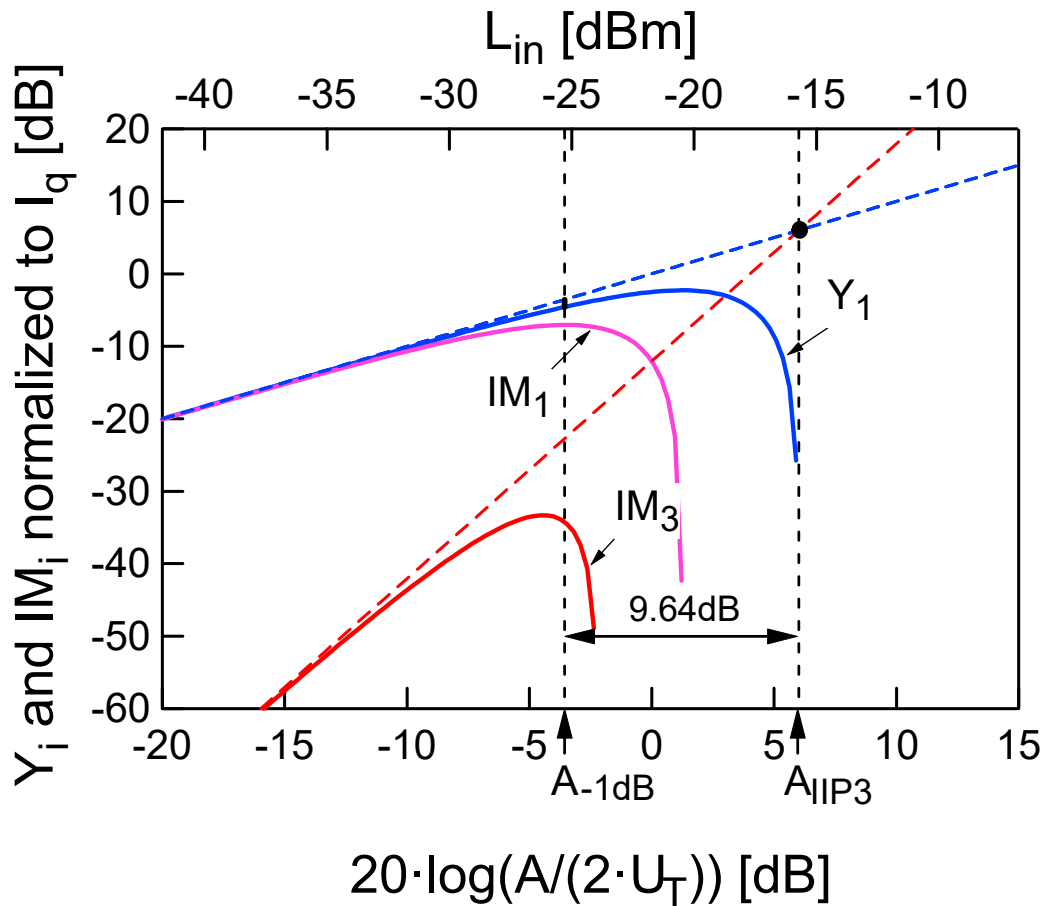

Taylor series expansion:

$$y = \tanh(x) \cong x - \frac{x^3}{3}$$

and hence:

$$\alpha_1 = 1 \quad \text{and} \quad \alpha_3 = -\frac{1}{3}$$

Examples: Differential Pair in WI – Single and Two-tone

$$\alpha_1 = 1 \quad \text{and} \quad \alpha_3 = -\frac{1}{3}$$

$$x_{-1dB} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}\left(1-10^{-\frac{1}{20}}\right)} \cong 0.66$$

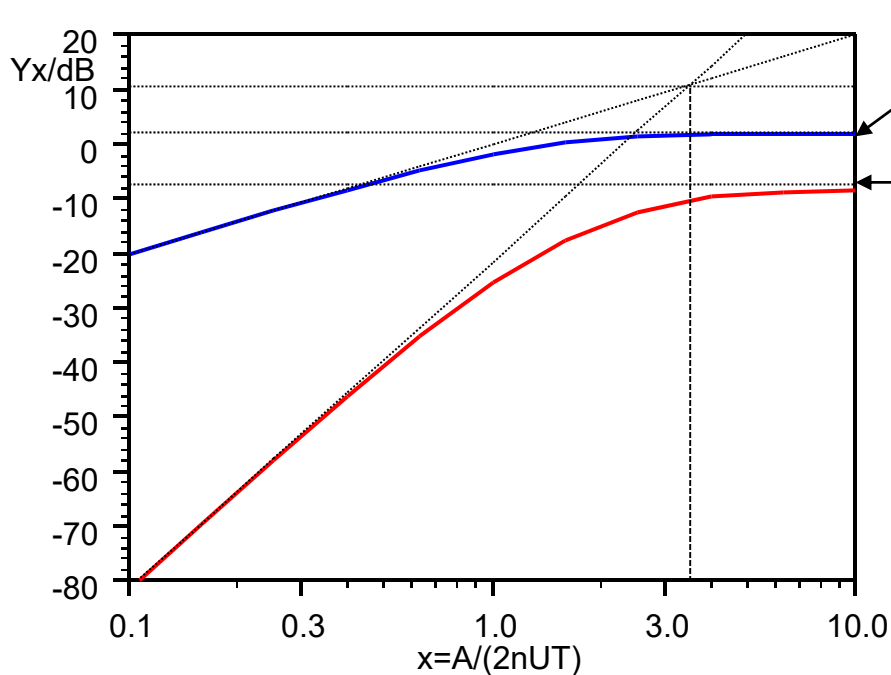$$20\log\left(x_{-1dB}\right) = 10\log\left(\frac{|\alpha_1|}{|\alpha_3|}\right) - 8.4 \cong -3.6dB$$

$$x_{IIP3} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}} = 2$$

$$20\log\left(x_{IIP3}\right) = 10\log\left(\frac{|\alpha_1|}{|\alpha_3|}\right) + 1.25 \cong 6dB$$

Examples: Differential Pair in WI – HB Simulations

- These results are confirmed by harmonic balanced (HB) simulations (Aplac)

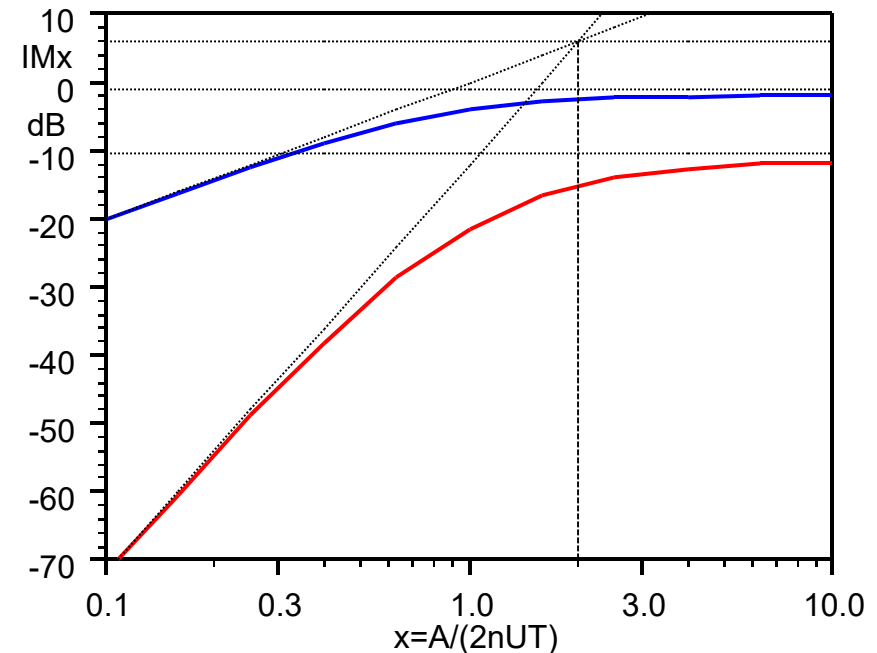Single tone analysis for diff. pair in weak inversion tanh(x)

