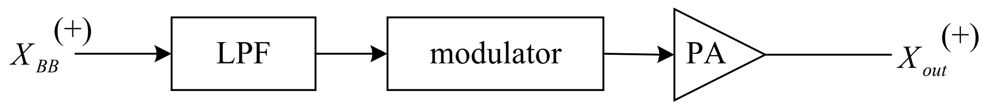
Nonlinearity and Distortion

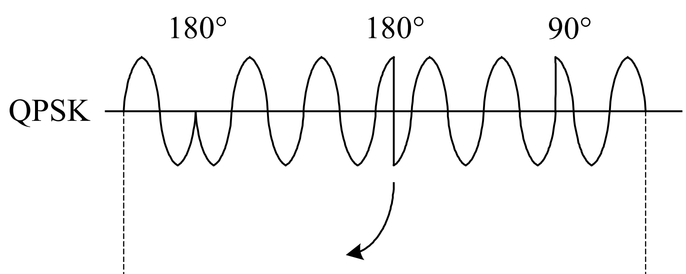
Let's consider two examples

Example 1

you have a QPSK transmitter



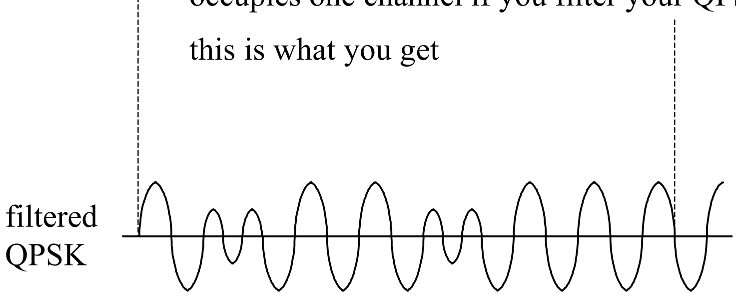
If you do not filter your QPSK this is what you have



but to send this out you need an infinite bandwidth

→ you do not have infinite bandwidth

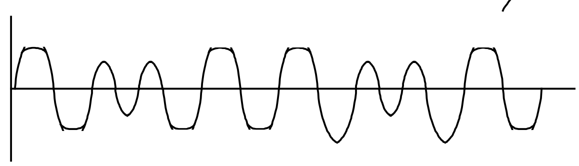
So you have to filter your signal such that it only occupies one channel if you filter your QPSK with a LPF this is what you get



because of limited bandwidth your phase information turns into amplitude information

Now if your PA clamps the peak of your filtered QPSK

you loose your information



distortion due to damping changes your amplitude

you loose information now you cannot correctly detect your phase in your receiver

$\longrightarrow$ QPSK is very sensitive to amp linearity

FSK or FM is not sensitive to amp linearity

since you do not have instant phase variation

$\longrightarrow$ Choice of your modulation depends on

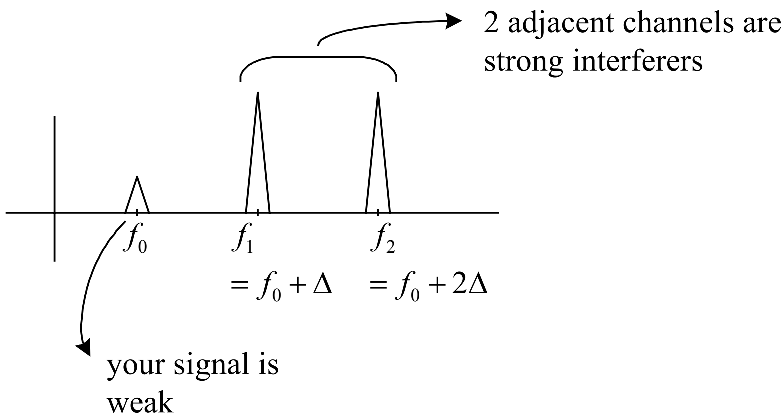
how linear your amplifier can be

QPSK needs very linear amplifier $\longrightarrow$ usually class A not very efficient

FSK, FM not need linear amplifier $\longrightarrow$ you can use class B, C, ,,, amplifier with much higher efficiency

Example 2

your receiver receives 3 signals

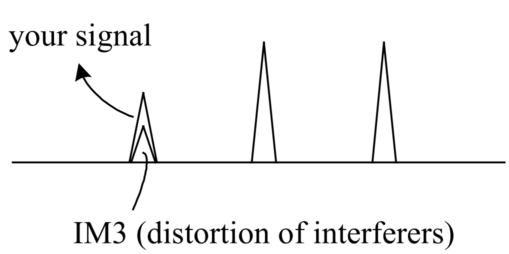


in your LNA you have nonlinearity so your strong interferers mix with each other due to LNA nonlinearity

