

CHAPTER 5

NOISE AND DISTORSION

NOISE AND SIGNAL-TO-NOISE RATIO

Noise can be generally defined as any undesirable signal that masks or degrades the useful signal. This definition includes deterministic noise, due to coupling with the clock signal for example, and random noise whose origin is the fluctuation of a physical quantity such as voltage or current.

Deterministic noise can usually be drastically reduced or even eliminated by techniques such as shielding, filtering, or others. By its nature, random noise cannot be predicted and therefore cannot be eliminated. It can only be manipulated and reduced by techniques such as filtering.

We will use the term noise uniquely for random noise.

Noise is important because it represents the resolution limit for many systems.

Noise, being a random signal, is described by its statistical properties such as its amplitude probability distribution at a certain instant. In most cases it is not actually necessary to know the probability density, but only the first- and second-order moments corresponding to the average and the variance.

In many cases, noise can be considered stationary.

The noise of a circuit or system must always be compared to the signal that carries the useful information. We therefore introduce the notion of Signal-to-Noise Ratio (SNR) which is the ratio between the signal power and the noise power:

$$SNR \equiv \frac{\text{Average signal power}}{\text{Average noise power}} = \frac{P_{\text{signal}}}{P_N} \quad (5.1)$$

SPECTRAL NOISE FACTOR

The spectral noise factor of a two-port network as shown in Fig. 5-1 is defined as the ratio of the Power Spectral Density (PSD) of the maximum output noise, N_o , and the PSD of the noise due to the source resistance connected to the two-port network input, N_i , measured at a temperature of 290 K:

$$F(f) \equiv \frac{\text{PSD of total noise at output}}{\text{PSD at output due to source resistance } R_S} = \frac{N_o(f)}{G(f) \cdot N_i} \quad (5.2)$$

where $G(f)$ represents the gain in power of the two-port network and N_i is the PSD of the noise at the input of the two-port network due to the source.

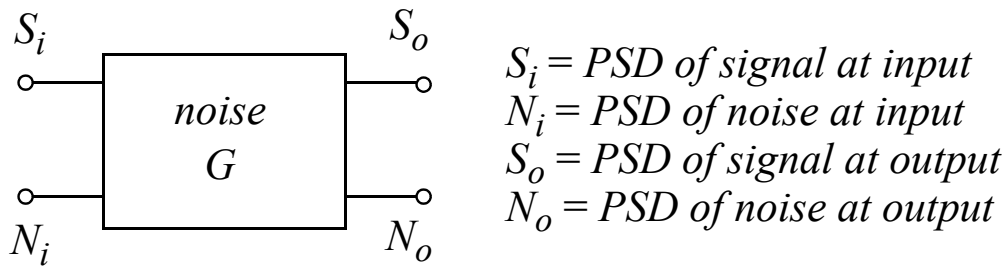


Fig 5-1: Two-port network with noise, and noise factor.

The contribution of basic noise at the output of the two-port network is thus given by:

$$N_o - GN_i = FGN_i - GN_i = (F - 1)GN_i \quad (5.3)$$

This basic contribution can be shifted to the two-port network input as a PSD N_p by dividing (5.3) by the gain G :

$$N_p = (F - 1)N_i \quad (5.4)$$

where $F - 1$ is the excess noise factor. The noise factor can then be written as:

$$F = 1 + \frac{N_p}{N_i} > 1 \quad (5.5)$$

The noise factor is thus always greater than one.

NOISE FIGURE AND AVERAGE NOISE FACTOR

We define the noise figure NF by:

$$NF \equiv 10 \log(F) > 0 \text{ dB} \quad (5.6)$$

The minimum (ideal) value of F being equal to 1, the minimum noise figure is equal to 0 dB. According to the definition given by (5.2), the noise factor is a function of frequency. It is generally defined for a set frequency. We can also define an average noise factor which takes into account the bandwidth B of a system:

$$\overline{F} \equiv \frac{\int_B N_o df}{\int_B G(f) N_i df} = \frac{\int_B F(f) G(f) N_i df}{\int_B G(f) N_i df} = \frac{\int_B F(f) G(f) df}{\int_B G(f) df} \quad (5.7)$$

which reduces to the average value of F when the gain is constant in the bandwidth B :

$$\overline{F} = \frac{1}{B} \int_B F(f) df \quad (5.8)$$

for $G(f) = G_0 = \text{const.}$ in the bandwidth B .

NOISE FACTOR AND SIGNAL-TO-NOISE RATIO

The noise factor can also be defined as a function of the signal-to-noise ratio (SNR) at the input SNR_i and at the output SNR_o :

$$SNR_i \equiv \frac{P_{Si}}{P_{Ni}} = \frac{\int_B S_i df}{\int_B N_i df} \quad \text{and:} \quad SNR_o \equiv \frac{P_{So}}{P_{No}} = \frac{\int_B G(f) S_i df}{\int_B N_o df} \quad (5.9)$$

Eqn. 5.7 can be rewritten:

$$\overline{F} = \frac{P_{No}}{\int_B G(f) N_i df} = \frac{1}{SNR_o} \frac{P_{So}}{\int_B G(f) N_i df} = \frac{1}{SNR_o} \frac{\int_B G(f) S_i df}{\int_B G(f) N_i df} \quad (5.10)$$

which for constant gain simplifies to:

$$\boxed{\overline{F} = \frac{1}{SNR_o} \frac{\int_B S_i df}{\int_B N_i df} = \frac{SNR_i}{SNR_o}} \quad (5.11)$$

The noise factor can thus be equivalently defined as the quotient of the signal-to-noise ratio at the input and the signal-to-noise ratio at the output. It is therefore a measure of the degradation of the signal-to-noise ratio at the output due to the basic noise generated by the two-port network.

NOISE TEMPERATURE

The power available from a source impedance Z_S corresponds to the maximum power that it can deliver to a load impedance Z_L . We know that this situation is achieved when the load impedance is equal to the complex conjugate of the source impedance:

$$Z_L = Z_S^* \quad (5.12)$$

In these conditions, the maximum power available from a source impedance R_S is given by:

$$P_S = \frac{V_S^2}{4R_S} \quad (5.13)$$

where V_S is the RMS source voltage. From this, we can deduce the power available from a source of thermal noise:

$$P_{Ni} = \frac{4kTR_S B}{4R_S} = kTB \quad (5.14)$$

We remark that this available thermal noise power is independent of the value of the source resistance. Thus, kTB is the maximum power available from any source that has an impedance with a resistive term. We also define the PSD available from a source in the bandwidth B as:

$$N_i = kT \quad (= 4 \times 10^{-21} J) \quad (5.15)$$

Note that this PSD is independent of the resistance value. The noise factor can then be expressed by:

$$F = 1 + \frac{N_a}{kT} \quad (5.16)$$

where T is the reference temperature (usually room temperature) and N_a the contribution of the two-port network.

NOISE TEMPERATURE

The noise factor is usually between 1 and 10. In certain cases, a larger scale is necessary. We use instead of the noise factor, the noise temperature T_r , defined by:

$$\boxed{F \equiv 1 + \frac{T_r}{T}} \quad (5.17)$$

From which: $T_r = (F - 1)T = N_a/k \quad (5.18)$

NOISE FACTOR OF CASCADING TWO-PORT NETWORKS

Consider the cascade connection of two 2-port networks, each characterized by its power gain and noise factor (cf Fig. 5-2).

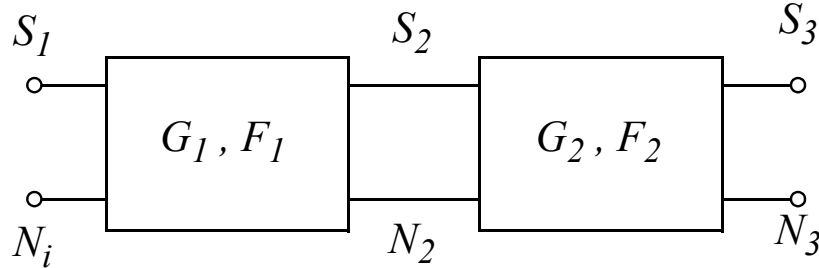


Fig 5-2: Noise factor of cascading two-port networks.
The global noise factor F is defined by:

$$F = \frac{N_3}{GN_i} = \frac{N_3}{G_1 G_2 N_i} \quad (5.19)$$

The noise at the output of the first two-port network is equal to the noise at the input N_i multiplied by the gain in power G_1 plus the contribution of the first two-port network at the output $(F_1 - 1)G_1 N_i$:

$$N_2 = G_1 N_i + (F_1 - 1)G_1 N_i = F_1 G_1 N_i \quad (5.20)$$

The noise at the output of the second two-port network is equal to the noise at the input N_2 multiplied by the gain in power G_2 plus the contribution of the second two-port network at the output $(F_2 - 1)G_2 N_i$:

$$N_3 = G_2 N_2 + (F_2 - 1)G_2 N_i = G_2 F_1 G_1 N_i + (F_2 - 1)G_2 N_i \quad (5.21)$$

From which:
$$F = F_1 + \frac{(F_2 - 1)G_2 N_i}{G_2 G_1 N_i} = F_1 + \frac{F_2 - 1}{G_1} \quad (5.22)$$

For the case in which $G_1 \gg (F_2 - 1)$, $F \cong F_1$ and the global noise factor is essentially determined by the first stage of the cascade. Eqn. 5.22 can be easily generalized for the case of n two-port networks in cascade (Friis Formula):

$$\left| F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \dots + \frac{F_n - 1}{G_1 G_2 \dots G_{n-1}} \right| \quad (5.23)$$

SENSITIVITY AND MINIMUM DETECTABLE SIGNAL

The power of the input signal P_{Si} corresponding to a given signal-to-noise ratio at the output SNR_o is called the sensitivity of the system. The level (in dBm) corresponding to P_{Si} is called the minimum detectable signal (MDS).

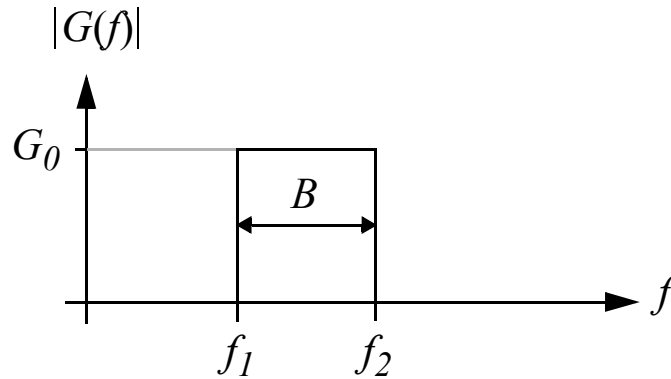


Fig 5-3: Gain in power of an ideal system.

For a system whose gain is constant in a frequency bandwidth B and zero beyond this band (cf Fig. 5-3), the average noise factor is given by:

$$\bar{F} = \frac{1}{SNR_o} \frac{\int_B G(f) S_i df}{\int_B G(f) N_i df} = \frac{G_0 \int_B S_i df}{SNR_o G_0 \int_B N_i df} = \frac{P_{Si}}{SNR_o kTB} \quad (5.24)$$

From this we get the signal power for a given signal-to-noise ratio at the output SNR_o and a given average noise factor:

$$P_{Si} = \bar{F} \cdot kTB \cdot SNR_o \quad (5.25)$$

or in terms of the input signal level in dBm:

$$L_{min} \equiv 10 \log \left(\frac{P_{Si}}{1mW} \right) = \bar{NF} + 10 \log \left(\frac{kTB}{1mW} \right) + (SNR_o)_{dB} \quad [dBm]$$

(5.26)

EQUIVALENT NOISE SOURCES

A two-port network with noise can be modeled by the same two-port network without internal noise sources, and two noise sources V_N and I_N , independent of the values of the source resistance R_S and the load resistance R_L (cf Fig. 5-4).

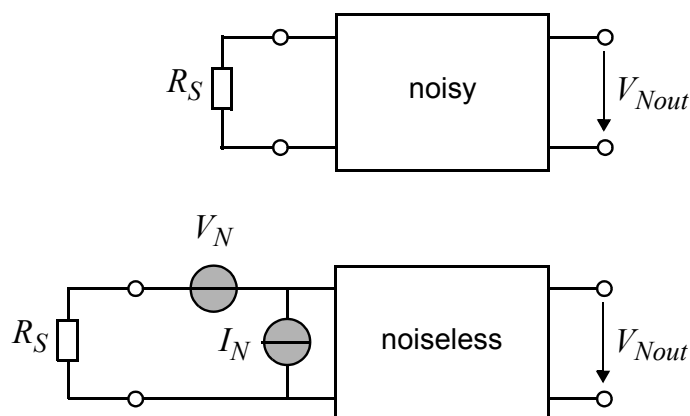


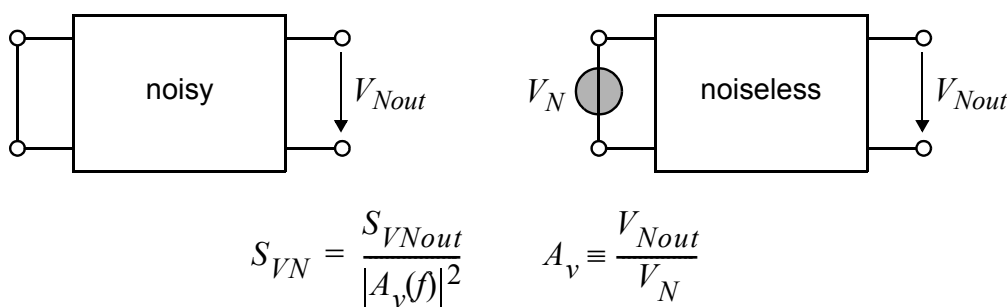
Fig 5-4: *Noisy two-port network and its noiseless model, with equivalent noise sources at the input.*

Note that the two noise sources are necessary in order to have a complete description of the two-port network noise for all source resistance values. In fact, when $R_S = 0$, the noise at the output V_{Nout} is due only to the noise source V_N , while if $R_S \rightarrow \infty$, it is due to the current noise source I_N . Since each of these sources considers the effects of the same physical causes of noise, internal to the two-port network, they are usually not independent. But in most cases, the correlation between V_N and I_N is weak and can be neglected.