$$20\log\left(\frac{4}{\pi}\right) = 2dB$$

$$20\log\left(\frac{4}{3\pi}\right) = -7.4dB$$

For large amplitude, the output becomes a square wave with $Y_1 = 4/\pi$ and $Y_3 = 4/(3*\pi)$

Dual tone analysis for diff. pair in weak inversion tanh(x)

Examples: Differential Pair in Strong Inversion

- The differential output current ΔI_{out} of a MOS differential pair biased in strong inversion is given by

$$y = \frac{\Delta I_{out}}{I_q} = \begin{cases} -1 & \text{for} \quad x/v_{gt} < -\sqrt{2} \\ x/v_{gt} \cdot \sqrt{1-\left(x/(2v_{gt})\right)^2} & \text{for} \quad -\sqrt{2} \leq x/v_{gt} \leq \sqrt{2} \\ +1 & \text{for} \quad \sqrt{2} < x/v_{gt} \end{cases} \quad \text{with} \quad \begin{aligned} x &= \frac{A}{nU_T} \\ v_{gt} &= \frac{V_{Gq}-V_{T0}}{nU_T} \end{aligned}$$
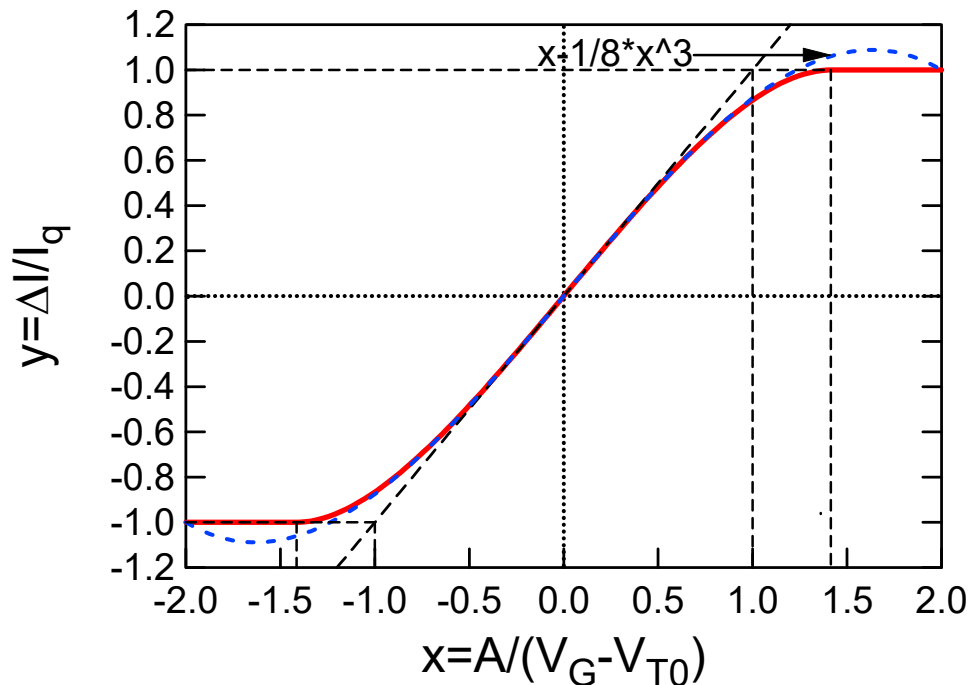
Taylor series expansion (in linear range):

$$y = \frac{x}{v_{gt}} \cdot \sqrt{1-\left(\frac{x}{2v_{gt}}\right)^2} \cong \frac{x}{v_{gt}} - \frac{1}{8}\left(\frac{x}{v_{gt}}\right)^3$$

and hence:

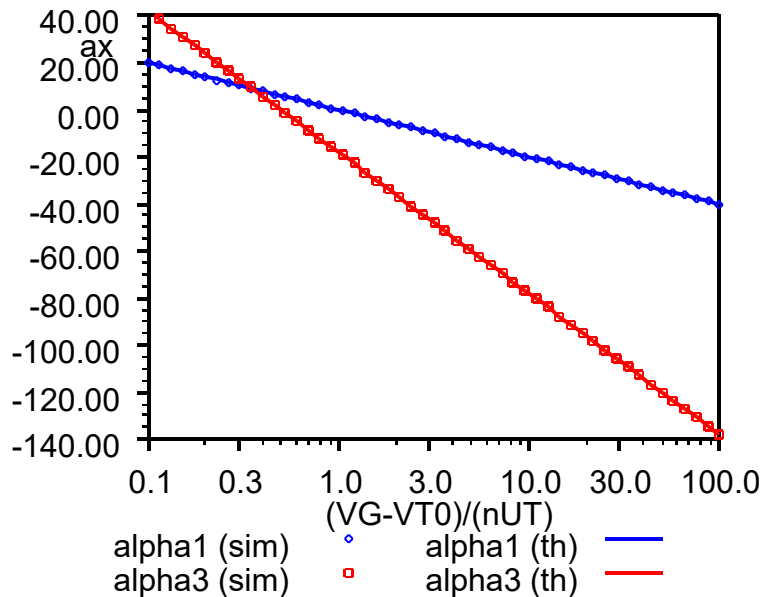$$\alpha_1 = \frac{1}{v_{gt}} \quad \text{and} \quad \alpha_3 = -\frac{1}{8v_{gt}^3}$$

Examples: DP in Strong Inversion – HB Simulations

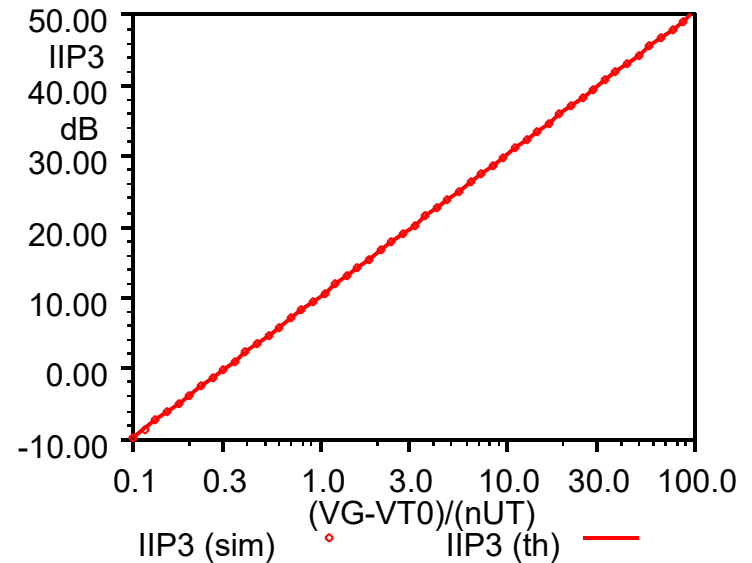$$\alpha_1 = \frac{1}{v_{gt}} = \frac{nU_T}{V_{Gq} - V_{T0}}$$