This mixing generates frequencies such as

$$2f_1 - f_2 = 2(f_0 + \Delta) - (f_0 + 2\Delta) = f_0$$

This is what you get at the output of LNA



This distortion makes your reception difficult

$\longrightarrow$ we have to understand distortion and nonlinearity

Non-linearity

- **Weakly non-linear systems** (such as in receivers)
  LNA
- also clamping
  Strongly non-linear systems (such as in high efficiency transmitter)
  class B-E PA's

to deal with weakly non-linear system there are two

approaches

- **power series** ⟶ represent transfer function using a power series
- **volterra series** ⟶ power series with information about phase

for strongly non linear system ⟶ use **envelope analysis**

instead of frequency domain analysis

use time domain analysis

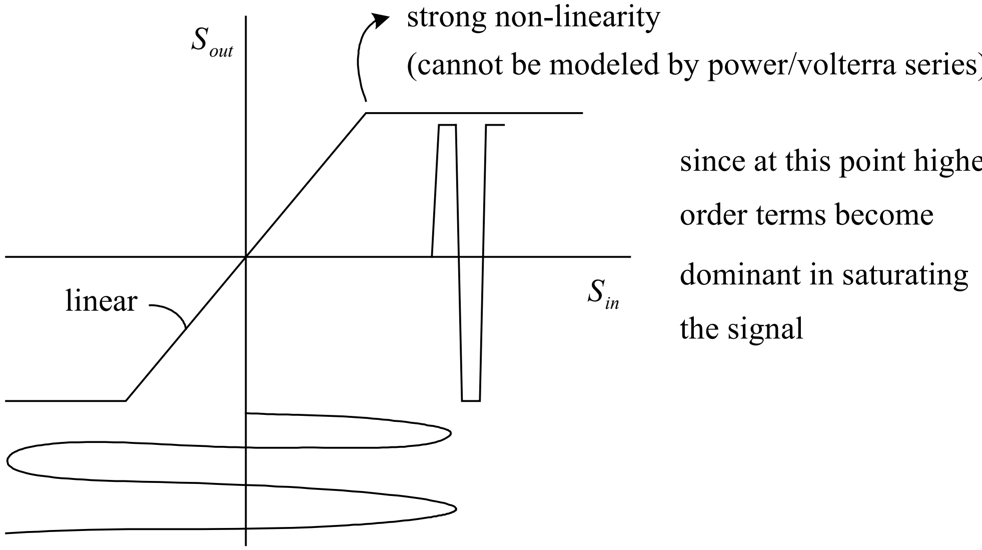
* At midband frequency

parasitic caps and inductors
bias and bypass caps

can be neglected
in terms of the phase
they introduce ⟶ use
power series

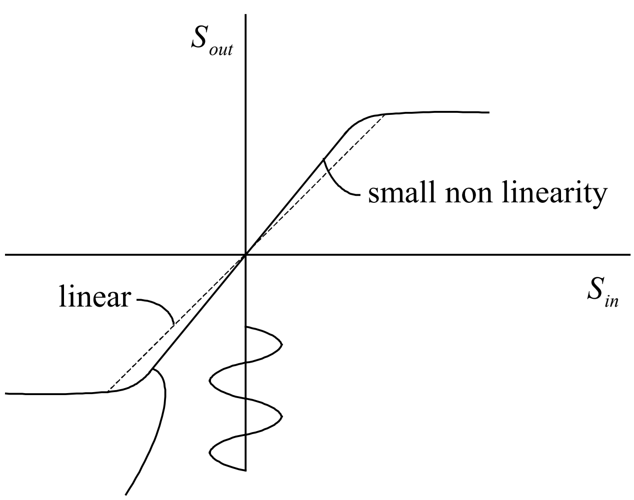
* At low or high frequencies

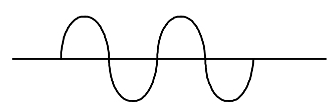
you need to know the phase of the signal
that is subjected to non-linearily ⟶ use
volterra series



since at this point higher order terms become dominant in saturating the signal

* most often this is what you have





you can write this transfer function in general by a power series

$$S_{out} = a_1 S_{in} + a_2 S_{in}^2 + a_3 S_{in}^3 + \cdots$$

where a_1, a_2, a_3, ,,, are constants

* they usually vary with signal amplitude
* they vary with the choice of bias
* they vary with frequency
* they vary with input/output matching