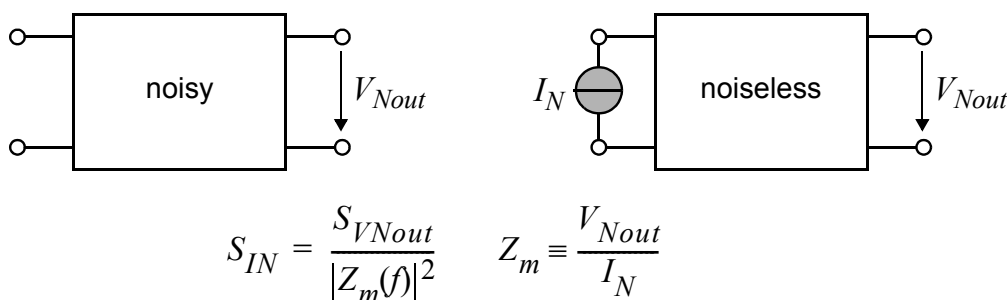
CALCULATION OF EQUIVALENT NOISE SOURCES (1/2)

Voltage amplifier

The PSD S_{VN} of the source V_N of a voltage amplifier is calculated by evaluating the PSD of the output noise voltage of the two-port network when the input is short-circuited, and dividing it by the square of the voltage gain A_v (cf Fig. 5-5 a). The PSD S_{IN} of the source I_N is obtained by evaluating the PSD of the output noise voltage of the two-port network when the input is an open circuit, and dividing it by the square of the transimpedance Z_m (cf Fig. 5-5 b).



a) Source of noise voltage.



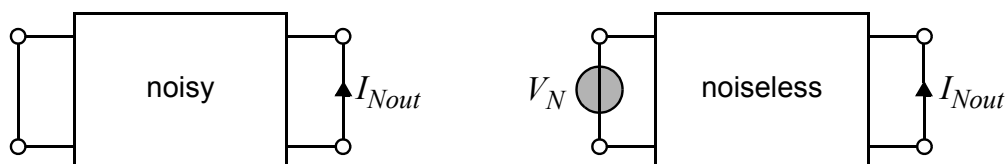
b) Source of noise current.

Fig 5-5: Calculation of PSD of equivalent noise sources of a voltage amplifier.

CALCULATION OF EQUIVALENT NOISE SOURCES (2/2)

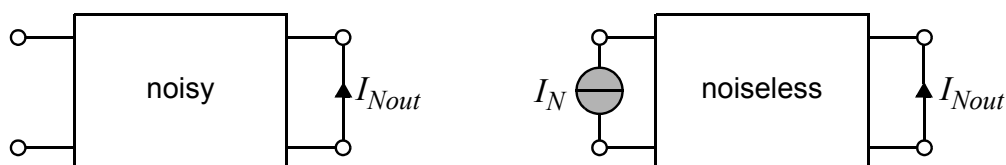
Transconductance amplifier

The PSD S_{VN} of the source V_N of a transconductance amplifier is calculated by evaluating the PSD of the output noise current of the two-port network when the input is short-circuited, and dividing it by the square of the transadmittance Y_m (cf Fig. 5-5 a). The PSD S_{IN} of the source I_N is obtained by evaluating the PSD of the output noise current of the two-port network when the input is an open circuit, and dividing by the square of the current gain A_i (cf Fig. 5-5 b).



$$S_{VN} = \frac{S_{INout}}{|Y_m(f)|^2} \quad Y_m \equiv \frac{I_{Nout}}{V_N}$$

a) Source of noise voltage.



$$S_{IN} = \frac{S_{INout}}{|A_i(f)|^2} \quad A_i \equiv \frac{I_{Nout}}{I_N}$$

b) Source of noise current.

Fig 5-6: *Calculation of PSD of equivalent noise sources of a transconductance amplifier.*

EQUIVALENT NOISE VOLTAGE SOURCE FOR THE BIPOLAR TRANSISTOR

The equivalent noise sources for the bipolar transistor can be calculated from the small-signal model shown in Fig. 5-7.

$$S_{VNB} = 4kTr_{bb'} \quad S_{INB} = 2qI_B + K_f \frac{I_B^{AF}}{f} \quad S_{INC} = 2qI_C$$

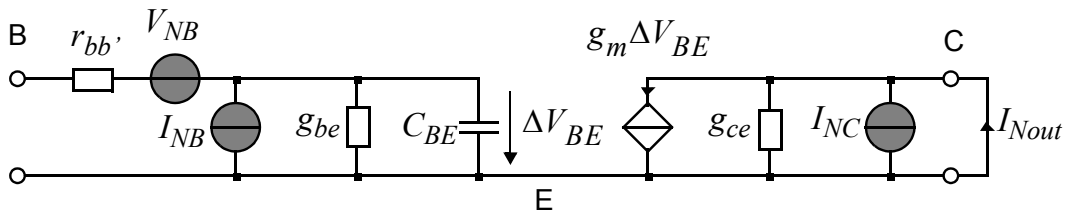


Fig 5-7: Small-signal model of the bipolar transistor, including the noise sources.

By considering that $g_{ce} \ll g_m$ and $g_{be}r_{bb'} \ll 1$, the low-frequency noise current when the input is short-circuited is given by:

$$I_{Nout} = I_{NC} + g_m \Delta V_{BE} = I_{NC} - g_m V_{NB} \quad (5.27)$$

The PSD of the current I_{Nout} is thus:

$$S_{INout} = S_{INC} + g_m^2 S_{VNB} \quad (5.28)$$

from which we find the PSD of the source V_N :

$$\begin{aligned} S_{VN} &= \frac{S_{INout}}{g_m^2} = \frac{S_{INC}}{g_m^2} + S_{VNB} = \frac{2qI_C}{g_m^2} + 4kTr_{bb'} \\ &= 4kT \left(r_{bb'} + \frac{1}{2g_m} \right) = 4kTR_N \end{aligned} \quad (5.29)$$

with:

$$R_N \equiv r_{bb'} + \frac{1}{2g_m} \quad (5.30)$$

EQUIVALENT NOISE CURRENT SOURCE FOR THE BIPOLAR TRANSISTOR

The noise current when the input is an open circuit is:

$$I_{Nout} = I_{NC} + \frac{g_m \cdot I_{NB}}{g_{be} + j\omega C_{BE}} = I_{NC} + \beta(j\omega) I_{NB} \quad (5.31)$$

Since the current gain is equal to $\beta(j\omega)$, the PSD of the source I_N is:

$$S_{IN} = S_{INB} + \frac{S_{INC}}{|\beta(j\omega)|^2} = 2q \left(I_B + \frac{I_C}{|\beta(j\omega)|^2} \right) + K_f \frac{I_B^{AF}}{f} \quad (5.32)$$

The current gain as a function of the frequency is given by:

$$\beta(f) = \frac{\beta_F}{1 + j\beta_F(f/f_T)} \quad (5.33)$$

so
$$S_{IN} = 2q \left(I_B + \frac{I_C}{\beta_F^2} \left(1 + \beta_F^2 \left(\frac{f}{f_T} \right)^2 \right) \right) + K_f \frac{I_B^{AF}}{f} \cong \quad (5.34)$$

$$2qI_C \left(\frac{1}{\beta_F} + \left(\frac{f}{f_T} \right)^2 \right) + K_f \frac{I_B^{AF}}{f}$$

The PSD of the source I_N as a function of the frequency is shown in Fig. 5-8.

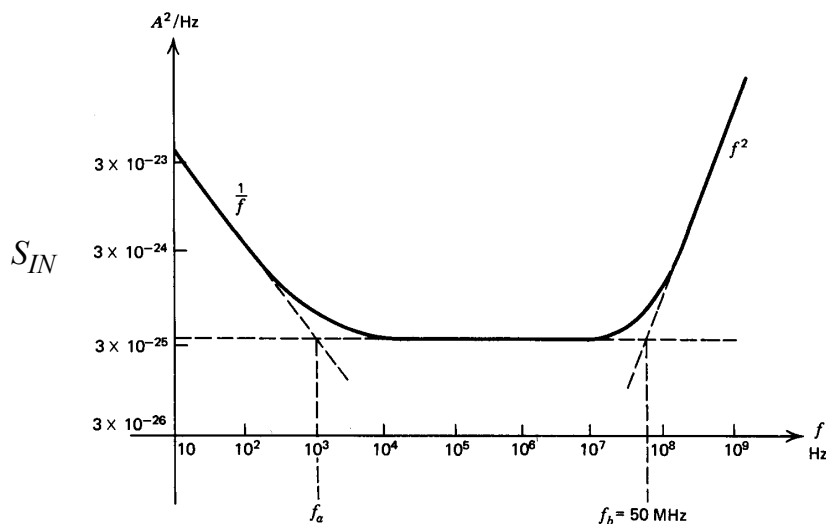


Fig 5-8: PSD of the equivalent noise current source I_N .

EQUIVALENT NOISE SOURCE OF AN AMPLIFIER CONNECTED TO A SOURCE

Amplifier noise associated with a voltage source can be modeled by a single noise source V_{Neq} (cf Fig. 5-9), including the noise from the source resistance $4kTR_S$ and the noise from the amplification device.

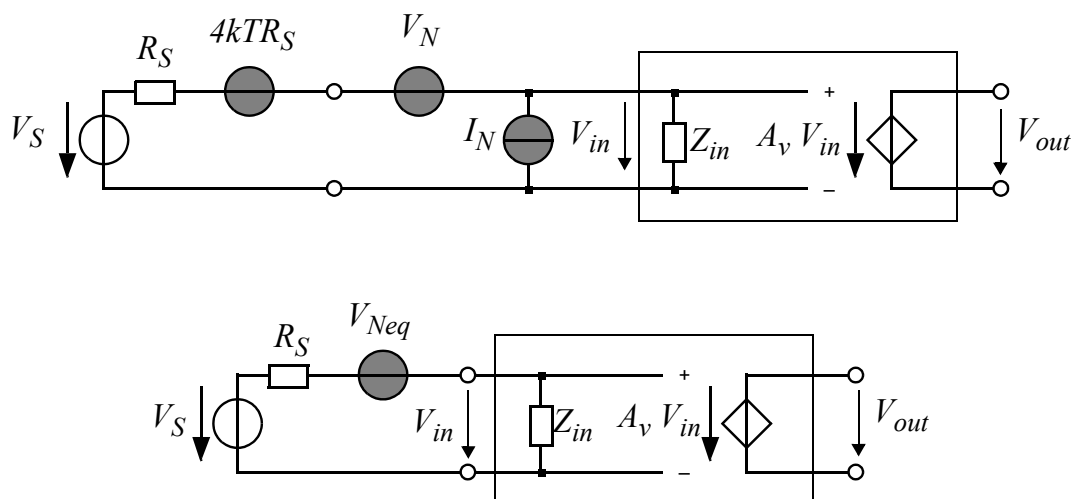


Fig 5-9: Noisy amplifier and equivalent noise sources.

The PSD of the output noise is given by:

$$S_{VNout} = |A_v(f)|^2 \left\{ \underbrace{\left| \frac{Z_{in}}{Z_{in} + R_S} \right|^2 S_{VN}}_{\text{contribution of the amplifier}} + \underbrace{\left| \frac{R_S Z_{in}}{Z_{in} + R_S} \right|^2 S_{IN}}_{\text{contribution of the source}} + \left| \frac{Z_{in}}{Z_{in} + R_S} \right|^2 4kTR_S \right\} \quad (5.35)$$

The voltage gain between the source and the output being equal to $(Z_{in}/(Z_{in} + R_S))A_v(f)$, the PSD of the equivalent noise voltage source at the input is thus given by:

$$\boxed{S_{VNeq} = 4kTR_S + S_{VN} + R_S^2 S_{IN}} \quad (5.36)$$

This PSD is independent of the parameters Z_{in} and A_v of the amplifier, but depends on the source resistance R_S .

NOISE FACTOR OF AN AMPLIFIER AND OPTIMUM SOURCE RESISTANCE

The spectral noise factor of the amplifier in Fig. 5-9 is simply given by the ratio between S_{VNeq} and the contribution of the source $4kTR_S$:

$$F = \frac{S_{VNeq}}{4kTR_S} = 1 + \frac{S_{VN}}{4kTR_S} + \frac{S_{IN}}{4kT(1/R_S)} \quad (5.37)$$

This noise factor is independent of the amplifier parameters Z_{in} and A_v . In addition, it has a minimum for a value R_{Sopt} of the source resistance:

$$R_{Sopt} = \sqrt{\frac{S_{VN}}{S_{IN}}} \quad (5.38)$$

The corresponding minimum noise factor is thus:

$$F_{opt} = 1 + \frac{\sqrt{S_{VN} \cdot S_{IN}}}{2kT} \quad (5.39)$$

The noise figure corresponding to (5.37) is represented as a function of R_S in Fig. 5-10.

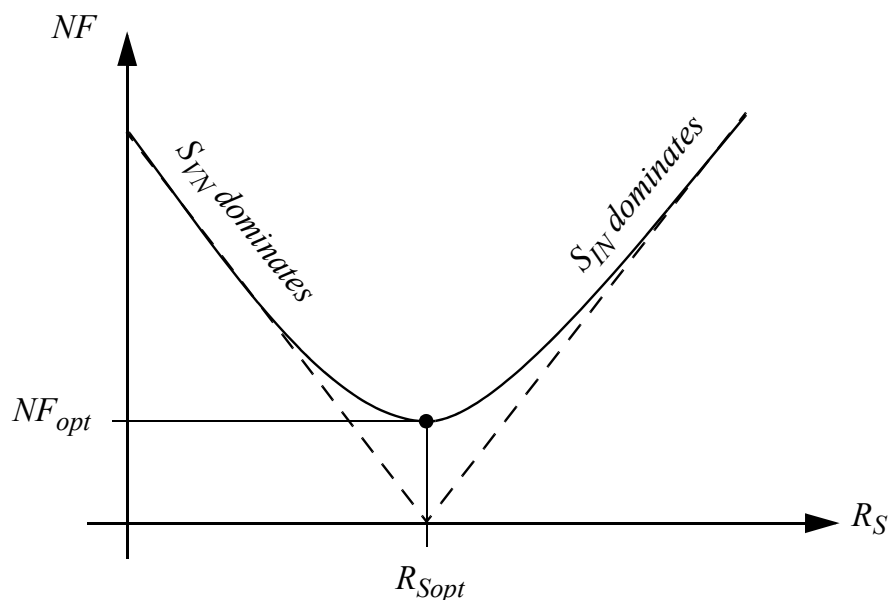


Fig 5-10: Noise figure corresponding to (5.37).