$$\alpha_3 = -\frac{1}{8 v_{gt}^3} = -\frac{1}{8}\left(\frac{nU_T}{V_{Gq} - V_{T0}}\right)^3$$

$$x_{IIP3} = \sqrt{\frac{32}{3}} \cdot v_{gt} = \sqrt{\frac{32}{3}} \cdot \frac{V_{Gq} - V_{T0}}{nU_T}$$

alpha1 and alpha3 versus overdrive voltage

IIP3 versus versus overdrive voltage

Examples: MOST in Weak Inversion

- In the case of weak inversion, the nonlinearity is a simple exponential function

- Assuming the sinewave is applied to the gate and the transistor is in saturation with $V_S = 0$, we have

$$I_D(t) = I_{D0} \cdot \exp\left[\frac{V_{Gq} + A\cos(\omega t)}{nU_T}\right] = I_q \cdot \exp\left[x \cdot \cos(\omega t)\right]$$

with $\quad I_q = I_{D0} \cdot \exp\left[\dfrac{V_{Gq}}{nU_T}\right] \quad$ and $\quad x = \dfrac{A}{nU_T}$

- The output current can then be normalized to the quiescent value (value for $A = 0$)

$$y(t) = \frac{I_D(t)}{I_q} = \exp\left[x \cdot \cos(\omega t)\right]$$

- The Fourier series of $y(t)$ is then given by

$$y(t) = \exp\left[x \cdot \cos(\omega t)\right] = I_0(x) + 2 \cdot \sum_{n=1}^{+\infty} I_n(x) \cdot \cos(n\omega t)$$

- where $I_n(x)$ are the modified Bessel function of the 1st kind. The harmonics can be calculated and compared to the approximation obtained from the Taylor series

$$f(x) = e^x \cong 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots \qquad \alpha_1 = 1 \quad \alpha_2 = \frac{1}{2} \quad \alpha_3 = +\frac{1}{6}$$

Examples: MOST in Weak Inversion – Single Tone

- Since the characteristic is expanding instead of compressing, the −1dB compression point becomes a +1dB expansion point
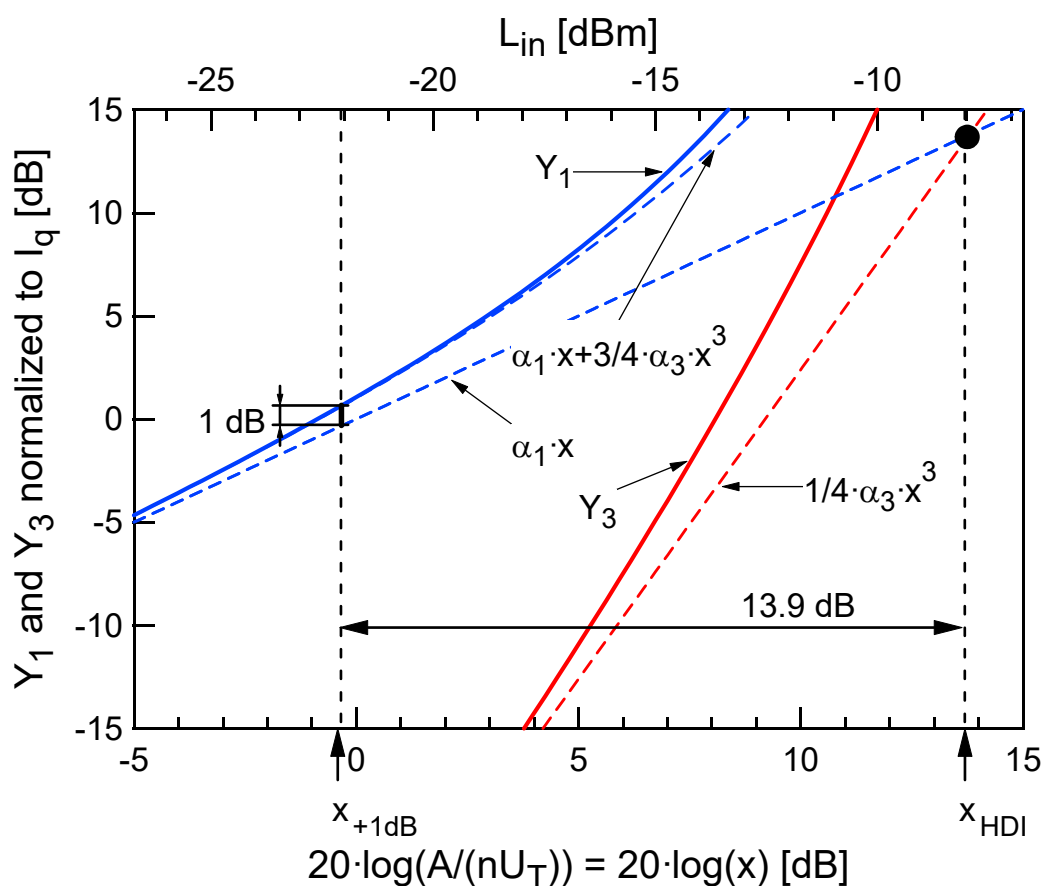
$$\alpha_1 = 1 \quad \text{and} \quad \alpha_3 = +\frac{1}{6}$$