let's assume : they are constant (ignore variation with signal amplitude)
(fixed bias, small frequency range, fixed matching)

now apply an input function $S_{in} = f(\omega t)$

$$S_{out} = a_1 f(\omega t) + a_2 f^2(\omega t) + a_3 f^3(\omega t) + \cdots$$

$$let \quad f(\omega t) = m_1 \cos\omega_1 t + m_2 \cos\omega_2 t \qquad \omega_1 \sim \omega_2$$

$$S_{out} = a_1(m_1 \cos\omega_1 t + m_2 \cos\omega_2 t) + a_2(m_1 \cos\omega_1 t + m_2 \cos\omega_2 t)^2 + a_3(m_1 \cos\omega_1 t + m_2 \cos\omega_2 t)^3 + \cdots$$

use trigonometric
expansions

$$\begin{aligned} S_{out} &= a_1(m_1 \cos\omega_1 t + m_2 \cos\omega_2 t) \\ &+ a_2\left(m_1^2 \frac{1+\cos 2\omega_1 t}{2} + m_2^2 \frac{1+\cos 2\omega_2 t}{2} + m_1 m_2 \frac{\cos(\omega_1+\omega_2)t + \cos(\omega_1-\omega_2)t}{2}\right) \\ &+ a_3\left\{\left(m_1^3\left(\frac{\cos\omega_1 t}{2} + \frac{\cos\omega_2 t}{4} + \cdots \right)\right)\right\} \end{aligned}$$

$$S_{out} = a_2\left(\frac{m_1^2 + m_2^2}{2}\right) \quad \Big\} \; dc$$

$$\left.\begin{aligned} &+ a_1\left(m_1 \cos\omega_1 t + m_2 \cos\omega_2 t\right) \\ &+ a_3\left(\frac{3}{4} m_1^3 \cos\omega_1 t + \frac{3}{4} m_2^3 \cos\omega_2 t + \frac{3}{2} m_1 m_2^2 \cos\omega_1 t + \frac{3}{2} m_1^2 m_2^2 \cos\omega_2 t\right)\end{aligned}\right\} \; \omega_1, \omega_2$$

$$+ a_2\left(\frac{m_1}{2}^2 \cos 2\omega_1 t + \frac{m_2}{2}^2 \cos 2\omega_2 t\right) \quad \Big\} \; 2\omega_1,\, 2\omega_2$$

$$+ a_2\left(\frac{m_1 m_2}{2} \cos(\omega_1 + \omega_2)t + \frac{m_1 m_2}{2} \cos(\omega_1 - \omega_2)t\right) \quad \left\}\begin{array}{l} \omega_1 + \omega_2 \\ \omega_1 - \omega_2 \end{array}\right.$$

$$+ a_3\left(\frac{3}{4} m_1^2 m_2 \cos(2\omega_1 - \omega_2)t + \cdots\right) \quad \left\}\begin{array}{l} 2\omega_1 - \omega_2 \\ 2\omega_2 - \omega_1 \\ 2\omega_1 + \omega_2 \\ 2\omega_2 + \omega_1 \end{array}\right.$$

$$+ a_3\left(\frac{1}{4} m_1^3 \cos(3\omega_1)t + \cdots\right) \quad \left\}\begin{array}{l} 3\omega_1 \\ 3\omega_2 \end{array}\right.$$

Let's look at the important coefficient

① *coefficient of* $\cos\omega_1 t \Rightarrow$ $\left[a_1 m_1 + a_3\left(\frac{3}{4}m_1^3 + \frac{3}{2}m_1 m_2^2\right)\right]\cos\omega_1 t$

② *coefficient of* $\cos(\omega_1 - \omega_2) \Rightarrow$ $\left[a_2\left(\frac{m_1 m_2}{2}\right)\right]\cos(\omega_1 - \omega_2)t$

③ *coefficient of* $\cos 2\omega_1 t \Rightarrow$ $\left[a_2 \frac{m_1^2}{2}\right]\cos 2\omega_1 t$

④ *coefficient of* $\cos(2\omega_1 - \omega_2)t \Rightarrow$ $\left[a_3 \frac{3}{4}m_1^2 m_2\right]\cos(2\omega_1 - \omega_2)t$

⑤ *coefficient of* $\cos 3\omega_1 t \Rightarrow$ $\left[a_3 \frac{1}{4}m_1^3\right]\cos 3\omega_1 t$