OPTIMUM SOURCE RESISTANCE OF BIPOLAR AND MOS TRANSISTORS

The optimum source resistance and the minimum noise factor for a bipolar transistor (neglecting the 1/f noise, that is by setting $K_f = 0$), are calculated from:

$$S_{VN} = 4kT \left(r_{bb'} + \frac{1}{2g_m} \right) \quad S_{IN} = 2qI_B \quad (5.40)$$

We find:

$$R_{Sopt} = \sqrt{\beta_F} \frac{\sqrt{1 + 2g_m r_{bb'}}}{g_m} \quad F_{opt} = 1 + \frac{\sqrt{1 + 2g_m r_{bb'}}}{\sqrt{\beta_F}} \quad (5.41)$$

From (5.41) we deduce that a low-noise bipolar transistor must have a small base resistance $r_{bb'}$, and a large current gain β_F .

The optimum source resistance and the minimum noise factor for a MOS transistor (without 1/f noise) are calculated from the PSD of the equivalent noise sources at the input:

$$S_{VN} = 4kT \frac{\gamma}{g_m} \quad S_{IN} \cong 0 \quad (5.42)$$

$$\text{from which:} \quad R_{Sopt} \rightarrow \infty \quad F_{opt} \rightarrow 1 \quad (5.43)$$

The MOS transistor is thus well matched for large source resistances.

In practice, for source resistances higher than 1 M Ω , the MOS transistor offers a noise factor for white noise that is smaller than that of the bipolar (for the same transconductance). But the noise factor can be deteriorated by the presence of 1/f noise, which is generally higher in a MOS transistor.

IMPEDANCE MATCHING FOR MINIMUM NOISE

If the source resistance R_S is fixed, it is useless to try to obtain the condition $R_S = R_{opt}$ by adding series or parallel resistors. In fact, they would only worsen the noise of the system. We can, however, obtain the minimum noise factor by using an (ideal) transformer of ratio n as indicated in Fig. 5-11.

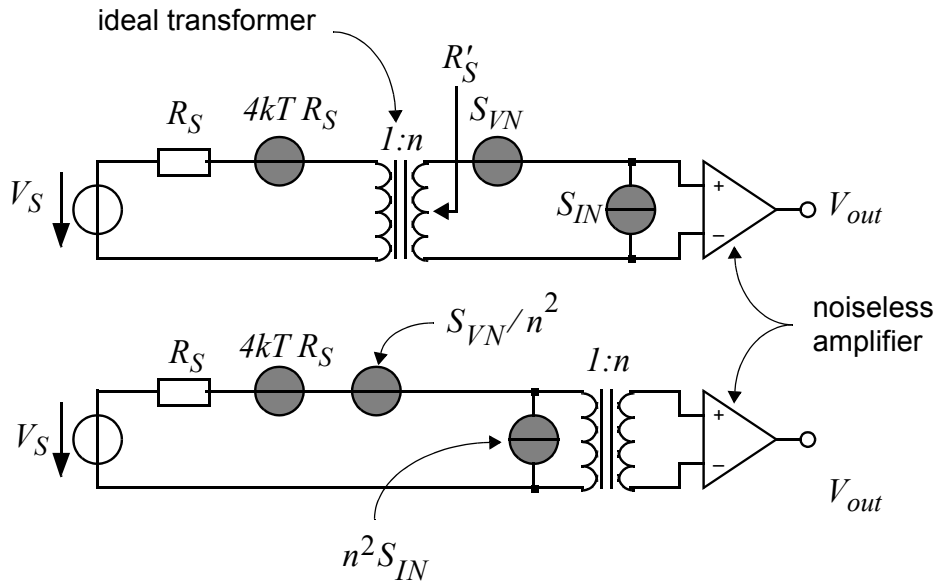


Fig 5-11: *Source impedance matching for a minimum noise factor.*

The PSD of the equivalent noise at the input is then given by:

$$S_{VNeq} = 4kTR_S + \frac{S_{VN}}{n^2} + n^2 \cdot R_S^2 S_{IN} \quad (5.44)$$

This PSD is minimum for a transformation ratio n_{opt} :

$$n_{opt}^2 = \frac{R'_S}{R_S} = \frac{R_{Sopt}}{R_S} = \frac{1}{R_S} \sqrt{\frac{S_{VN}}{S_{IN}}} \quad (5.45)$$

$$S_{VNeq} = 4kTR_S + 2R_S \sqrt{S_{VN} S_{IN}} = 4kTR_S \left(1 + \frac{\sqrt{S_{VN} S_{IN}}}{2kT} \right) = 4kTR_S \cdot F_{opt} \quad (5.46)$$

This optimum ratio also gives the minimum noise factor.

IMPEDANCE MATCHING BY AUTOTRANSFORMER

The impedance level of narrowband circuits can be modified with an autotransformer as indicated in Fig. 5-12.

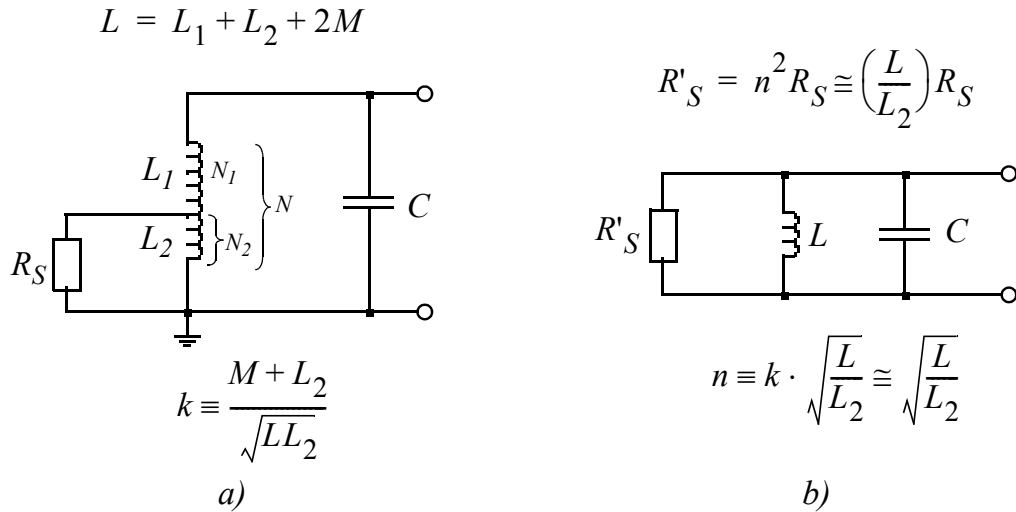


Fig 5-12: Changing the impedance level by using an autotransformer.

For perfect coupling, the coupling factor $k \cong 1$. The diagram in Fig. 5-12 a) can then be replaced by that of Fig. 5-12 b), which shows that the source impedance is multiplied by the square of the transformation ratio n .

$$n = \sqrt{\frac{L}{L_2}} \quad (5.47)$$

The impedance as seen from the secondary is therefore:

$$R'_S = n^2 R_S \cong \left(\frac{L}{L_2}\right) R_S \quad (5.48)$$

The noise can thus be minimized by choosing n according to (5.45).

IMPEDANCE MATCHING BY CAPACITIVE DIVIDER

For the case in which the input impedance of the amplifier stage is inductive, the source impedance level can be changed by using the circuit shown in Fig. 5-13 if $\omega \gg (R_S C_2)^{-1}$.

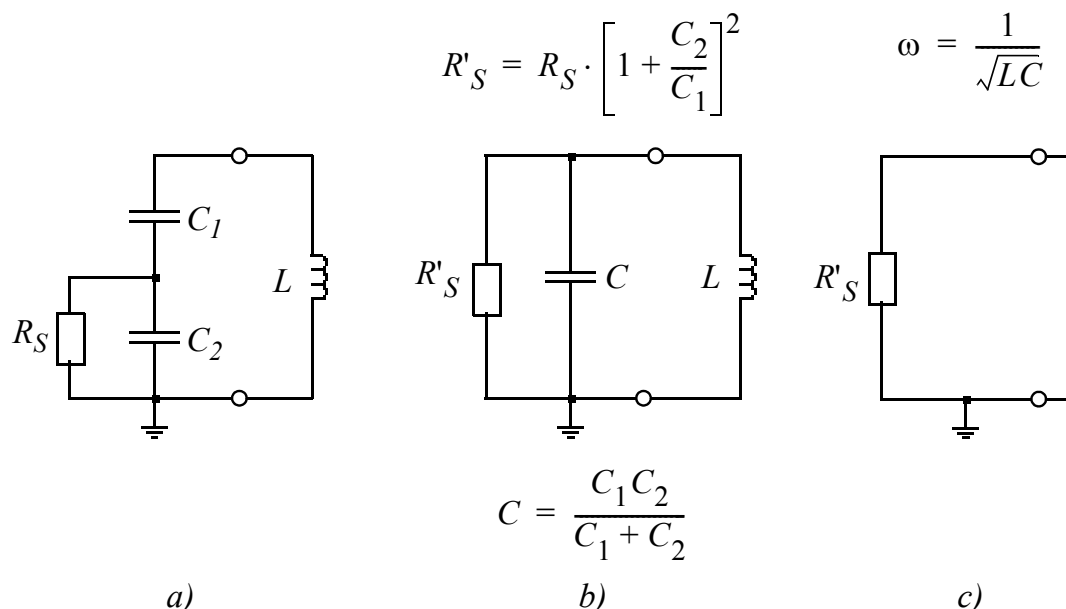


Fig 5-13: *Changing the impedance level with a capacitive divider.*

For frequencies $\omega \gg (R_S C_2)^{-1}$, the resistance seen by the inductance equals the source resistance multiplied by n^2 :

$$R'_S = n^2 \cdot R_S = \left[1 + \frac{C_2}{C_1} \right]^2 \cdot R_S \quad (5.49)$$

The noise can thus be minimized by choosing n according to (5.45).

TOTAL EQUIVALENT INPUT NOISE FOR A RESISTIVE SOURCE

As indicated in Fig. 5-14, the block made up of an amplifier, of which the noise is modeled by two noise sources S_{VN} and S_{IN} , and an impedance source R_S , can be modeled by a single noise voltage source in series with the source V_S with PSD S_{VNeq} :

$$S_{VNeq} = 4kTR_S + S_{VN} + R_S^2 \cdot S_{IN} \quad (5.50)$$

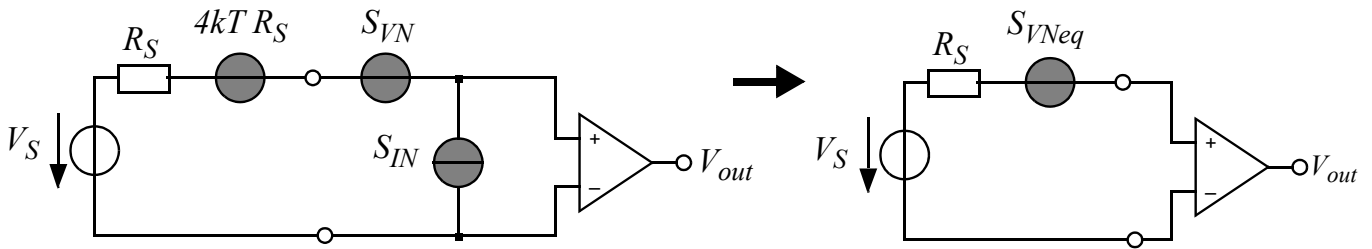


Fig 5-14: Total equivalent noise source at the input.

Notice that this PSD is minimum when $R_S \rightarrow 0$ and not for $R_S = R_{opt}$, value for which the contribution of S_{VN} is equal to that of S_{IN} :

$$S_{VNeq} \big|_{R_S=R_{Sopt}} = 4kTR_{Sopt} + 2 \cdot S_{VN} \quad (5.51)$$

Eqn. 5.50 is graphed as a function of the source resistance R_S in Fig. 5-15.

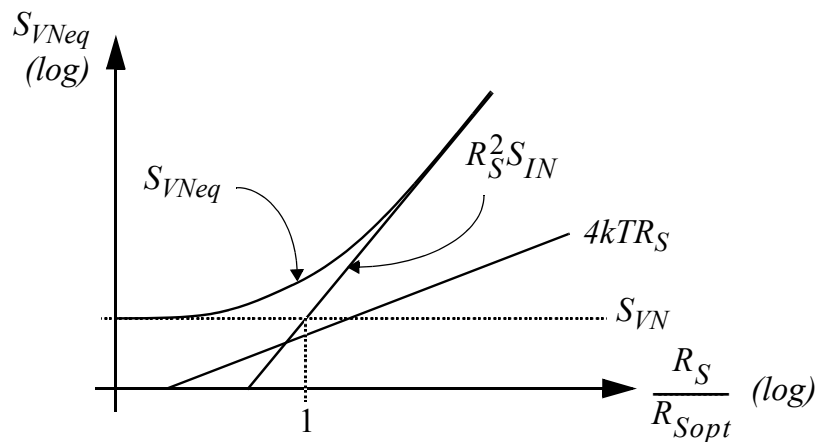


Fig 5-15: S_{VNeq} as a function of R_S .