$$x_{+1dB} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}\left(10^{+\frac{1}{20}}-1\right)} \cong 0.988$$
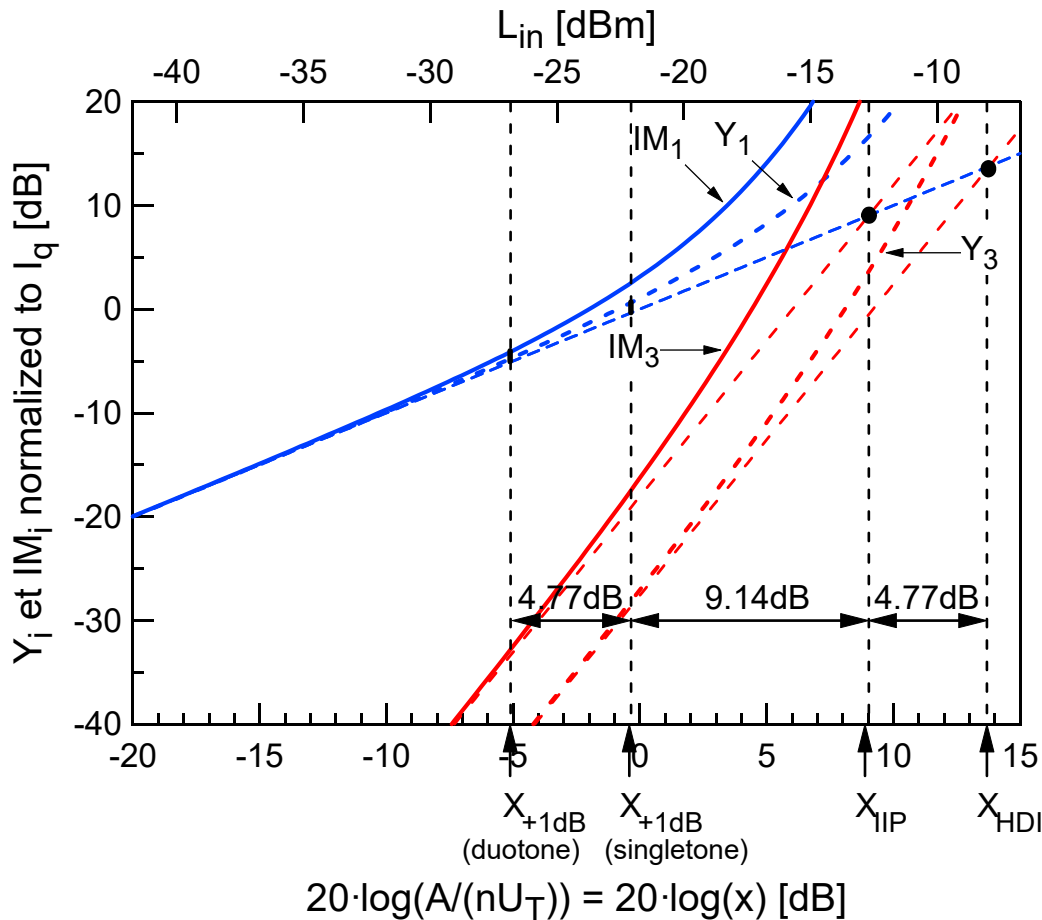
$$20\log(x_{+1dB}) = 10\log\left(\frac{|\alpha_1|}{|\alpha_3|}\right) - 7.9 \cong 0.1dB$$

$$x_{HDI} = 2\sqrt{\frac{|\alpha_1|}{|\alpha_3|}} \cong 4.9$$

$$20\log(x_{HDI}) = 10\log\left(\frac{|\alpha_1|}{|\alpha_3|}\right) + 6 \cong 13.8dB$$

Examples: MOST in Weak Inversion – Two-tone

- The two-tone characteristics are plotted below and compared to the single tone

$$\alpha_1 = 1 \quad \text{and} \quad \alpha_3 = +\frac{1}{6}$$

$$x_{IIP3} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}} = \sqrt{8} \cong 2.8$$

$$20\log\left(x_{IIP3}\right) = 10\log\left(\frac{|\alpha_1|}{|\alpha_3|}\right) + 1.25 \cong 9dB$$
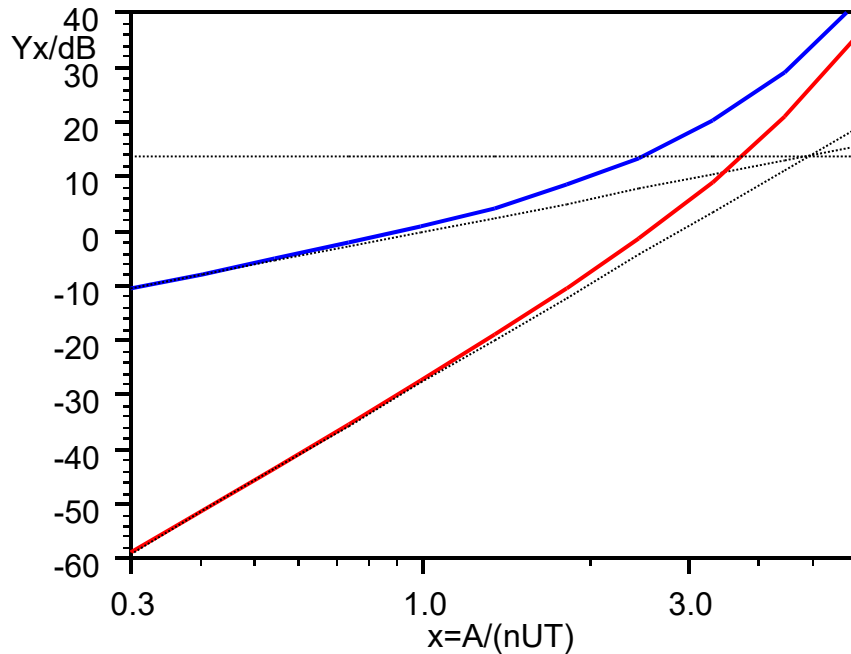
$$\frac{x_{HDI}}{x_{IIP3}} = 2 \cdot \sqrt{\frac{3}{4}} = 1.73$$

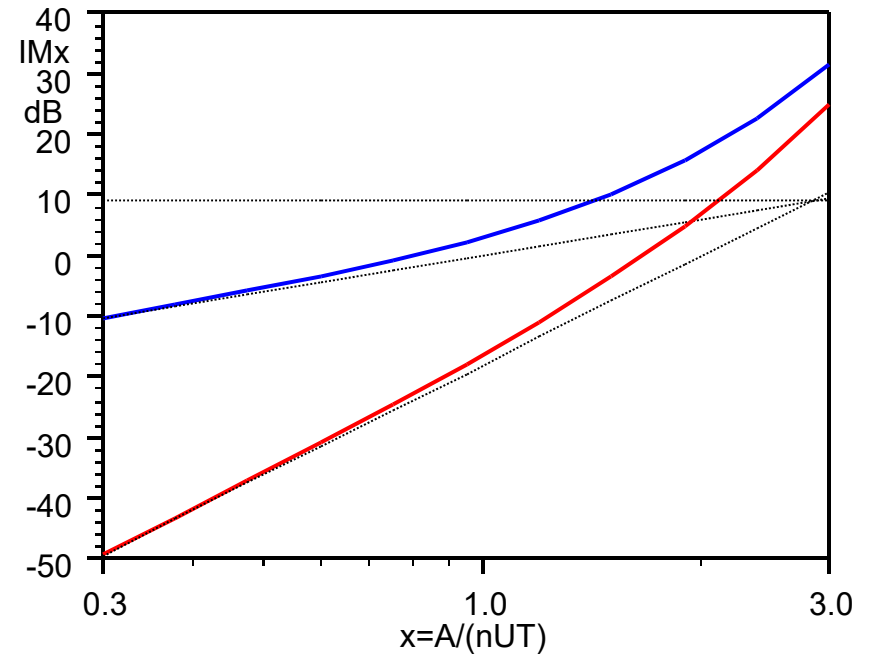$$20\log\left(\frac{x_{HDI}}{x_{IIP3}}\right) = 4.77dB$$

Examples: MOST in Weak Inversion – HB Simulations

- These results are confirmed by harmonic balanced (HB) simulations (Aplac)

Single tone analysis for transistor in weak inversion exp(x)

Dual tone analysis for MOS transistor in weak inversion exp(x)

Examples: MOST in Strong Inversion

- In the case of strong inversion and assuming a long channel and constant mobility, the nonlinearity is a simple quadratic function. Therefore there is no 3rd-order term and hence the IP3 is infinite!