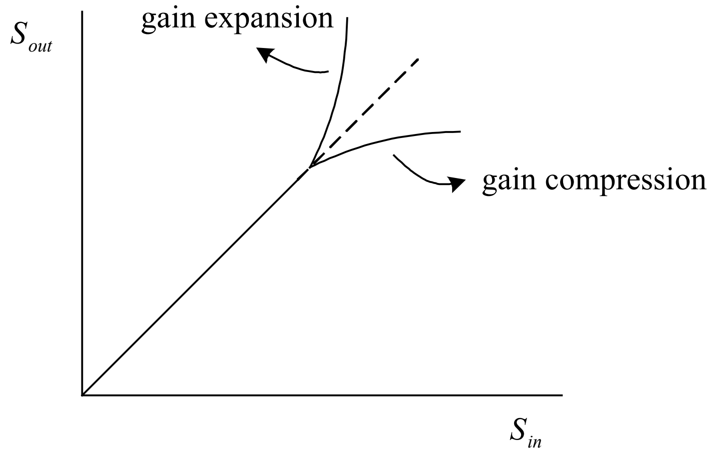
First non-linearity

① **Gain compression**

for only one tone : $f(\omega t) = m_1 \cos\omega_1 t$

$$\longrightarrow \left(\underbrace{a_1 m_1}_{\text{linear gain}} + \underbrace{a_3 \frac{3}{4}m_1^3}_{\text{gain compression / expansion}}\right)\cos\omega_1 t$$

as signal becomes stronger



usually we observe gain comperssion that means a3 is usually negative

② **IM2**

$$a_2 \frac{m_1 m_2}{2} \cos(\omega_1 - \omega_2) + a_2 \frac{m_1 m_2}{2} \cos(\omega_2 - \omega_1)t = a_2 m_1^2 \cos(\omega_1 - \omega_2)t \qquad \text{good mixer}$$

$$m_1 = m_2$$

assume $m_1 = m_2$

$$\longrightarrow \quad IM2 = \frac{amplitude\ of\ 2nd\ order}{amplitude\ of\ fundamental} = \frac{a_2}{a_1} m_1$$

mixing comes from 2nd order (4th order inter modulation etc? terms (IM2)

$\longrightarrow$ good news for MOSFETs since their input/output relationship is of 2nd order nature

$$i_d = const\ V_{gs}^2$$

3rd order term also exists in MOSFETs due to two mechanisms

strong non-linearly gain clamping + body effect $\longrightarrow$ weakly non-linear effect

③ **HD2**

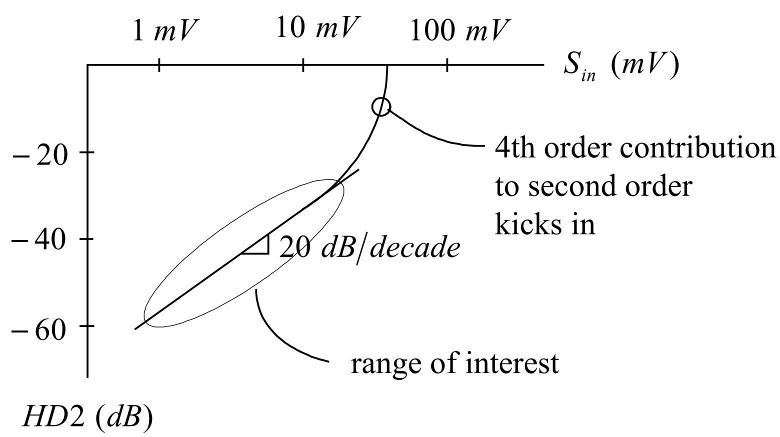
$$a_2 \frac{m_1^2}{2} \cos 2\omega_1 t \longrightarrow \text{2nd harmonic}$$

Define : 2nd harmonic distortion factor HD2

$$HD2 = \frac{second\ \ hamonic\ \ amplitude}{fundamental\ \ amplitude} = \frac{\frac{a_2}{2} m_1^2}{a_1 m_1}$$