VOLTAGE AMPLIFIER WITH RESISTIVE SOURCE (1/4)

Equivalent circuit

Consider the voltage amplifier presented in Fig. 5-16.

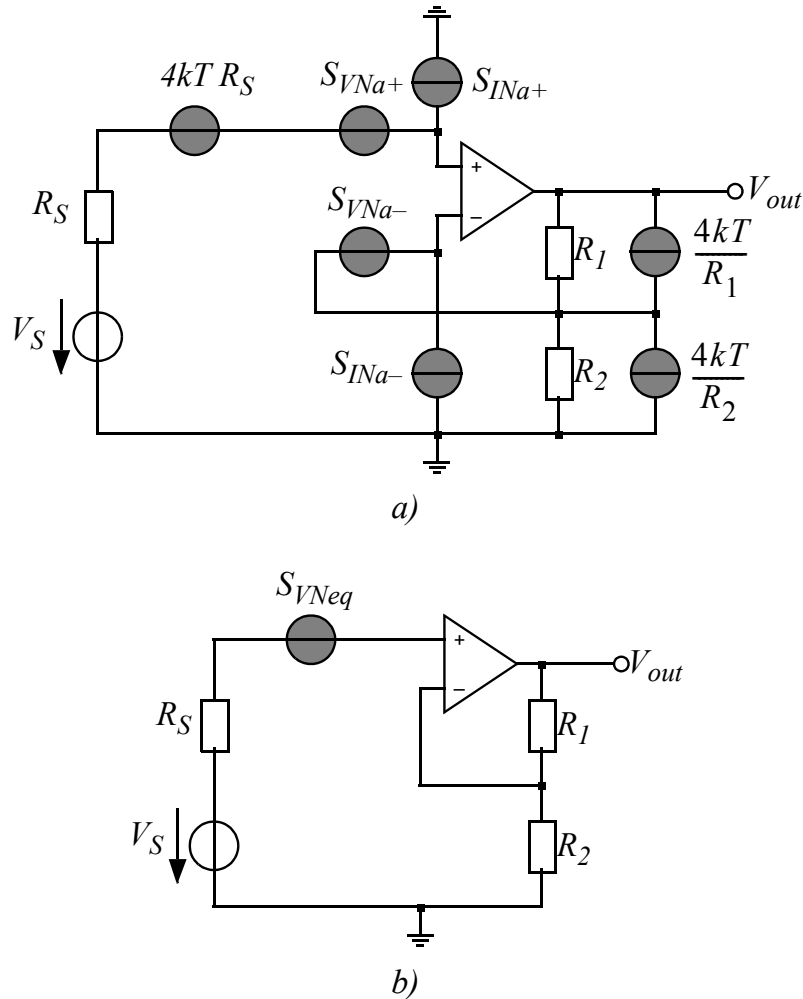


Fig 5-16: Voltage amplifier and equivalent noise model.

Ignoring the correlations that exist between S_{VNa+} and S_{INa+} (S_{VNa-} and S_{INa-}), the PSD of the equivalent noise voltage at the input is given by:

$$S_{VNeq} = 4kT(R_S + R_{12}) + S_{VNa+} + S_{VNa-} + R_S^2 S_{INa+} + R_{12}^2 S_{INa-} \quad (5.52)$$

with:

$$R_{12} = R_1 \parallel R_2 \quad (5.53)$$

VOLTAGE AMPLIFIER WITH RESISTIVE SOURCE (2/4)

Bipolar input stage

For the case in which the input stage of the amplifier is realized with bipolar transistors, the PSD S_{VNa+} and S_{INa+} are given by the equations (5.29) and (5.34). Ignoring the $1/f$ noise ($K_f = 0$), we find the PSD S_{VNeq} :

$$\begin{aligned} S_{VNeq} &= 4kT \left(R_S + R_{12} + 2r_{bb'} + \frac{U_T}{I_C} \right) + (R_S^2 + R_{12}^2) 2qI_C \left(\frac{1}{\beta_F} + \left(\frac{f}{f_T} \right)^2 \right) \\ &= 4kT \left(R_S + R_{12} + 2r_{bb'} + \frac{U_T}{I_C} + (R_S^2 + R_{12}^2) \frac{I_C}{2\beta_F U_T} \right) \end{aligned} \quad (5.54)$$

for $f \ll f_T$ and $\beta_F \gg I$. Notice that S_{VNeq} has a minimum for one particular value of the polarization current I_{Copt} :

$$I_{Copt} \equiv \sqrt{\frac{2\beta_F}{R_S^2 + R_{12}^2}} \cdot U_T \quad g_{mopt} \equiv \sqrt{\frac{2\beta_F}{R_S^2 + R_{12}^2}} \quad (5.55)$$

The minimum value of the PSD S_{VNeq} is thus given by:

$$S_{VNeqopt} = 4kT \underbrace{\left(R_S + R_{12} + 2r_{bb'} + 2 \sqrt{\frac{R_S^2 + R_{12}^2}{2\beta_F}} \right)}_{= R_{Neq}} = 4kT \cdot R_{Neq} \quad (5.56)$$

VOLTAGE AMPLIFIER WITH RESISTIVE SOURCE (3/4)

Optimum polarization current

The equivalent noise resistance R_{Neq} normalized to the source resistance R_S is graphed in Fig. 5-17 as a function of the polarization current for a bipolar transistor. For $r_{bb'} = R_S = R_I = 50 \Omega$ and $\beta_F = 200$, we find $I_{Copt} = 7.35 \text{ mA}$.

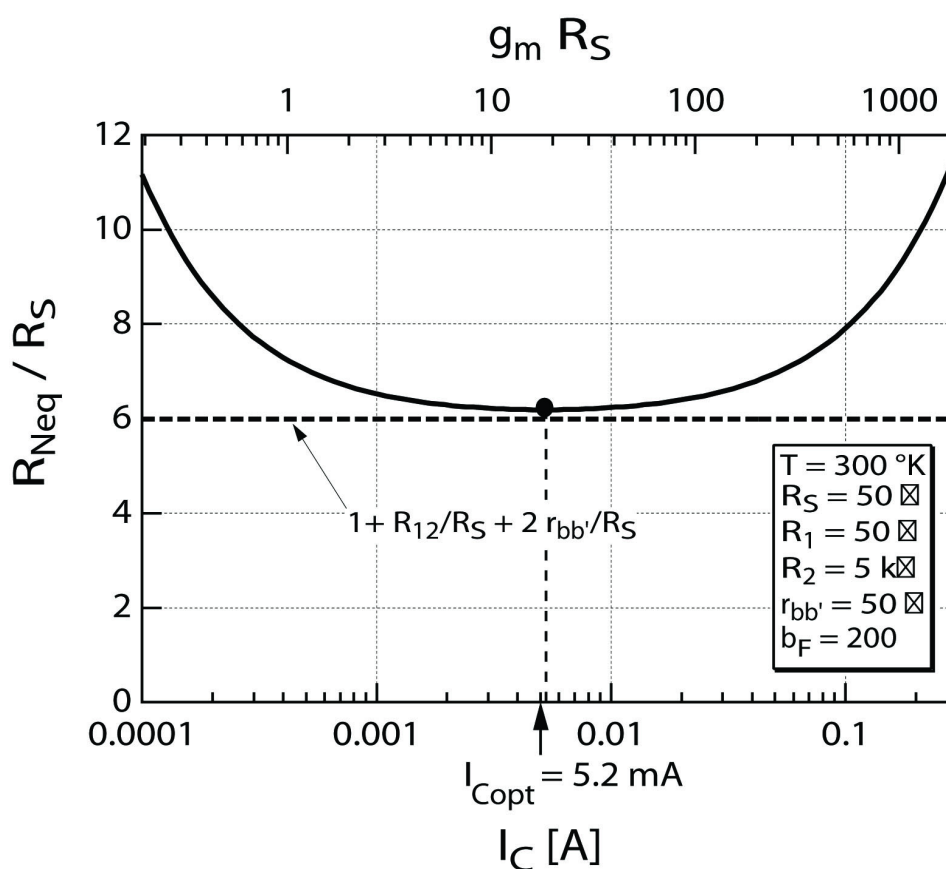


Fig 5-17: Optimum polarization current.

VOLTAGE AMPLIFIER WITH RESISTIVE SOURCE (4/4)

MOS input stage

For the case in which the input transistors are MOS transistors, we have:

$$S_{VNa+} = S_{VNa-} = 4kT \left(\frac{\gamma}{g_m} + \frac{K_f}{WLf} \right) \quad (5.57)$$

$$S_{INa+} = S_{INa-} = (\omega C_{in})^2 S_{VNa+}$$

The equivalent PSD of the input noise is thus given by:

$$S_{VNeq} = 4kT \left\{ R_S + R_{12} + \left[\frac{\gamma}{g_m} + \frac{K_f}{WLf} \right] [2 + (\omega C_{in})^2 (R_S^2 + R_{12}^2)] \right\} \quad (5.58)$$

which simplifies at low frequency ($\omega \ll 1 / (C_{in} \sqrt{R_S^2 + R_{12}^2})$) to:

$$S_{VNeq} = 4kT \left(R_S + R_{12} + 2 \left[\frac{\gamma}{g_m} + \frac{K_f}{WLf} \right] \right) \quad (5.59)$$

In this case, there is no optimum as a function of polarization. The noise contribution of the MOS transistors can be made negligible compared to the noise term due to the resistances, by reducing S_{VNa+} and S_{VNa-} . This is achieved by increasing the transconductance and the gate area WL of the MOS transistor. The frequency for which the white noise is equal to the $1/f$ noise (the corner frequency) is given by:

$$f_k = \frac{2K_f}{WL \left(R_S + R_{12} + \frac{2\gamma}{g_m} \right)} \quad (5.60)$$

HARMONIC DISTORSION

Single tone input signal

Although components such as amplifiers and transistors are often considered as linear elements, they have nonlinear transfer characteristics. Fig. 5-18 shows the typical transfer characteristic of an amplifier. It is made up of a linear portion and two saturation zones. We apply a sinusoidal input signal:

$$x(t) = X_1 \cos(\omega_0 t) \quad \text{or} \quad x(\phi) = X_1 \cos(\phi) \quad (5.61)$$

with $\phi \equiv \omega_0 t$. As long as the signal amplitude is less than X_{max} , the output signal is likewise sinusoidal with the same frequency. When $X_1 > X_{max}$, the output signal will be subject to distortion. It then contains frequency components which are multiples of ω_0 or *harmonics*. The harmonic components will depend on the input amplitude, the maximum amplitude X_{max} , and the nonlinear characteristics.

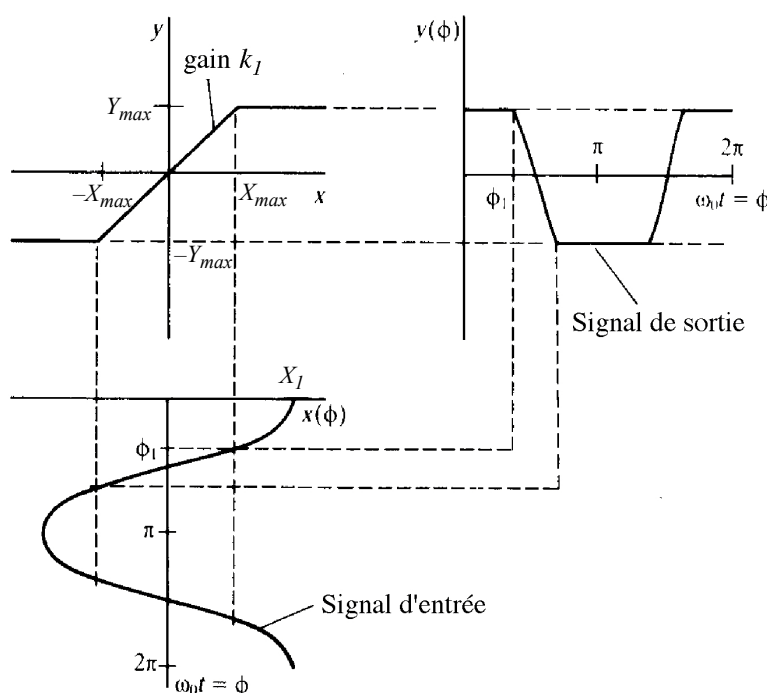


Fig 5-18: *Transfer characteristic of an amplifier.*

FOURIER SERIES OF THE OUTPUT SIGNAL

The output signal $y(t)$, remaining periodic, can be broken down into a Fourier series:

$$y(\phi) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos(n\phi) + b_n \sin(n\phi) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} r_n \cos(n\phi - \alpha_n) \quad (5.62)$$

The coefficients a_n , b_n , r_n and α_n are given by:

$$\begin{aligned} a_n &\equiv \frac{1}{\pi} \int_{-\pi}^{\pi} y(\phi) \cos(n\phi) d\phi & \text{and:} & & b_n &\equiv \frac{1}{\pi} \int_{-\pi}^{\pi} y(\phi) \sin(n\phi) d\phi \\ r_n &= a_n^2 + b_n^2 & \text{and:} & & \alpha_n &= \text{atan}(b_n/a_n) \end{aligned} \quad (5.63)$$

In the case of the saturation characteristic of Fig. 5-18, we remark that the average value of the output signal is zero and thus $a_0 = 0$. In addition, the output signal $y(\phi)$ shown in Fig. 5-18 being an even function of ϕ , the coefficients b_n are all zero. Since the transfer characteristic is odd, the output only contains odd harmonics. It can be shown that the coefficient corresponding to the fundamental is given by:

$$a_1 = k_1 \cdot X_1 \cdot f(\xi) \quad \text{with:} \quad \xi \equiv X_1/X_{max} \quad (5.64)$$

$$\text{where:} \quad f(\xi) = \begin{cases} \frac{2}{\pi} \left[\text{asin}(1/\xi) + \frac{\sqrt{1 - (1/\xi)^2}}{\xi} \right] & \xi > 1 \\ 1 & \xi \leq 1 \end{cases} \quad (5.65)$$

When $X_1 \leq X_{max}$, the system is linear, and therefore there are no harmonics. In this case the output signal is simply given by $y(\phi) = k_1 \cdot x(\phi) = k_1 \cdot X_1 \cos(\phi)$ and so $a_1 = k_1 \cdot X_1$ which corresponds to $f(\xi) = 1$.