- A more realistic model accounting for mobility reduction due to the vertical field is given by

$$\frac{I_D}{I_{spec}} = i_d = \frac{\left(\dfrac{V_P - V_S}{2U_T}\right)^2}{1 + \Theta \cdot V_P} = \frac{\left(\dfrac{V_G - V_{T0} - nV_S}{2nU_T}\right)^2}{1 + \Theta \cdot \dfrac{V_G - V_{T0}}{n}} \qquad \text{where parameter } \Theta \text{ is typically 0.1 1/V}$$

- For $V_S = 0$ it can be written in simple normalized form as

$$i_d = \frac{\left(v_{gt}/2\right)^2}{1 + \theta \cdot v_{gt}/2} \quad \text{with} \quad v_{gt} = \frac{V_G - V_{T0}}{nU_T} \quad \text{and} \quad \theta = 2U_T \cdot \Theta$$

- Adding the sine wave at the gate it becomes

$$i_d(t) = \frac{\left(\dfrac{v_{gt} + v_{in}(t)}{2}\right)^2}{1 + \theta \cdot \dfrac{v_{gt} + v_{in}(t)}{2}} \quad \text{with} \quad v_{in}(t) = x \cdot \cos(\omega t) \quad \text{and} \quad x = \frac{A}{nU_T}$$

Examples: MOST in Strong Inversion

- The current can be rewritten as $i_d(t) = i_q \cdot f(x)$, where i_q is the normalized bias current

$$f(x) = \frac{i_d}{i_q} = \frac{\left(1 + x/v_{gt}\right)^2}{1 + \theta' \cdot x/v_{gt}} \quad \text{with} \quad \theta' = \frac{\theta \, v_{gt}/2}{1 + \theta \, v_{gt}/2} \quad \text{and} \quad i_q = \frac{\left(v_{gt}/2\right)^2}{1 + \theta \cdot v_{gt}/2}$$

- The nonlinear behavior is then captured by the single function $f(x)$ which can be developed in a Taylor series as

$$f(x) \cong 1 + \alpha_1' \frac{x}{v_{gt}} + \alpha_2' \left(\frac{x}{v_{gt}}\right)^2 + \alpha_3' \left(\frac{x}{v_{gt}}\right)^3 + \ldots = 1 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \ldots$$

$$\alpha_1' = 2 - \theta' \qquad \alpha_2' = \left(1 - \theta'\right)^2 \qquad \alpha_3' = -\theta' \cdot \left(1 - \theta'\right)^2$$

with
$$\alpha_1 = \frac{\alpha_1'}{v_{gt}} = \frac{2 - \theta'}{v_{gt}} \quad \alpha_2 = \frac{\alpha_2'}{v_{gt}^2} = \frac{\left(1 - \theta'\right)^2}{v_{gt}^2} \quad \alpha_3 = \frac{\alpha_3'}{v_{gt}^3} = -\frac{\theta' \cdot \left(1 - \theta'\right)^2}{v_{gt}^3}$$

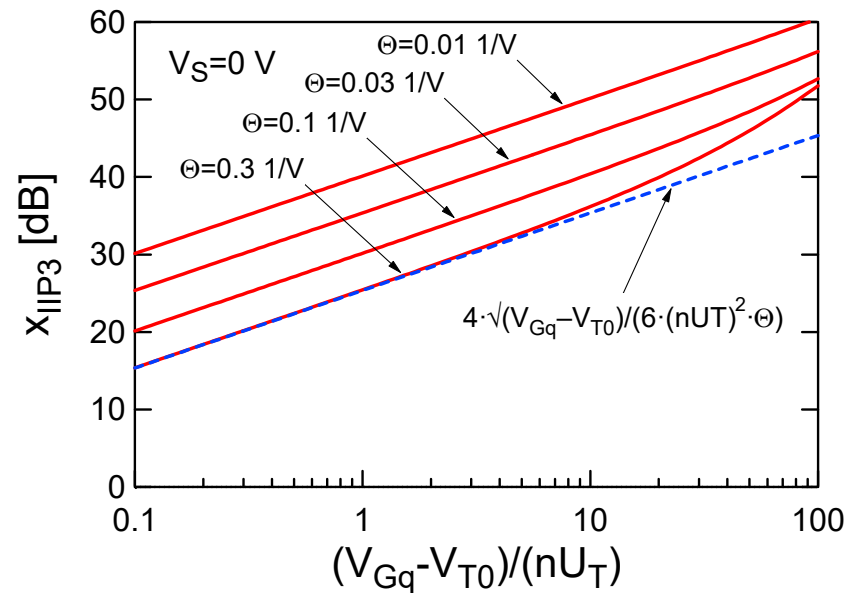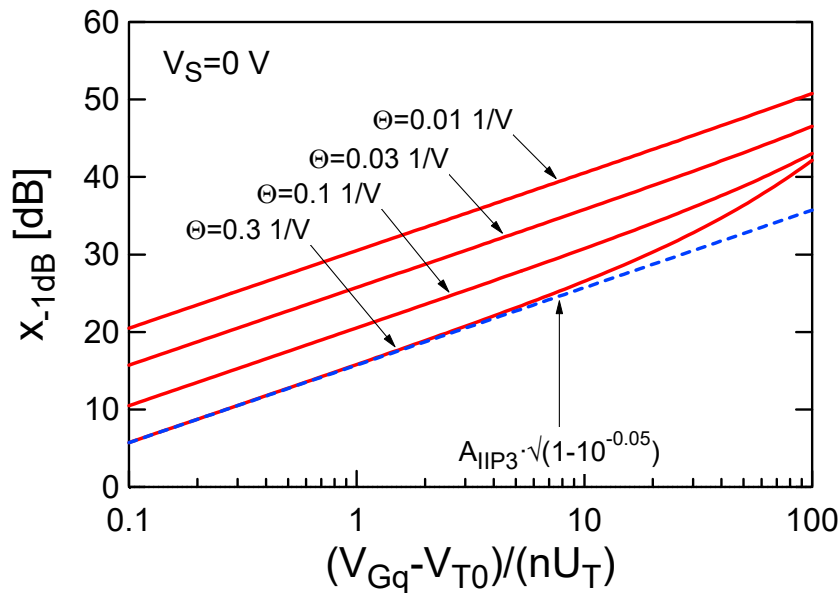- Usually $\theta' \ll 1$ and the coefficients can be simplified as

$$\alpha_1 \cong \frac{2}{v_{gt}} = \frac{1}{\sqrt{IC}} \quad \alpha_2 \cong \frac{1}{v_{gt}^2} = \frac{1}{4 \cdot IC} \quad \alpha_3 \cong -\frac{\theta'}{v_{gt}^3} \cong -\frac{\theta}{2v_{gt}^2} = -\frac{\theta}{8 \cdot IC}$$

Examples: MOST in Strong Inversion

- The IIP3 and -1dB points are then given by

$$x_{IIP3} = \frac{A_{IIP3}}{nU_T} = \sqrt{\frac{4}{3}\frac{|\alpha_1|}{|\alpha_3|}} \cong \sqrt{\frac{8}{3}\frac{v_{gt}^2}{\theta'}} \cong 4\sqrt{\frac{v_{gt}}{3\theta}} = 4\sqrt{\frac{V_{Gq}-V_{T0}}{6(nU_T)^2\,\Theta}}$$

$$x_{-1dB} = \frac{A_{-1dB}}{nU_T} = \sqrt{\left(1-10^{-\frac{1}{20}}\right)\cdot\frac{4}{3}\cdot\frac{|\alpha_1|}{|\alpha_3|}} = \sqrt{\left(1-10^{-\frac{1}{20}}\right)\cdot x_{IIP3}} \cong \frac{x_{IIP3}}{3}$$