assuming gain compression has not occurred yet

$$HD2 = \frac{1}{2} \frac{a_2}{a_1} m_1$$



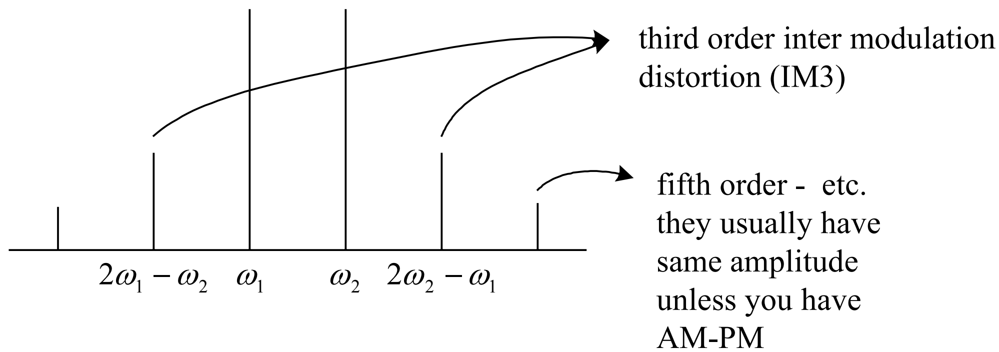
IM2 = 2HD2

2nd intermod

2nd harm distortion

④ **IM3** *assume* $m_1 = m_2$

$$a_3 \frac{3}{4} m_1^3 \cos(2\omega_1 - \omega_2)t = a_3 \frac{3}{4} m_1^3 \cos(2\omega_2 - \omega_1)t$$



IM3 = third order inter modulation

$m_1 = m_2 \longrightarrow$ $$IM3 = \frac{\text{amplitude of third order intermod}}{\text{amplitude of fundamental}}$$

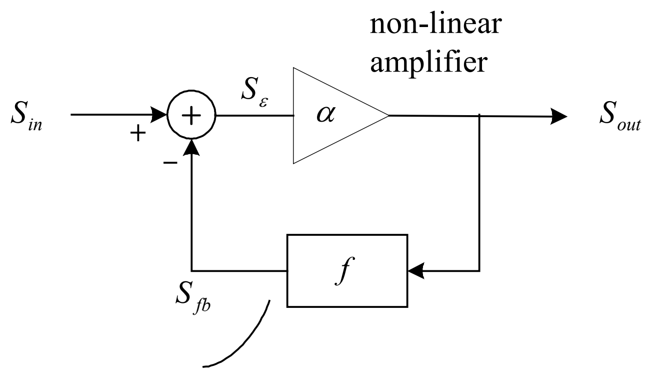
$$IM3 = \frac{a_3 \frac{3}{4} m_1^3}{a_1 m_1} = \frac{3}{4} \frac{a_3}{a_1} m_1^2$$

⑤ **HD3** : $\cos 3\omega_1 t$

$$HD3 = \frac{\text{third order amplitude}}{\text{fundamental amplitude}} = \frac{a_3 \times \frac{1}{4} m_1^3}{a_1 m_1} = \frac{1}{4} \frac{a_3}{a_1} m_1^2$$

$\boxed{\text{IM3 = 3HD3}}$

Effect of feedback on distortion



assumption : feedback network is linear

use power series :

$$\left.\begin{aligned} S_{out} &= a_1 S_\varepsilon + a_2 S_\varepsilon^2 + a_3 S_\varepsilon^3 + \cdots \\ S_\varepsilon &= S_{in} - f\, S_{out} \end{aligned}\right\} \rightarrow$$

$$S_{out} = a_1\left(S_{in} - f\, S_{out}\right) + a_2\left(S_{in} - f\, S_{out}\right)^2 + a_3\left(S_{in} - f\, S_{out}\right)^3 + \cdots$$

substitute

$$S_{out} = b_1 S_{in} + b_2 S_{in}^2 + b_3 S_{in}^3$$ ← we are looking for b_1, b_2, b_3 coefficients

$$\begin{aligned} b_1 S_{in} + b_2 S_{in}^2 + b_3 S_{in}^3 = {} & a_1 S_{in} - a_1 f\, b_1 S_{in} - a_1 f\, b_2 S_{in}^2 - a_1 f\, b_3 S_{in}^3 \\ & + a_2\left(S_{in} - f\, b_1 S_{in} - f\, b_2 S_{in}^2 - f\, b_3 S_{in}^3\right)^2 \\ & + a_3\left(S_{in} - f\, b_1 S_{in} - f\, b_2 S_{in}^2 - f\, b_3 S_{in}^3\right)^3 \end{aligned}$$

1st order terms $\longrightarrow$ $b_1 \cancel{S_{in}} = a_1 \cancel{S_{in}} - a_1 f\, b_1 \cancel{S_{in}}$

$$\boxed{b_1 = \frac{a_1}{1 + a_1 f}}$$

standard feedback equation for a linear system

2nd order terms

$$b_2 \cancel{S_{in}^2} = -a_1 f\, b_2 \cancel{S_{in}^2} + a_2 (1 - f\, b_1)^2 \cancel{S_{in}^2}$$

$$b_2 (1 + a_1 f) = a_2 (1 - f\, b_1)^2 = a_2 \frac{1}{(1 + a_1 f)^2} \longrightarrow \boxed{b_2 = \frac{a_2}{(1 + a_1 f)^3}}$$