The amplitude of the harmonics greater than 1 is given by:

$$a_n = \frac{2k_1 X_1}{n\pi} \left[\frac{\sin((n+1)\phi)}{n+1} - \frac{\sin((n-1)\phi)}{n-1} \right] \quad \text{for: } n = 3, 5, 7, \dots \quad (5.66)$$

COMPRESSION

The function $f(\xi)$ given by (5.65) and permitting the calculation of the amplitude of the fundamental is shown in Fig. 5-19 as a function of ξ . This figure shows that the amplitude of the output signal fundamental decreases as a function of the input signal amplitude. This phenomenon is called compression.

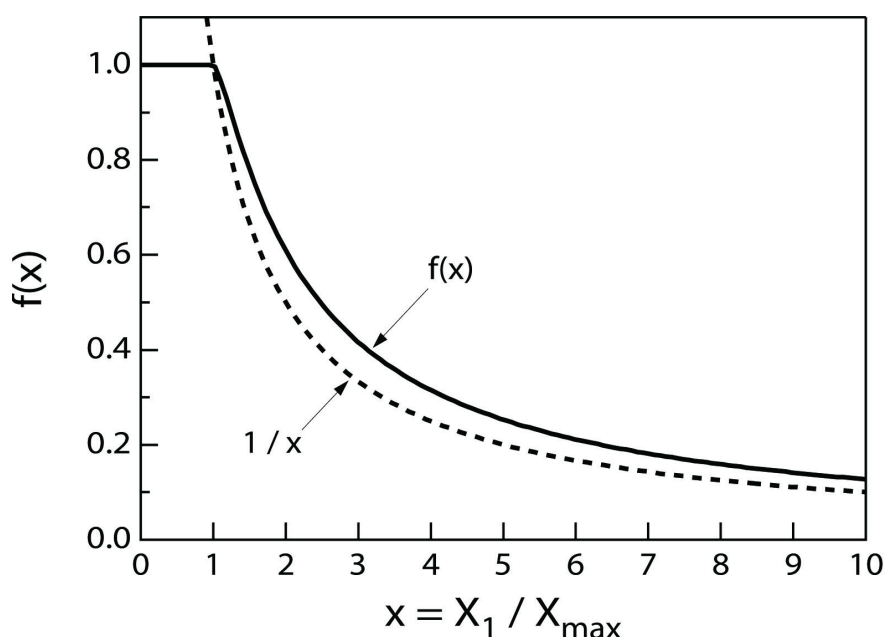


Fig 5-19: Function $f(\xi)$ given by (5.65) as a function of ξ .

The compression rate (CR) is defined as the ratio between the amplitude of the fundamental output signal of the nonlinear system and the amplitude of the fundamental output signal of an ideal linear system:

$$CR \equiv \frac{\text{fundamental amplitude}}{\text{linear system fundamental amplitude}} \cdot 100\% \quad (5.67)$$

For the characteristic in Fig. 5-18, the compression rate is:

$$CR = \frac{a_1}{k_1 X_1} = \frac{k_1 X_1 \cdot f(\xi)}{k_1 X_1} = f(\xi) \quad (5.68)$$

It is generally expressed in dB. It is common to refer to the amplitude corresponding to a compression rate of -1 dB.

APPROXIMATION OF HARMONIC AMPLITUDE

We can estimate the amplitude of the harmonics at the output of a nonlinear memoryless system excited by a sinusoidal signal, by expanding the nonlinear characteristic with a Taylor series:

$$y = k_0 + k_1x + k_2x^2 + k_3x^3 + \dots \quad (5.69)$$

By introducing the sinusoidal signal (5.61) and developing, we obtain:

$$y = Y_0 + Y_1 \cos(\phi) + Y_2 \cos(2\phi) + Y_3 \cos(3\phi) + \dots \quad (5.70)$$

$$\begin{aligned} Y_0 &= k_0 + \frac{1}{2}k_2X^2 + \dots & Y_1 &= k_1X + \frac{3}{4}k_3X^3 + \dots \\ \text{with:} & & & \\ Y_2 &= \frac{1}{2}k_2X^2 + \dots & Y_3 &= \frac{1}{4}k_3X^3 + \dots \end{aligned} \quad (5.71)$$

For a perfectly linear system without offset ($k_i = 0 \quad \forall i \neq 1$) the output signal is sinusoidal and its amplitude Y_1 simply reduces to k_1X . The factor k_1 then corresponds to the gain of the linear system. For a nonlinear system, the output signal is no longer sinusoidal (but is still periodic with the same period as the input signal). The amplitude of the fundamental Y_1 is modified by the cubic term $(3/4)k_3X^3$ and can thus be either larger or smaller than that obtained when considering the system to be linear. For a nonlinear system dominated by the cubic term, we then refer to an expansion characteristic when $k_3 > 0$ and a compression characteristic when $k_3 < 0$.

COMPRESSION POINT AT -1 dB

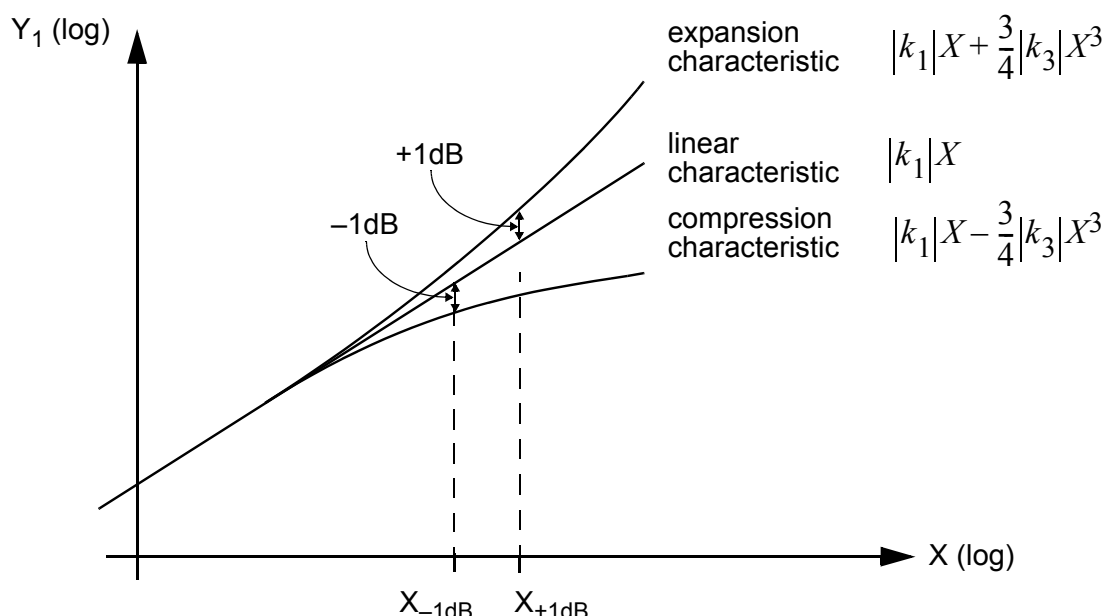


Fig 5-20: Expansion or compression point at ± 1 dB.

The compression point at -1 dB (or expansion point at +1 dB) corresponds to the amplitude X_{-1dB} (or X_{+1dB}) for which the fundamental is 1 dB from the value obtained when considering the system as perfectly linear (cf Fig. 5-20). These amplitudes are given respectively by:

$$X_{-1dB} = \sqrt{\frac{4}{3}(1 - 10^{-1/20})} \left| \frac{k_1}{k_3} \right| \quad X_{+1dB} = \sqrt{\frac{4}{3}(10^{1/20} - 1)} \left| \frac{k_1}{k_3} \right| \quad (5.72)$$

or expressed in dB:

$$\begin{aligned} 20\log(X_{-1dB}) &= 10\log\left(\left|\frac{k_1}{k_3}\right|\right) - 8.386 \quad [dB] \\ 20\log(X_{+1dB}) &= 10\log\left(\left|\frac{k_1}{k_3}\right|\right) - 7.886 \quad [dB] \end{aligned} \quad (5.73)$$

HARMONIC DISTORTION INTERCEPT (HDI)

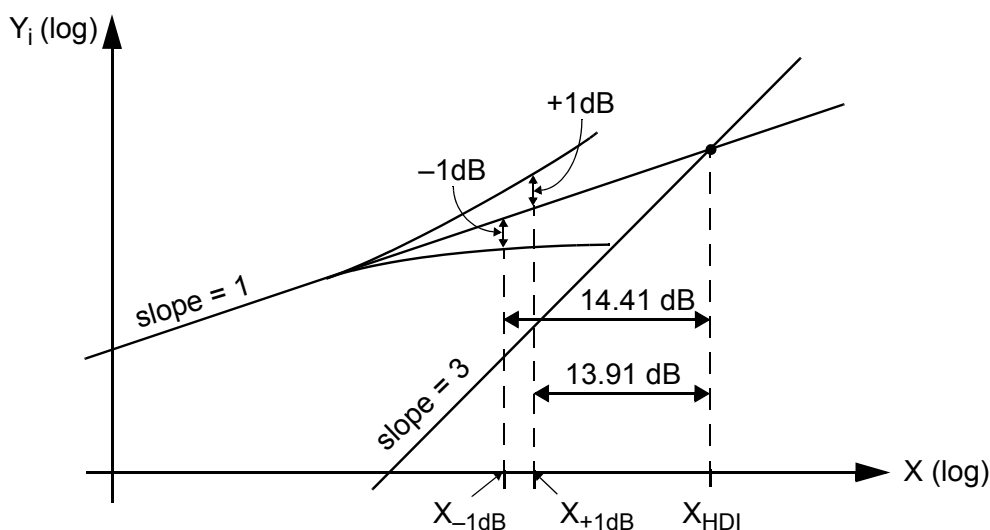


Fig 5-21: *Harmonic distortion intercept.*

Fig. 5-21 shows the harmonic components Y_1 and Y_3 as a function of the input signal amplitude X in a log-log graph for the case in which the nonlinearity is dominated by the cubic term. The compression of the fundamental is thus mainly due to the cubic term $\frac{3}{4}|k_3|X^3$ and the compression of the third harmonic is negligible. Notice that the amplitude of the 3rd harmonic increases three times as quickly as that of the fundamental (in a log-log diagram). The amplitude X_{HDI} (HDI=Harmonic Distortion Intercept) corresponding to the intersection point between the fundamental when considering the system to be linear, and the 3rd harmonic (without compression) is given by:

$$k_1 X_{HDI} = \frac{1}{4} k_3 X_{HDI}^3 \rightarrow X_{HDI} = 2 \cdot \sqrt{|k_1/k_3|} \quad (5.74)$$

or as: $20\log(X_{HDI}) = 10\log(|k_1/k_3|) + 6 \text{ dB} \quad (5.75)$

The compression point at -1 dB is located 14.41 dB below the intersection point:

$$\Delta X \equiv 20\log(X_{HDI}/X_{-1dB}) = -10\log\left(\frac{1}{3}(1 - 10^{-1/20})\right) = 14.41 \text{ dB} \quad (5.76)$$

EXAMPLE 1: EXPONENTIAL CHARACTERISTIC (1/2)

For the exponential characteristic of a bipolar transistor, the collector current for a base-emitter control voltage $v_{be}(t) \equiv V_{BEq} + \Delta V_{BE} \cdot \cos(\omega_0 t)$ is given by:

$$i_c(t) = I_s \cdot \exp\left[\frac{V_{BEq} + \Delta V_{BE} \cdot \cos(\omega_0 t)}{U_T}\right] = I_q \cdot \exp[X \cdot \cos(\omega_0 t)] \quad (5.77)$$

where $I_q \equiv I_s \cdot \exp[V_{BEq}/U_T]$ is the bias current defined for $\Delta V_{BE} = 0$ and $X \equiv \Delta V_{BE}/U_T$. We can thus normalize the current $i_c(t)$ to I_q :

$$y(t) \equiv \frac{I_c(t)}{I_q} = \exp[X \cdot \cos(\omega_0 t)] \quad (5.78)$$

It can be shown that the Fourier series development of the function (5.78) is given by:

$$y(t) = \exp[X \cdot \cos(\omega_0 t)] = I_0(X) + 2 \cdot \sum_{n=1}^{+\infty} I_n(X) \cdot \cos(n\omega_0 t) \quad (5.79)$$

where $I_0(X)$ is the modified Bessel function of order 0 and $I_n(X)$ is the modified Bessel function of order n . For an exponential function, it is thus possible to calculate the harmonics Y_n exactly. But unfortunately, this is not always possible. So we can compare the harmonics calculated exactly by (5.79) to those obtained by expanding the normalized exponential characteristic $y(t)$ in a Taylor series:

$$y(t) \equiv \frac{I_c(t)}{I_q} = e^x \cong 1 + x + \frac{x^2}{2} + \frac{x^3}{6} \quad (5.80)$$

where $x(t) \equiv X \cdot \cos(\omega_0 t)$.

EXAMPLE 1: EXPONENTIAL CHARACTERISTIC (2/2)

From (5.80) we get the coefficients $k_1 = 1$ and $k_3 = 1/6$ and the expansion point at +1 dB:

$$20\log(X_{+1dB}) = -0.1 \text{ dB} \quad (5.81)$$

and the intersection point of the harmonics:

$$20\log(X_{HDI}) = 13.8 \text{ dB} \quad (5.82)$$

The expansion point at +1 dB and the intersection point of the harmonic components are represented in Fig. 5-22. Notice that the approximation of the compression by using the expansion (5.80) is very close to the exact characteristic. In addition, we see that the third harmonic is subject to expansion and moves away from the right.