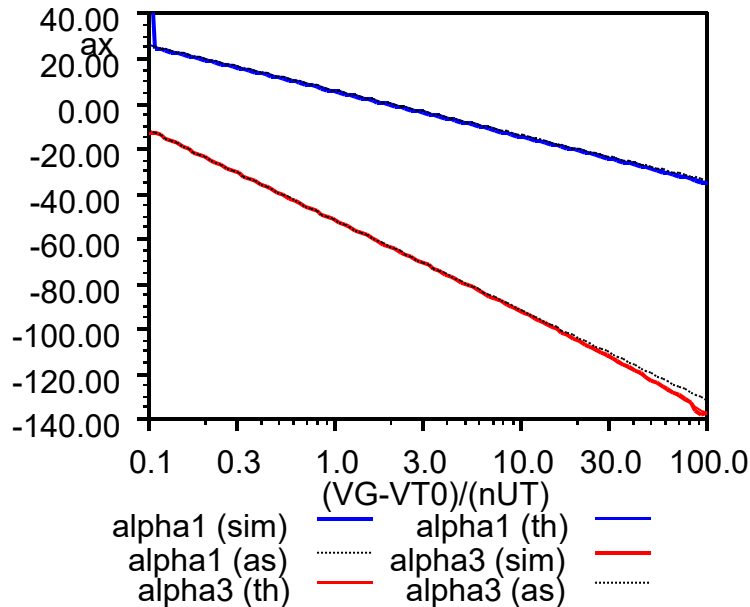# Examples: MOST in Strong Inversion – HB Simulations

- HB simulations (Aplac) for $\Theta = 0.1 \; 1/V$ for ideal MOST in strong inversion
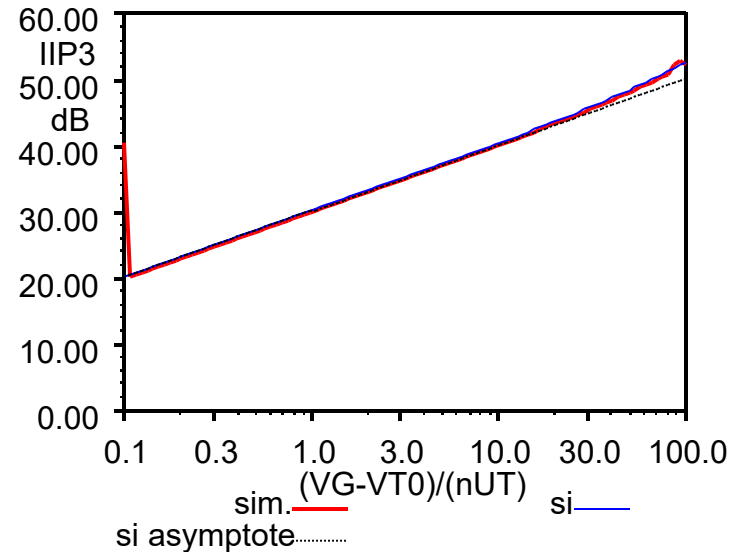
$$\alpha_1 \cong \frac{2}{v_{gt}}$$

$$\alpha_3 \cong -\frac{\theta'}{v_{gt}^3} \cong -\frac{\theta}{2v_{gt}^2}$$

$$x_{IIP3} = \frac{A_{IIP3}}{nU_T} \cong \sqrt{\frac{8}{3}\frac{v_{gt}^2}{\theta'}} \cong 4\sqrt{\frac{v_{gt}}{3\theta}}$$
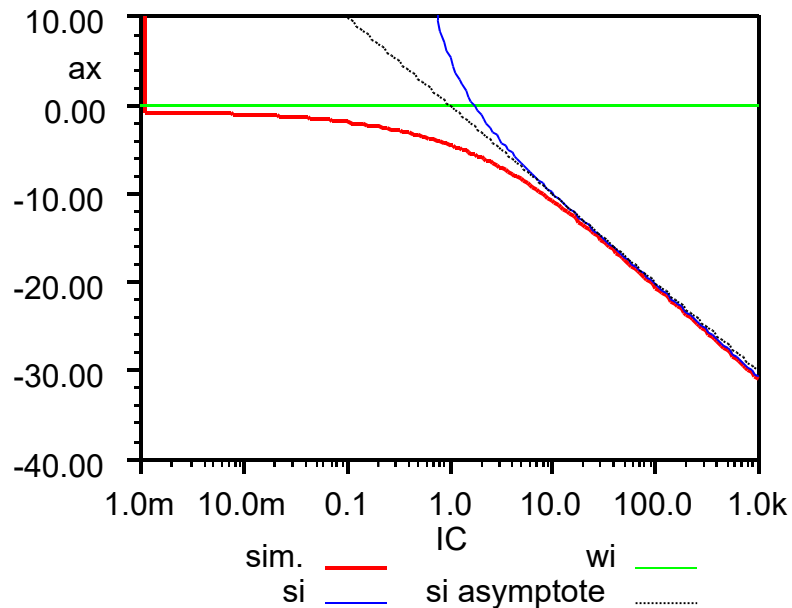
alpha1 and alpha3 versus overdrive voltage

IIP3 versus versus overdrive voltage
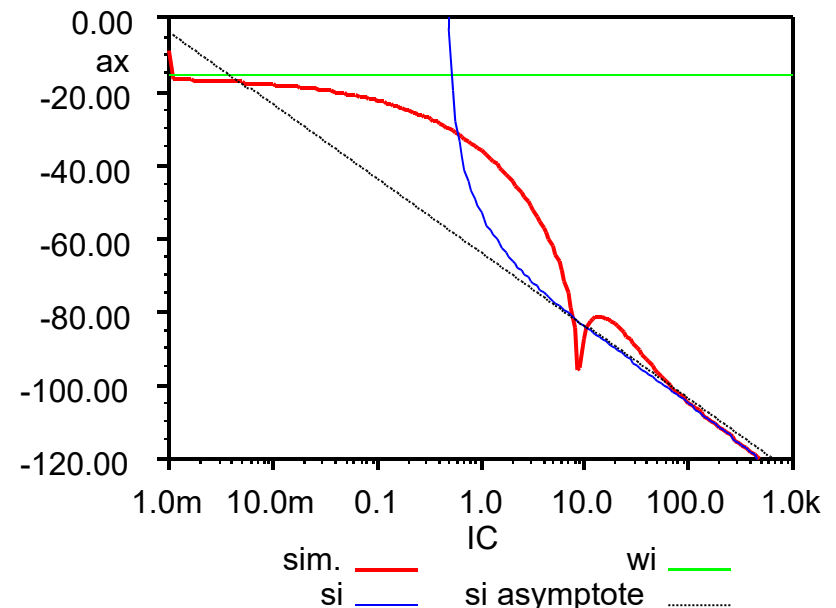
Examples: Continuous Model of MOST– HB Simulations