3rd order terms

$$b_3 \cancel{S_{in}^3} = -a_1 f\, b_3 \cancel{S_{in}^3} - 2a_2 (1 - f\, b_1) f\, b_2 \cancel{S_{in}^3} + a_3 (1 - f\, b_1)^3 \cancel{S_{in}^3}$$

$$b_3 (1 + a_1 f) = a_3 \frac{1}{(1 + a_1 f)^3} - 2a_2 \frac{f}{1 + a_1 f} \cdot \frac{a_2}{(1 + a_1 f)^3}$$

$$\boxed{b_3 = \frac{a_3 (1 + a_1 f) - 2a_2^2 f}{(1 + a_1 f)^5}}$$

Now let's look at different distortion coefficients

no feedback $\quad HD2 = \frac{1}{2} IM2 = \frac{a_2}{2a_1} m_1$

with feedback $$HD2_{feedback} = \frac{b_2}{2b_1} m_1 = \frac{a_1/(1+a_1 f)}{2a_2/(1+a_1 f)^3} \cdot m_1$$

$$HD2_{feedback} = \frac{a_1}{2a_2} m_1 \Big/ (1+a_1 f)^2 = \boxed{\frac{HD2_{no-feedback}}{(1+a_1 f)^2}}$$

$\longrightarrow$ Dramatic reduction of HD2 (IM2) with feedback when measured based on same input power (m_1)

if HD2 is measured based on output power amplitude

(since total gain is reduced by $\dfrac{1}{1+a_1 f}$,

therefore $HD2_{feedback} = \dfrac{HD2_{no-feedback}}{1+a_1 f}$ at the same output power)

no feedback $$HD3 = \frac{1}{3} IM3 = \frac{1}{4}\frac{a_3}{a_1} m_1^2$$

with feedback

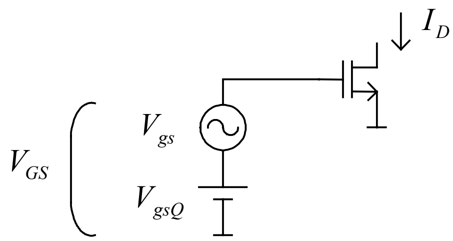
$$HD3_{feedback} = \frac{1}{4}\frac{b_3}{b_1} m_1^2 = \frac{1}{4} \frac{\dfrac{\overbrace{a_3(1+a_1 f) - 2a_2^2 f}}{(1+a_1 f)^5}}{\dfrac{a_1}{1+a_1 f}} m_1^2$$

$$= \frac{1}{4} \frac{\dfrac{a_3}{a_1}(1+a_1 f) - 2\dfrac{a_2^2}{a_1} f}{(1+a_1 f)^4} m_1^2$$

you can make $HD3_{feedback}$ *or* $IM3_{feedback} = 0$

by using right amount of feedback!

Distortion in MOS transistors



assume : ideal MOSFET

(no body effect)

body effect introduces third order terms!

$$I_D = \frac{\mu C_{OX}}{2}\frac{W}{L}(V_{GS} - V_{Th})^2$$

$$V_{GS} = V_{gsQ} + V_{gs}$$

$$I_D = \frac{\mu C_{OX}}{2}\frac{W}{L}\left[(V_{gsQ} - V_{Th})^2 + V_{gs}^2 + 2V_{gs}(V_{gsQ} - V_{Th})\right]$$

$$I_D = I_{DQ} + \frac{\mu C_{OX}}{2}\frac{W}{L}V_{gs}^2 + \underbrace{\frac{\mu C_{OX} W}{L}(V_{gsQ} - V_{Th})}_{g_m}\cdot V_{gs}$$

$$I_D = I_{DQ} + g_m V_{gs} + \frac{\mu C_{OX}}{2}\frac{W}{L}V_{gs}^2$$

$$a_1 = g_m = \frac{\mu C_{OX} W}{L}\left(V_{gsQ} - V_{Th}\right)$$

$$a_2 = \frac{\mu C_{OX}}{2}\frac{W}{L}$$

$$a_3 = 0$$

non-linear coefficients

$V_{gs} = \hat{V}_{gs}\cos\omega t$

$$HD2 = \frac{1}{2} IM2 = \frac{a_2}{2a_1} m_1 = \frac{\frac{\mu C_{OX}}{2}\frac{W}{L}\hat{V}_{gs}}{\frac{2\mu C_{OX} W}{L}\left(V_{gsQ} - V_{Th}\right)} = \frac{\hat{V}_{gs}}{4\left(V_{gsQ} - V_{Th}\right)}$$

$$\boxed{HD2 = \frac{\hat{V}_{gs}}{4\left(V_{gsQ} - V_{Th}\right)}}$$ in terms of input signal