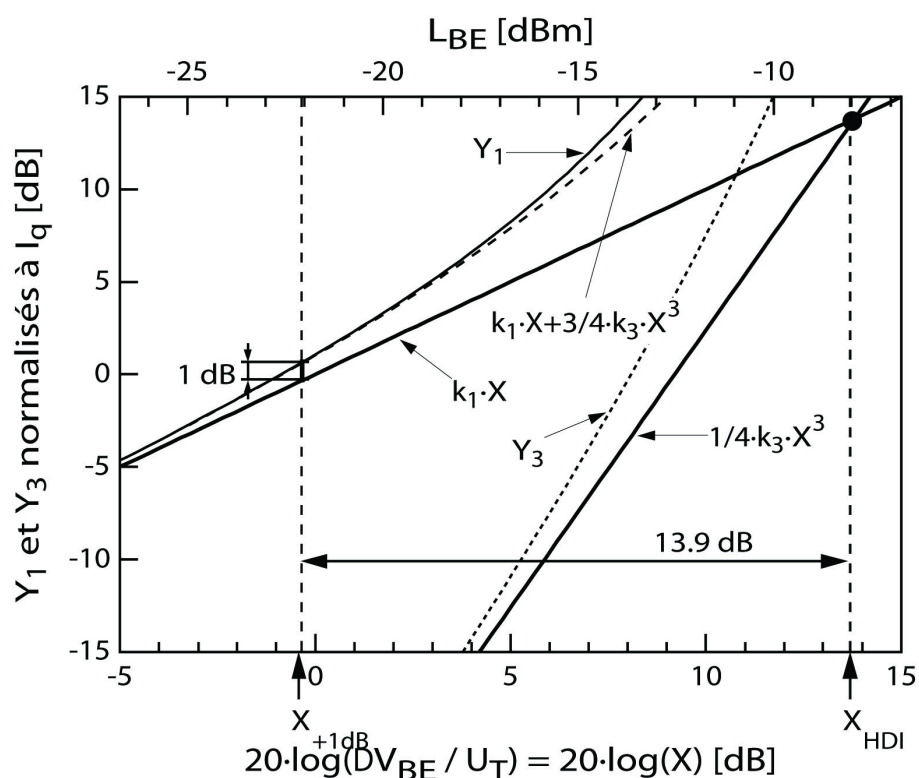


Fig 5-22: Expansion point at 1 dB and intersection of harmonic components for an exponential.

EXAMPLE 2: THE DIFFERENTIAL PAIR (1/2)

The differential current of a bipolar differential pair is simply given by:

$$y \equiv \frac{\Delta i}{I_q} = \tanh(x) \quad \text{with} \quad x \equiv \frac{V_{in}}{2U_T} \quad (5.83)$$

where V_{in} is the differential input voltage and I_q half of the tail current (or the bias current of each of the transistors when $V_{in} = 0$). Unlike the case of the exponential function, there is no analytic expression for the Fourier series expansion of (5.83) with $x(t) = X \cdot \cos(\omega_0 t)$. This characteristic can nevertheless be approximated with a Taylor series expansion:

$$y \equiv \tanh(x) \cong x - \frac{x^3}{3} \quad (5.84)$$

Notice that there is no even term. From (5.84) we get $k_1 = 1$ and $k_3 = -1/3$. This approximation is represented in Fig. 5-23 with the function $\tanh(x)$ as well as another approximation:

$$y \equiv \tanh(x) \cong x - 0.254 \cdot x^3 \quad (5.85)$$

We remark that the approximation (5.85) is better (error less than 2% for $|x| \leq 1$) than that given by the Taylor expansion (5.84) (error less than 2% for $|x| \leq 0.59$).

The amplitude of the fundamental calculated using approximations (5.84) and (5.85), along with the amplitude of the linear system, are shown in Fig. 5-24. The compression point calculated from (5.73) is given by -3.6 dB for (5.84) and -2.44 dB for (5.85). The first and third harmonics intersect at the point 10.8 dB using the approximation (5.84).

EXAMPLE 2: THE DIFFERENTIAL PAIR (2/2)

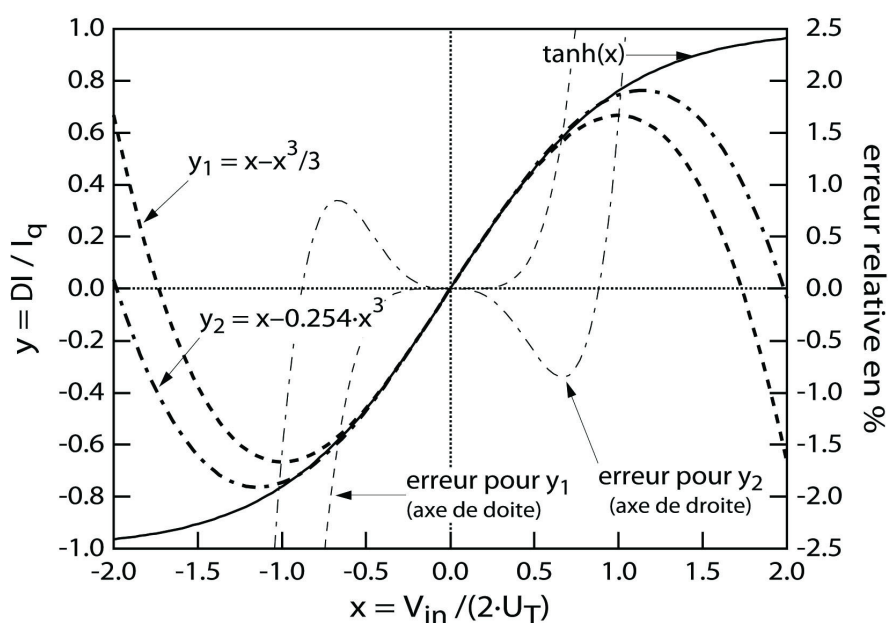


Fig 5-23: $\tanh(x)$ and 3rd degree approximation.

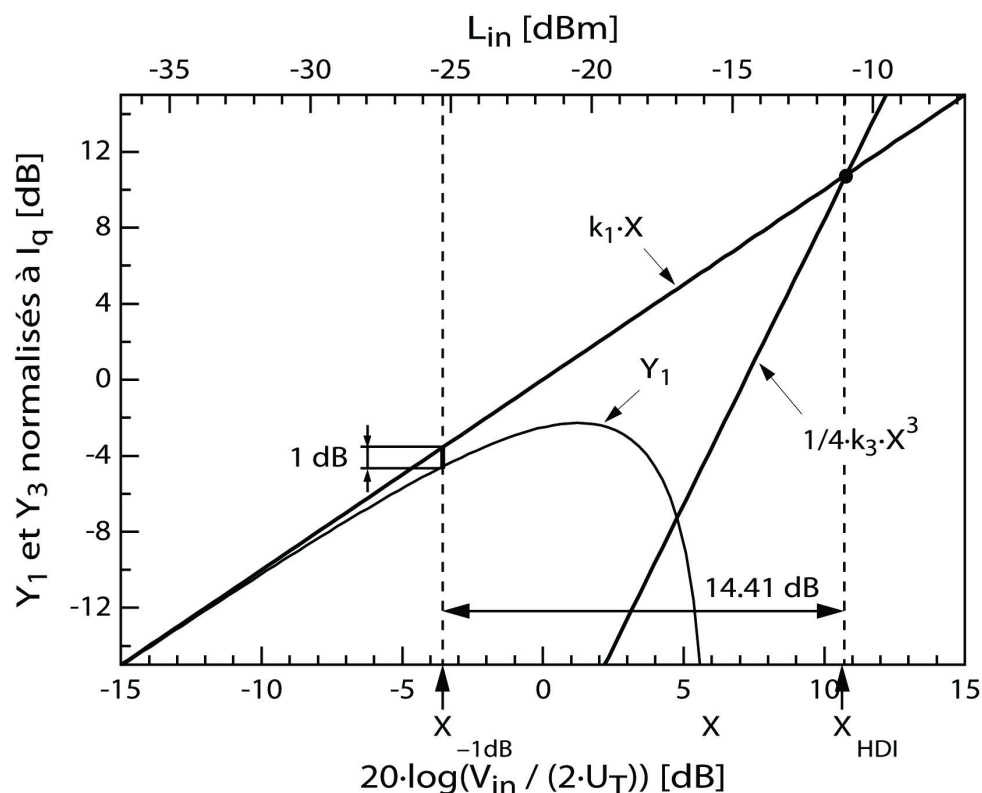


Fig 5-24: Compression point at -1 dB and intersection point of harmonics 1 and 3 for a $\tanh$ characteristic.

HARMONIC DISTORTION RATE

One measure of the distortion of a signal is given by the ratio between the harmonic amplitude n (r_n) and the amplitude of the fundamental r_1 , measured as a percentage:

$$\text{Distorsion rate for the harmonique } n \equiv \frac{r_n}{r_1} \cdot 100\% \quad (5.86)$$

The total harmonic distortion rate (THD) is the ratio between the RMS value of the harmonics $n \neq 1$ and the RMS value of the distorted signal at the output for a sinusoidal input signal:

$$THD \equiv \frac{\sqrt{\sum_{n=2}^{+\infty} a_n^2 + b_n^2}}{\sqrt{\sum_{n=1}^{+\infty} a_n^2 + b_n^2}} = \frac{\sqrt{\sum_{n=2}^{+\infty} r_n^2}}{\sqrt{\sum_{n=1}^{+\infty} r_n^2}} \cong \frac{\sqrt{r_2^2 + r_3^2}}{\sqrt{r_1^2 + r_2^2 + r_3^2}} \quad (5.87)$$

The THD can often be approximated by considering only the first harmonics, because the amplitude of the harmonics generally decreases very quickly.

INTERMODULATION

Sinusoidal “duo tone” input signal

Consider a nonlinear system with a static polynomial characteristic:

$$y = k_0 + k_1x + k_2x^2 + k_3x^3 + k_4x^4 + \dots \quad (5.88)$$

Let us consider the case of a signal composed of two sinusoidal waves:

$$x(t) = X_1 \cos(\omega_1 t) + X_2 \cos(\omega_2 t) = X_1 \cos(\phi_1) + X_2 \cos(\phi_2) \quad (5.89)$$

where $\phi_i \equiv \omega_i t$ ($i = 1, 2$). The output signal is obtained by substituting (5.89) into (5.88). Let's look at the result term by term. The linear term is given by:

$$k_1x = k_1X_1 \cos(\phi_1) + k_1X_2 \cos(\phi_2) \quad (5.90)$$

The amplitudes of the sinusoidal waves are simply multiplied by the gain k_1 . The quadratic term is given by:

$$\begin{aligned} k_2x^2 &= k_2[X_1 \cos(\phi_1) + X_2 \cos(\phi_2)]^2 \\ &= k_2[X_1^2 \cos^2(\phi_1) + 2X_1X_2 \cos(\phi_1)\cos(\phi_2) + X_2^2 \cos^2(\phi_2)] \\ &= \left. \frac{k_2X_1^2}{2} + \frac{k_2X_2^2}{2} \right\} = \text{constant term} \end{aligned} \quad (5.91)$$

$$+ \frac{k_2X_1^2}{2} \cos(2\phi_1) + \frac{k_2X_2^2}{2} \cos(2\phi_2) \left\} = \text{second harmonics}$$

$$+ k_2X_1X_2 [\cos(\phi_1 + \phi_2) + \cos(\phi_1 - \phi_2)] \left\} = \begin{array}{l} \text{products of second order} \\ \text{intermodulation} \end{array}$$

In addition to the constant term (at zero frequency) and second harmonics, the signal contains two equal amplitude components for which the frequencies are the sum and the difference of the input signal frequencies. These components are called second order intermodulation products (IM2). Thus, the nonlinearity corresponding to the quadratic term generates second order intermodulation products whose amplitudes depend linearly on X_1 and X_2 .

3rd ORDER INTERMODULATION PRODUCTS

The cubic term is given by:

$$\begin{aligned}
 k_3 x^3 &= k_3 [X_1 \cos(\phi_1) + X_2 \cos(\phi_2)]^3 \\
 &= k_3 [X_1^3 \cos^3(\phi_1) + 3X_1^2 X_2 \cos^2(\phi_1) \cos(\phi_2) \\
 &\quad + 3X_1 X_2^2 \cos(\phi_1) \cos^2(\phi_2) + X_2^3 \cos^3(\phi_2)]
 \end{aligned} \tag{5.92}$$

After expansion, we get:

$$\begin{aligned}
 k_3 x^3 &= \frac{3k_3}{2} \left(X_1 X_2^2 + \frac{X_1^3}{2} \right) \cos(\phi_1) + \frac{3k_3}{2} \left(X_1^2 X_2 + \frac{X_2^3}{2} \right) \cos(\phi_2) \Big\} = \text{fundamentals} \\
 &\quad + \frac{k_3 X_1^3}{4} \cos(3\phi_1) + \frac{k_3 X_2^3}{4} \cos(3\phi_2) \Big\} = \text{3rd harmonics} \\
 &\quad + \frac{3k_3 X_1^2 X_2}{4} [\cos(2\phi_1 + \phi_2) + \cos(2\phi_1 - \phi_2)] \\
 &\quad + \frac{3k_3 X_1 X_2^2}{4} [\cos(2\phi_2 + \phi_1) + \cos(2\phi_2 - \phi_1)] \Big\} = \text{3rd order IM products}
 \end{aligned} \tag{5.93}$$

Eqn. 5.93 shows that the 3rd order nonlinearity produces fundamental components along with 3rd order harmonics. In addition, we find 3rd order IM products at the frequencies $2f_1 + f_2$, $2f_2 + f_1$, $2f_1 - f_2$ and $2f_2 - f_1$. When the frequencies f_1 and f_2 are close, the components at frequencies $2f_1 - f_2$ and $2f_2 - f_1$ become particularly bothersome, because they fall into the useful transmission band and are thus difficult to eliminate by selective filtering.

The same approach can be used for nonlinearities of order higher than 3. In general, an n th order nonlinearity produces harmonics that are multiples of the frequencies f_1 and f_2 up to n . If n is even (odd), they are all even (odd) multiples of f_1 and f_2 . If n is even, there is no fundamental, but there is a constant term.