- HB simulations (Aplac) for $\Theta = 0.1\ 1/V$ for ideal MOST from weak to strong inversion

alpha1 versus inversion coefficient

alpha3 versus inversion coefficient

$$\alpha_1 = \begin{cases} 1 & WI \\ \dfrac{2}{v_{gt}} = \dfrac{1}{\sqrt{IC}} & SI \end{cases}$$
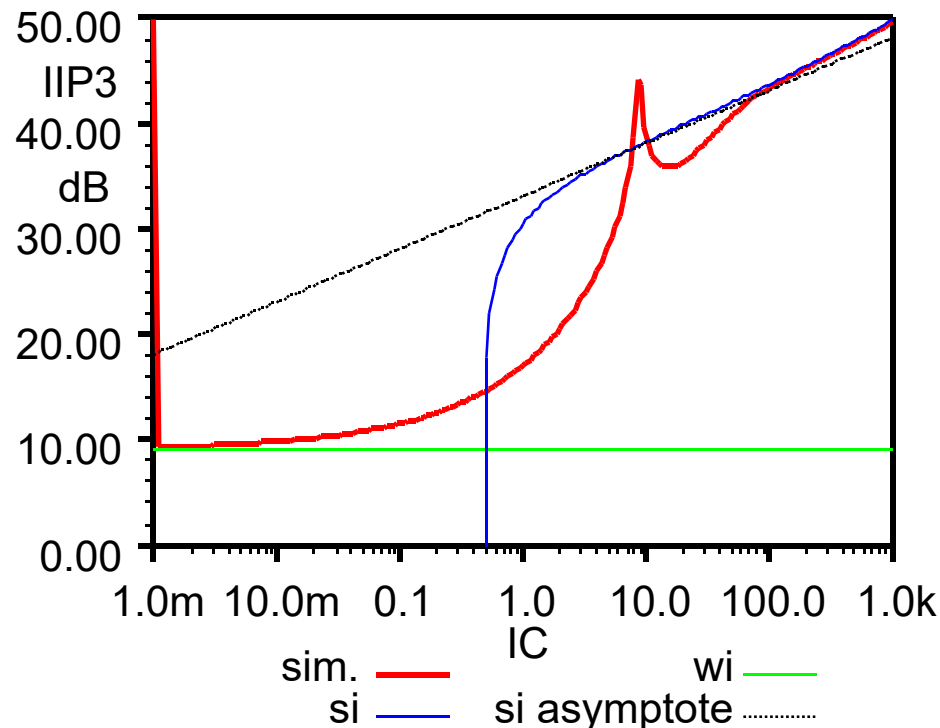
$$\alpha_3 = \begin{cases} \dfrac{1}{6} & WI \\ -\dfrac{\theta}{2v_{gt}^2} = -\dfrac{\theta}{8 \cdot IC} & SI \end{cases}$$

Examples: Continuous Model of MOST– HB Simulations

- HB simulations (Aplac) for $\Theta = 0.1 \; 1/V$ for ideal MOST from weak to strong inversion

$$x_{IIP3} = \frac{A_{IIP3}}{nU_T} = \begin{cases} \sqrt{8} & WI \\ 4\sqrt{\dfrac{v_{gt}}{3\theta}} = 4\sqrt{\dfrac{2\sqrt{IC}}{3\theta}} & SI \end{cases}$$

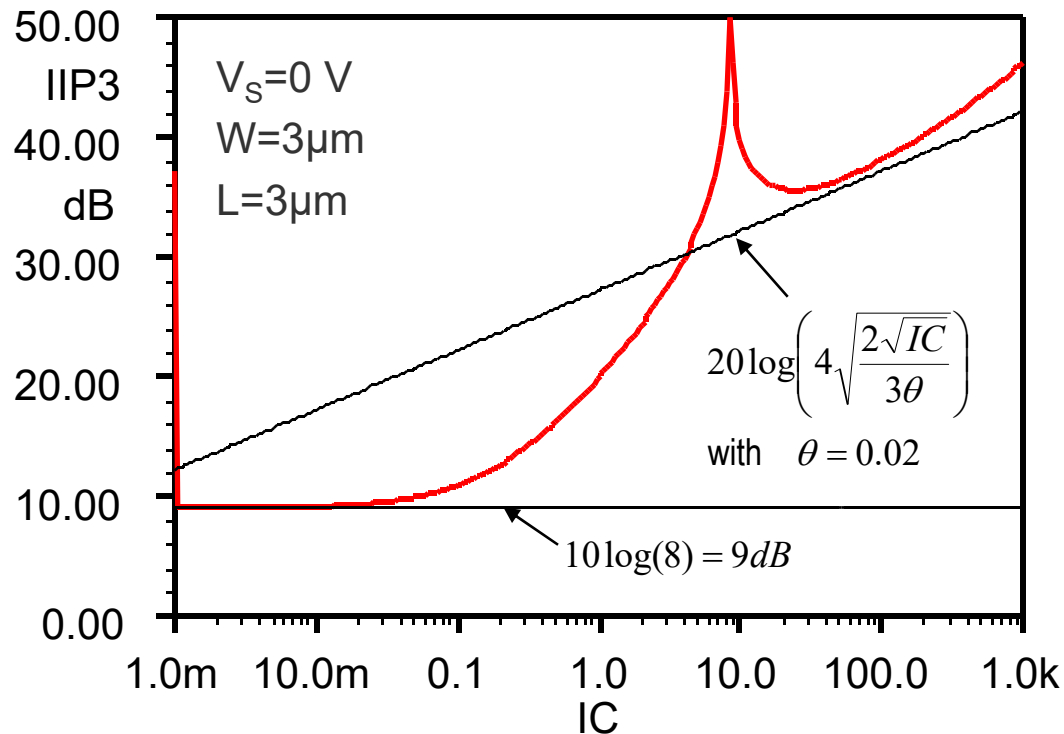IIP3 versus inversion coefficient

sim. ——   IC   wi ——
si ——   si asymptote ··········

Examples: Compact Model of MOST– HB Simulations

- Simulations with complete EKV v2.6 model for a generic 0.18μm CMOS process

IIP3 versus inversion coefficient

$V_S=0$ V
W=3μm
L=3μm

$20\log\left(4\sqrt{\dfrac{2\sqrt{IC}}{3\theta}}\right)$

with $\theta = 0.02$

$10\log(8) = 9dB$
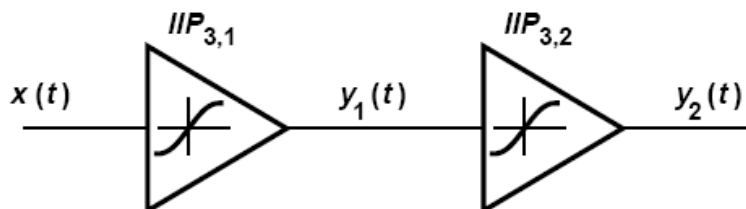
Outline

- General considerations

- Nonlinearity and time variance

- Effects of nonlinearity
  - Single tone input
    - Harmonic distortion
    - Gain compression
  - Two-tone input
    - Desensitization and blocking
    - Cross modulation
    - Intermodulation
    - Third-order Intercept Point

- **Cascade of nonlinear stages**

Cascade of Nonlinear Stages

$$y_1(t) = \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t)$$

$$y_2(t) = \beta_1 y_1(t) + \beta_2 y_1^2(t) + \beta_3 y_1^3(t)$$

Considering only the 1st- and 3rd-order terms

$$y_2(t) = \alpha_1 \beta_1 x(t) + \left(\alpha_3 \beta_1 + 2\alpha_1 \alpha_2 \beta_2 + \alpha_1^3 \beta_3 \right) x^3(t) + \ldots$$