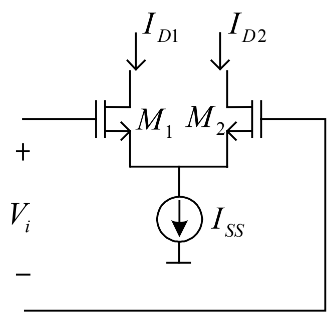
in terms of output signal $\hat{i}_D = g_m \hat{V}_{gs} = \hat{V}_{gs}\frac{\mu C_{OX} W}{L}\left(V_{gsQ} - V_{Th}\right)$

$$\boxed{HD2 = \frac{1}{8}\frac{\hat{i}_D}{I_{DQ}}}$$

→ ideal MOSFET does not show

* gain compression
* IM3, HD3 as long as it is in saturation!

MOS Differential Pair



$$I_{out} = I_{D1} - I_{D2}$$

* Assumptions :

- M1 & M2 identical
- ignore body effect
- ideal MOS

$$I_D = \frac{k'}{2}\left(\frac{W}{L}\right)\left[V_{gs} - V_{Th}\right]^2 \qquad \left(k' = \mu C_{OX}\right)$$

$$\left.\begin{aligned} V_{gs1} &= V_{Th} + \sqrt{\frac{2I_{D1}}{k'(W/L)}} \\ V_{gs2} &= V_{Th} + \sqrt{\frac{2I_{D2}}{k'(W/L)}} \end{aligned}\right\} \quad V_i = V_{gs1} - V_{gs2} = \frac{\sqrt{I_{D1}} - \sqrt{I_{D2}}}{\sqrt{\frac{k'}{2}\frac{W}{L}}} \qquad (1)$$

on the other hand $\quad I_{D1} + I_{D2} = I_{SS} \qquad (2)$

$(1),\ (2) \longrightarrow \quad find \ \ I_{D1},\ I_{D2} \quad V_S \quad V_i$

$$\longrightarrow \quad I_{out} = I_{D1} - I_{D2} = \frac{k'}{\cancel{2}}\frac{W}{L}V_i\sqrt{\frac{\cancel{4}I_{SS}}{k'(W/L)} - V_i^2}$$

$$I_{out} = \sqrt{k'\frac{W}{L}I_{SS}}\;V_i\left(1 - \frac{k'(W/L)}{4I_{SS}}V_i^2\right)^{1/2}$$

$$\boxed{\text{use power expansion} \quad (1+x)^j \approx 1 + jx + \frac{j\,x^2}{2!} + \cdots}$$

$$\longrightarrow \quad I_{out} = \sqrt{\frac{k'WI_{SS}}{L}}\left(V_i - \frac{k'(W/L)}{8I_{SS}}V_i^3 - \qquad V_i^5 \ \cdots\right)$$

$$I_{out} = a_1 V_i + a_2 V_i^2 + a_3 V_i^3 + \cdots$$

$$a_1 = \sqrt{\frac{k'W\,I_{SS}}{L}}$$

$$a_2 = 0$$

$$a_3 = \frac{-k'W}{2L}\sqrt{\frac{k'W}{2L \cdot I_{SS}}}$$

$$a_4 = 0$$

$$\longrightarrow \quad HD2 = \frac{1}{2}IM2 = 0$$

$$HD3 = \frac{1}{3}IM3 = \frac{1}{4}\frac{a_3}{a_1}m_1^2$$

$$\boxed{HD3 = \frac{-k'W/L}{8\sqrt{2}\,I_{SS}}\left|\hat{V}_i\right|^2}$$

$$when \ \hat{V}_i = 0 \quad \longrightarrow \quad I_{D1} = I_{D2} = \frac{I_{SS}}{2} = \frac{k'}{2}W/L\left[V_{gsQ} - V_{Th}\right]^2$$