INTERMODULATION PRODUCTS

In general, the intermodulation products generated by a nonlinearity of order n are given by:

$$\pm k \cdot f_1 \mp l \cdot f_2 \quad \text{where:} \quad k + l = m \quad \text{with:} \quad m = 2, 3, \dots, n \quad (5.94)$$

where the upper frequencies are obtained by taking the two plus signs and the lower frequencies are obtained by taking opposite signs such that the resulting frequencies remain positive. The two minus signs are thus never used. Taking $\Delta f \equiv |f_2 - f_1|$, certain odd-order IM products are given by:

n	$l \cdot f_2 - k \cdot f_1$	$k \cdot f_1 - l \cdot f_2$	$l \cdot f_2 + k \cdot f_1$	$k \cdot f_1 + l \cdot f_2$
3	$2f_2 - f_1$ $= f_2 + \Delta f$	$2f_1 - f_2$ $= f_1 - \Delta f$	$2f_2 + f_1$ $= 3f_2 - \Delta f$	$2f_1 + f_2$ $= 3f_1 + \Delta f$
5	$3f_2 - 2f_1$ $= f_2 + 2\Delta f$	$3f_1 - 2f_2$ $= f_1 - 2\Delta f$	$3f_2 + 2f_1$ $= 5f_2 - 2\Delta f$	$3f_1 + 2f_2$ $= 5f_1 + 2\Delta f$
7	$4f_2 - 3f_1$ $= f_2 + 3\Delta f$	$4f_1 - 3f_2$ $= f_1 - 3\Delta f$	$4f_2 + 3f_1$ $= 7f_2 - 3\Delta f$	$4f_1 + 3f_2$ $= 7f_1 + 3\Delta f$

The IM product components closest to the fundamental frequencies f_1 and f_2 are presented in Fig. 5-25.

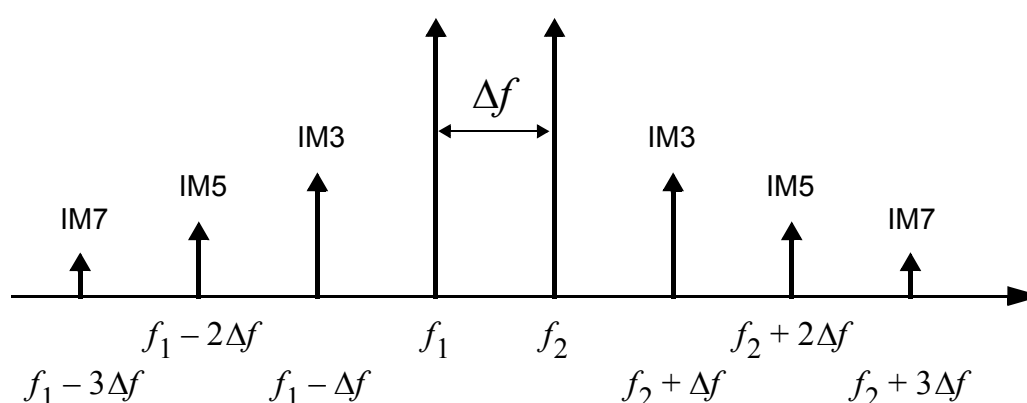


Fig 5-25: Odd IM products around f_1 and f_2

DESENSITIZATION AND BLOCKING (1/2)

A desired input signal $X_1 \cos(\phi_1)$ added to an interfering signal $X_2 \cos(\phi_2)$ is applied to a nonlinear system with the characteristic shown in Fig. 5-26 and described by:

$$y = \begin{cases} k_1 \cdot x + k_3 \cdot x^3 = 12x - x^3 & |x| \leq 2 \\ \pm 16 & |x| > 2 \end{cases} \quad (5.95)$$

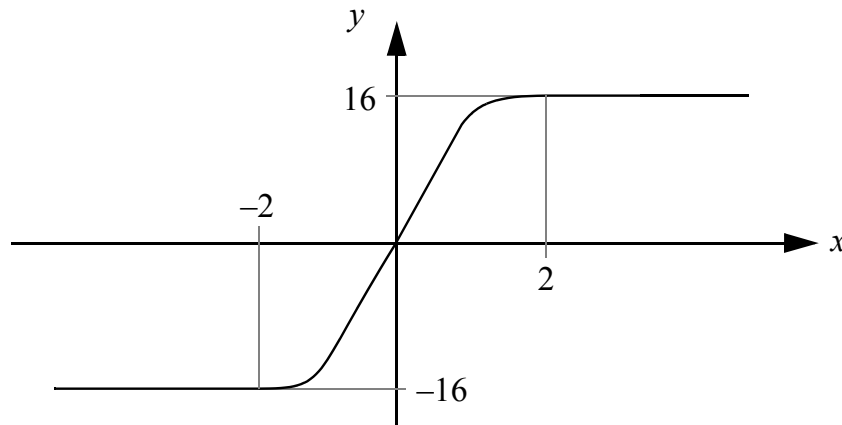


Fig 5-26: *Odd nonlinear characteristic.*

The output signal y_1 at the frequency f_1 for an input signal amplitude less than 2 (no saturation), is given by:

$$y_1 = 12X_1 \left(1 - \frac{X_1^2}{16} - \frac{X_2^2}{8} \right) \cos(\phi_1) \cong 12X_1 \left(1 - \frac{X_1^2}{16} \right) \cos(\phi_1) \quad \text{for } X_2 \ll X_1 \quad (5.96)$$

The compression rate for the case in which the amplitude of the interfering signal is much weaker than the amplitude of the desired signal is given by:

$$CR_{dB} = 20 \log \left(1 - \frac{X_1^2}{16} \right) \quad (5.97)$$

The amplitude corresponding to a compression rate of -1 dB is thus 1.32 V, which is in agreement with the hypothesis that $X_1 < 2V$. For small amplitudes X_1 , the system functions as a linear amplifier.

DESENSITIZATION AND BLOCKING (2/2)

Let us consider the case in which the amplitude of the useful signal is small (for example $X_1 < 0.4 V$) and the interfering signal becomes dominant, so:

$$y_1 = 12X_1 \left(1 - \frac{X_2^2}{8}\right) \cos(\phi_1) \quad X_1, X_2 \leq 2 \quad (5.98)$$

We notice that the effect of the interfering signal is to reduce the amplitude of the useful signal at the output. This phenomenon, called desensitization, is similar to compression, but is caused by a strong interfering signal at a different frequency. We define a desensitization ratio by:

$$DS \equiv \frac{|desired\ output\ signal\ in\ the\ presence\ of\ interfering\ signal|}{|desired\ output\ signal\ without\ interfering\ signal|} \quad (5.99)$$

In our case we have:

$$DS_{dB} = 20 \log \left(1 - \frac{X_2^2}{8}\right) \quad X_2 \leq 2 \quad (5.100)$$

For $X_1 < 0.4 V$ and $X_2 = 1.6 V$, we have $DS_{dB} = -3.35 dB$. We thus speak of a desensitization of about 3 dB that represents a reduction of the useful signal amplitude by a factor $\sqrt{2}$ and thus a reduction of the power by a factor of two.

In an extreme case, a large enough interfering signal can even cancel out the amplitude of the desired signal. This phenomenon is called blocking. In our case, the amplitude X_2 provoking the cancellation of the useful signal is given by $1 - X_2^2/8 = 0$, corresponding to $X_2 = 2.83 V$. This amplitude is larger than 2 V and therefore does not satisfy the hypotheses. Nevertheless, we can imagine a system with a useful signal that is cancelled out entirely by an interfering signal at another frequency.

THIRD-ORDER INTERCEPT POINT (IP3) (1/2)

Consider the case in which the amplitude of the interfering signal is equal to the amplitude of the desired signal ($X_1 = X_2 = V_{in}$). The average power of the input signal P_{in} and its corresponding level L_{in} are:

$$L_{in} \equiv 10\log\left(\frac{P_{in}}{1mW}\right) = 10\log\left(\frac{V_{in}^2}{2R_{in} \cdot 1mW}\right) \quad [dBm] \quad (5.101)$$

The factor 2 in (5.101) comes from the fact that V_{in} is a peak value and not an RMS value. The level of the linear term of the signal at the system output is given by:

$$L_{out1} = 10\log\left(\frac{P_{out1}}{1mW}\right) = 10\log\left(\frac{GP_{in}}{1mW}\right) = L_{in} + G_{dB} \quad (5.102)$$

where G is the gain in power $G \equiv P_{out1}/P_{in}$, which can be expressed as a function of the voltage gain k_1 and the ratio between the input resistance R_{in} and the load resistance R_L :

$$G = \frac{V_{out}^2}{2R_L} \cdot \frac{2R_{in}}{V_{in}^2} = \frac{(k_1 \cdot V_{in})^2}{V_{in}^2} \cdot \frac{R_{in}}{R_L} = k_1^2 \cdot \frac{R_{in}}{R_L} \quad (5.103)$$

The amplitude of the 3rd order IM product at $2f_1 - f_2$ is equal to $\frac{3}{4}k_3 V_{in}^3$. The average power corresponding to the load terminal R_L is thus:

$$P_{out3} = \frac{(\frac{3}{4}k_3 V_{in}^3)^2}{2R_L} \quad (5.104)$$

corresponding to a level L_{out3} :

$$\begin{aligned} L_{out3} &= 10\log\left(\frac{(\frac{3}{4}k_3 V_{in}^3)^2}{2R_L \cdot 1mW}\right) = 10\log\left(\frac{9}{4}k_3^2 \frac{R_{in}^3}{R_L} (1mW)^2 \left(\frac{V_{in}^2}{2R_{in} \cdot 1mW}\right)^3\right) \\ &= 3L_{in} + 20\log(k_3) + 10\log\left(\frac{R_{in}^3}{R_L}\right) - 56.5 = 3L_{in} + K \end{aligned} \quad (5.105)$$

where:

$$K \equiv 20\log(k_3) + 10\log\left(\frac{R_{in}^3}{R_L}\right) - 56.5 \quad (5.106)$$

THIRD-ORDER INTERCEPT POINT (IP3) (2/2)

The levels L_{out1} corresponding to the fundamental and L_{out3} corresponding to the 3rd order IM product at $2f_1 - f_2$, are shown as a function of the input level L_{in} in Fig. 5-27. The intersection point of the lines L_{out1} and L_{out3} is called the third-order intercept point (PI3). This point can be specified either by its projection on the x-axis IIP , or by its projection on the y-axis of L_{out1} (OIP). Unfortunately, this point cannot be measured directly because it corresponds to an extrapolation. The phenomena of compression, desensitization, and higher-order IM products deteriorate the linearity of the system for high power signals.

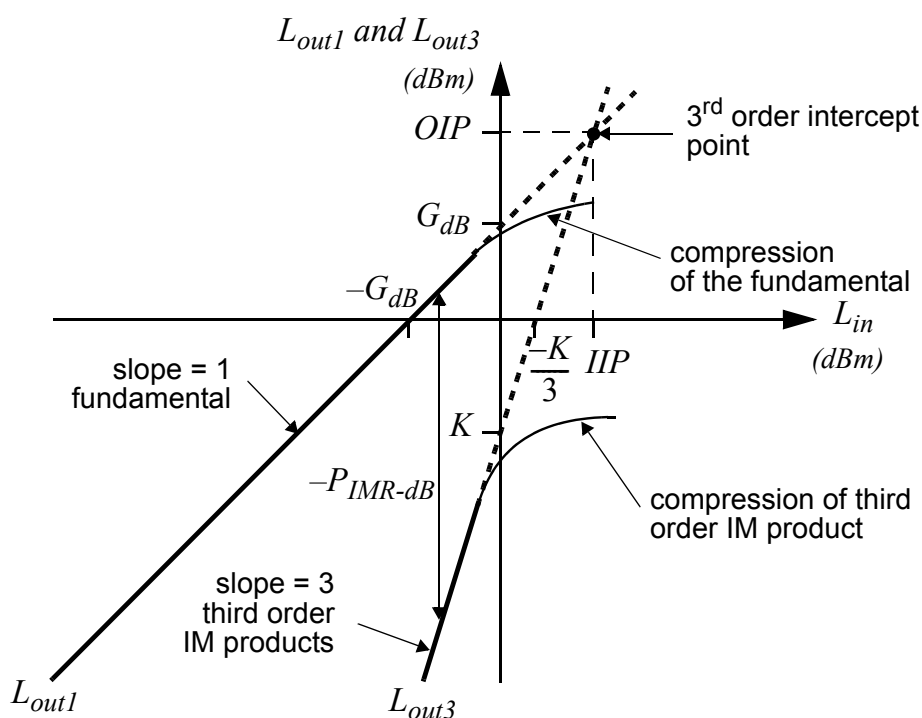


Fig 5-27: 3rd order intercept point.