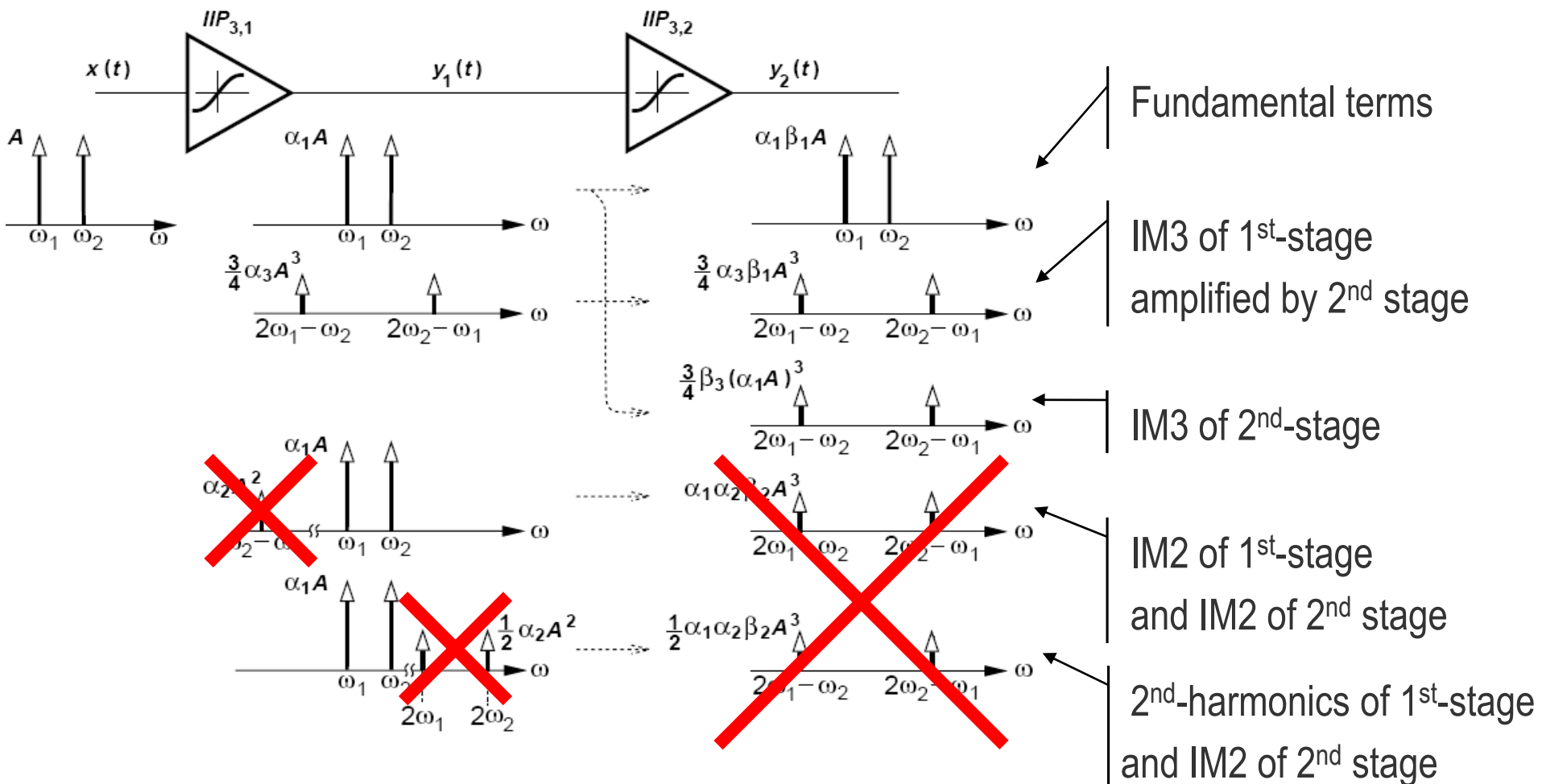- The overall IIP3 is then given by

$$IIP3_{tot} = \sqrt{ \frac{4}{3} \left| \frac{\alpha_1 \beta_1}{\alpha_3 \beta_1 + 2\alpha_1 \alpha_2 \beta_2 + \alpha_1^3 \beta_3} \right| }$$

- In worst case the above expression reduces to

$$\frac{1}{IIP3_{tot}^2} = \frac{3}{4} \frac{\left| \alpha_3 \beta_1 \right| + \left| 2\alpha_1 \alpha_2 \beta_2 \right| + \left| \alpha_1^3 \beta_3 \right|}{\left| \alpha_1 \beta_1 \right|} = \frac{1}{IIP3_1^2} + \frac{3\alpha_2 \beta_2}{2\beta_1} + \frac{\alpha_1^2}{IIP3_2^2}$$

- As the 1st-stage gain α_1 increases, the overall IIP3 decreases

Cascade of Nonlinear Stages

Fundamental terms

IM3 of 1st-stage amplified by 2nd stage

IM3 of 2nd-stage

IM2 of 1st-stage and IM2 of 2nd stage

2nd-harmonics of 1st-stage and IM2 of 2nd stage

- RF amplifiers are usually narrow-band and therefore the components at $2\omega_1$, $2\omega_2$ and $\omega_2 - \omega_1$ are filtered and hence the bottom components can be neglected

Cascade of Nonlinear Stages

- Considering only the 1st- and 3rd-order terms, we have

$$y_2(t) = \alpha_1\beta_1 x(t) + \left(\alpha_3\beta_1 + 2\alpha_1\alpha_2\beta_2 + \alpha_1^3\beta_3\right)x^3(t) + \ldots$$

- Assuming that the 2nd harmonics and IM2 are filtered before the 2nd stage, the 2nd term in front of the 3rd-order term can usually be neglected, resulting in considering only the IM3 of the 1st and 2nd stage

$$y_2(t) \cong \alpha_1\beta_1 x(t) + \left(\alpha_3\beta_1 + \alpha_1^3\beta_3\right)x^3(t) + \ldots$$

- The IIP3 for two stages is then given by

$$\frac{1}{IIP3_{tot}^2} \cong \frac{3}{4}\frac{\left|\alpha_3\beta_1\right| + \left|\alpha_1^3\beta_3\right|}{\left|\alpha_1\beta_1\right|} = \frac{1}{IIP3_1^2} + \frac{\alpha_1^2}{IIP3_2^2}$$

- Finally the general relation for cascaded stages under the above assumptions is given by

$$\frac{1}{IIP3_{tot}^2} \cong \frac{1}{IIP3_1^2} + \frac{\alpha_1^2}{IIP3_2^2} + \frac{\alpha_1^2\beta_1^2}{IIP3_3^2} + \ldots$$

References

Most of this Chapter is based on Chapter 2 of Reference [1]

[1] B. Razavi, *RF Microelectronics*, 2nd ed. Pearson, 2012.

[2] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, 2nd ed. Cambridge University Press, 2004.

[3] G. D. Vendelin, A. M. Pavio and U. L. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques,* 2nd ed. Wiley, 2006.