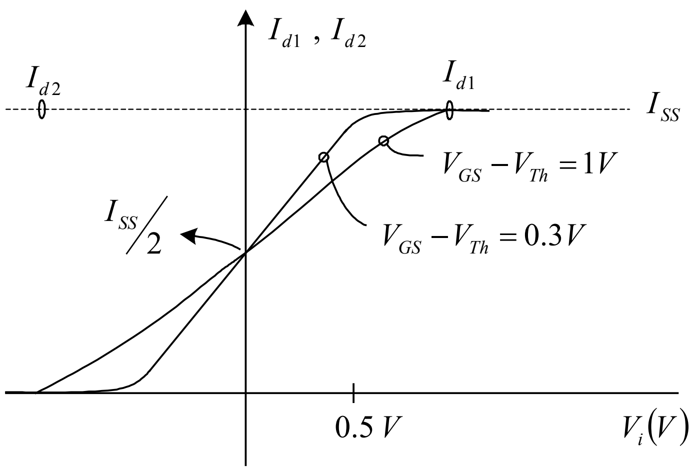
$$\longrightarrow \quad I_{SS} = \frac{k'W}{L}\left[V_{gsQ} - V_{Th}\right]^2$$

$$\longrightarrow \quad HD3 = -\frac{1}{8\sqrt{2}}\left(\frac{\hat{V}_i}{V_{gsQ} - V_{Th}}\right)^2$$

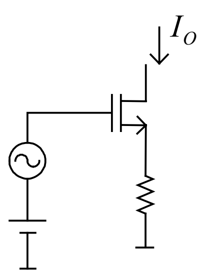
example :   *for* $HD3 = -40\,dB$,   what is the amplitude

of input signal ,   Assume $V_{gsQ} - V_{Th} \sim 1\,Volt$

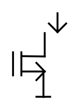
$HD3 = -40\,dB = -0.01 \longrightarrow \hat{V}_i = 0.34\,Volt$



<u>MOSFET with Rs</u>



$$I_O = \underbrace{I_{DQ}}_{\text{bias}} + \underbrace{i_d}_{\text{signal}}$$



+   feedback

$g_m R_S = f$

*** ignoring body effect**

$$i_d = b_1 V_{gs} + b_2 V_{gs}^2 + b_3 V_{gs}^3 + \cdots$$

$$b_1 = \frac{a_1}{1 + a_1 f} = \boxed{\frac{g_m}{1 + g_m R_S}}$$

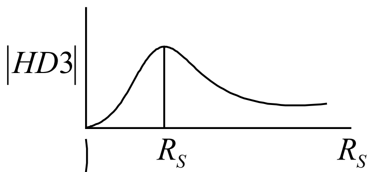
$$b_2 = \frac{a_2}{(1 + a_1 f)^3} = \boxed{\frac{\mu C_{OX} W/L}{2} \times \frac{1}{(1 + g_m R_S)^3}}$$

$$b_3 = \frac{\overset{0}{\cancel{a_3}}(1 + a_1 f) - 2a_2^2 f}{(1 + a_1 f)^5} = \frac{-2a_2^2 f}{(1 + a_1 f)^5} = \boxed{\frac{-2\left(\frac{\mu C_{OX} W/L}{2}\right)^2 R_S}{(1 + g_m R_S)^5}}$$

$$HD3 = \frac{1}{4}\frac{b_3}{b_1}\left|\hat{V}_{gs}\right|^2$$

$$= \frac{1}{4}\;\frac{-2\left(\frac{\mu C_{OX}}{2} W/L\right)^2 R_S \Big/ (1 + g_m R_S)^{\cancel{5}}}{g_m \Big/ \cancel{(1 + g_m R_S)}}\left|\hat{V}_{gs}\right|^2$$

$$HD3 = -\frac{1}{2}\;\frac{\left(\frac{\mu C_{OX}}{2} W/L\right) g_m R_S \left|\hat{V}_{gs}\right|^2}{(1 + g_m R_S)^4}$$



HD3 is zero for ideal MOS

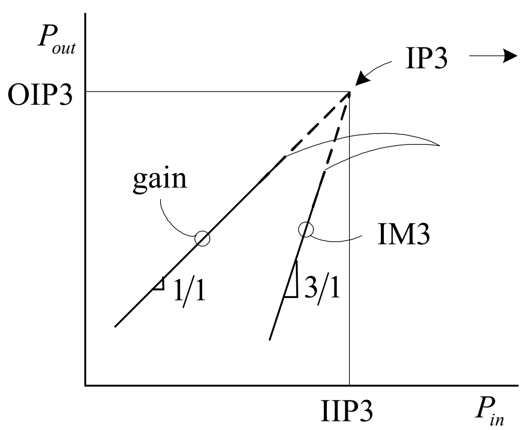
that gives worst HD3

$$\frac{\partial HD3}{\partial R_S} = 0$$

$$\boxed{R_S = \frac{1}{3g_m}}$$

Rs helps reducing HD2 but worsens HD3 !

Intercept Point



* IP3 is typically
10 ~ 15 dB higher than
1 dB compression point