The value of IIP can be expressed as a function of k_1 and k_3 by using equations (5.102), (5.103) and (5.105) in the following manner:

$$IIP = 10\log(k_1) - 10\log(k_3) - 10\log(R_{in}) + 28.24 \quad [dBm] \quad (5.107)$$

RELATIONSHIP BETWEEN THE COMPRESSION POINTS AND INTERCEPT POINTS IN “SINGLE TONE” AND “DUO TONE”

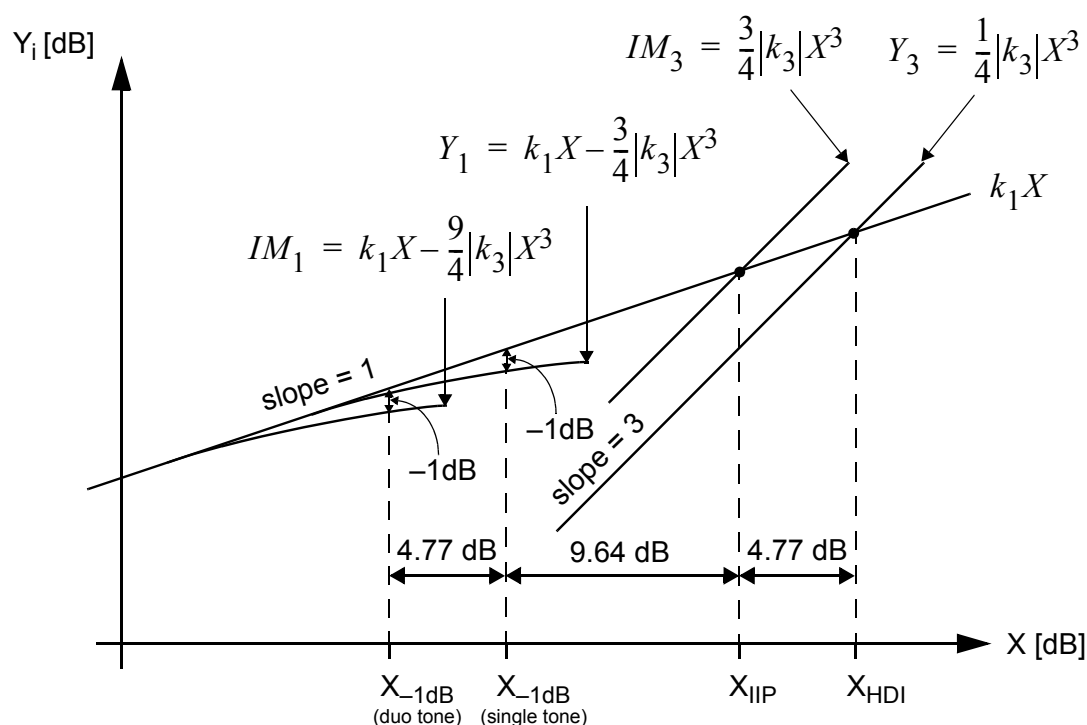


Fig 5-28: Relationship between the compression points at -1 dB and the intercept points measured in “single tone” and “duo tone.”

The relationship between the compression points at -1 dB and the intercept points measured in “single tone” and “duo tone” are illustrated in Fig. 5-28. Notice that the compression and intercept points measured in “single tone” are 4.77 dB higher than the compression and intercept points measured in “duo tone”. We can thus deduce the intercept point IIP from a “single tone” measurement, which is much easier to do than a “duo tone” measurement. This is not entirely correct in the case of a narrow-band system, because the harmonics of a sinusoidal signal generally fall outside of the useful band and are thus strongly attenuated while the 3rd order intermodulation products fall in the baseband and are not attenuated.

COMPARISON FOR A DIFFERENTIAL PAIR

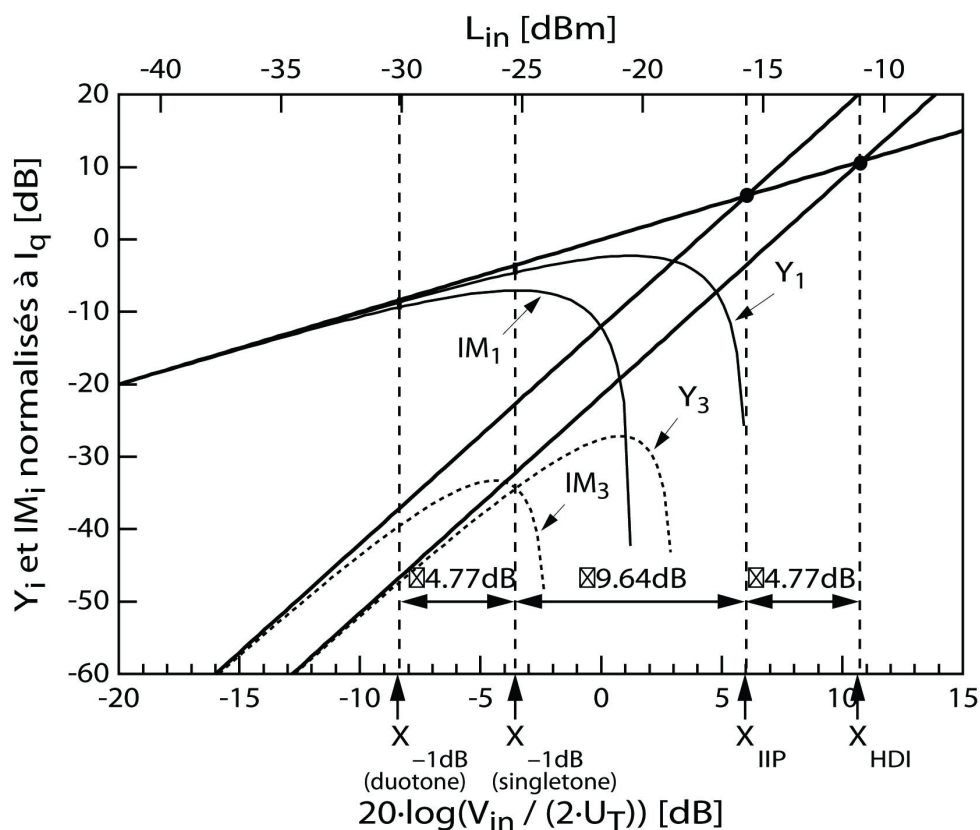


Fig 5-29: Comparison of compression and intercept points for “single tone” and “duo tone” signals for a differential pair.

A comparison between the compression and intercept points for “single tone” and “duo tone” signals in the case of a differential pair is presented in Fig. 5-29. Once again we find a difference of 4.77 dB.

INTERMODULATION DISTORTION RATIO

The Intermodulation Distortion Ratio (IMR) is the ratio between the amplitude of one of the 3rd order IM products and the amplitude of the linear term:

$$IMR \equiv \frac{\frac{3}{4}k_3 X_1^2 X_2}{k_1 X_1} = \frac{3k_3}{4k_1} X_1 X_2 \quad (5.108)$$

The ratio between the power of the 3rd order IM products at the output and the power of the linear term for the amplitudes $X_1 = X_2 = V_{in}$ is given by:

$$P_{IMR} \equiv \frac{P_{out3}}{P_{out1}} = \frac{(\frac{3}{4}k_3 V_{in}^3)^2}{2R_L} \cdot \frac{2R_L}{(k_1 V_{in})^2} = \left(\frac{3k_3}{4k_1} V_{in}^2 \right)^2 = \left(\frac{3k_3}{4k_1} 2R_{in} \right)^2 \cdot P_{in}^2 \quad (5.109)$$

By definition, this ratio is equal to one when the input signal level is equal to IIP corresponding to a power P_{IIP} , from which we get:

$$P_{IMR} = \left(\frac{P_{in}}{P_{IIP}} \right)^2 \quad (5.110)$$

In decibels we have:

$$P_{IMR-dB} = 2 \cdot (L_{in} - IIP) \quad (5.111)$$

We can thus express IIP as a function of P_{IMR-dB} and L_{in} :

$$IIP = L_{in} - \frac{P_{IMR-dB}}{2} \quad (5.112)$$

The projection of the intersection point at the output is then simply given by:

$$OIP = L_{in} + G_{dB} - \frac{P_{IMR-dB}}{2} \quad (5.113)$$

MINIMUM DETECTABLE SIGNAL AND NOISE FLOOR

We have seen that the average noise factor $\overline{F}$ represents the quotient of the signal-to-noise ratios at the input SNR_i and the output SNR_o of the two-port network. The minimum detectable signal (*MDS*) corresponds to the value of the signal which must be applied at the input in order to have a given signal-to-noise ratio at the output. From (5.26), we find:

$$MDS \equiv L_{min} = \overline{NF} + 10\log(kTB) + (SNR_o)_{dB} + 30 \quad [dBm] \quad (5.114)$$

For $T = 290\text{ K}$ and $SNR_o = 1$ we have:

$$MDS = \overline{NF} + 10\log(B_{kHz}) - 144 \quad [dBm] \quad (5.115)$$

where B_{kHz} is the bandwidth expressed in kHz. The output signal level corresponding to an *MDS* level at the input is called the Noise Floor. When the input signal exceeds the *MDS*, the output signal level increases linearly beyond the noise floor. The level of the IM product increases three times faster, but is initially lost under the noise floor.

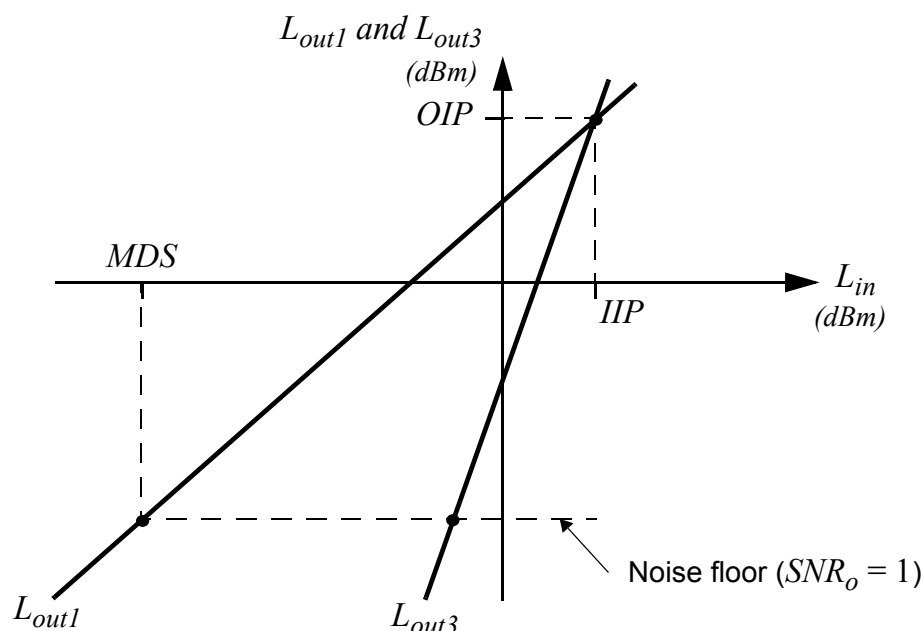


Fig 5-30: Relationship between the *MDS* and the noise floor.

DYNAMIC RANGE (1/2)

The dynamic range of a system is defined by the ratio between the maximum and minimum signals that the system can process. The minimum signal is limited by the noise while the maximum signal is limited by distortion. For a narrow-band system such as a receiver, the maximum signal is basically limited by the 3rd order IM products. The definition of the minimum signal level generally corresponds to the *MDS*. The maximum signal can be defined in multiple ways. One possible definition corresponds to the input signal level L_{IM} for which the level of the IM product is equal to the *MDS* (cf Fig. 5-31).

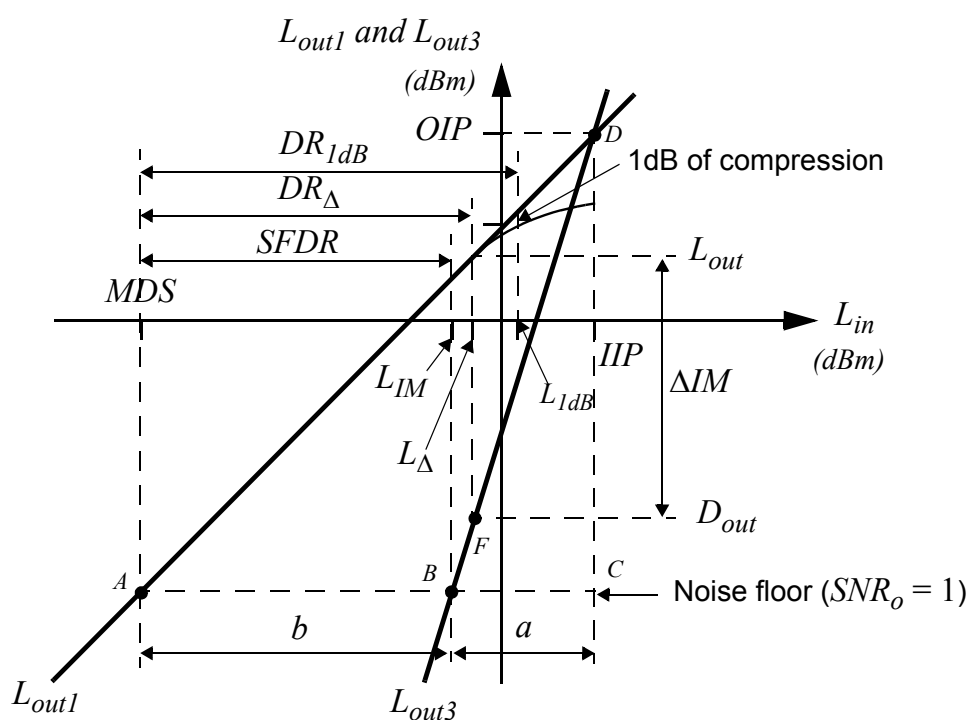


Fig 5-31: Relationship between the MDS and the dynamic range. The Spurious Free Dynamic Range (SFDR) of the system is then defined as the difference between L_{IM} and MDS:

$$SFDR \equiv L_{IM} - MDS \quad [dB] \quad (5.116)$$

DYNAMIC RANGE (2/2)

The maximum signal can be alternatively defined as the signal that produces a 1 dB compression of the fundamental. If L_{1dB} is the corresponding input level (cf Fig. 5-31), the dynamic range is equal to the difference between L_{1dB} and MDS :

$$DR_{1dB} \equiv L_{1dB} - MDS \quad [dB] \quad (5.117)$$

The dynamic range can also be defined as a function of the signal level L_{Δ} for which the third-order IM product L_{out3} is at a value $\Delta IM \equiv L_{out1} - L_{out3}$ (typically 60 dB) below the level L_{out1} corresponding to L_{Δ} :

$$DR_{\Delta} \equiv L_{\Delta} - MDS \quad (5.118)$$

One can define several relationships from the geometry shown in Fig. 5-31. Let b be the distance between the lines L_{out1} and L_{out3} and a be the distance between the line L_{out3} and the point IIP . Since the lines L_{out1} and L_{out3} have, respectively, slopes of 1 and 3, then $b = 2a$. Drawing a horizontal line that passes through the point F , we find that $b = \Delta IM$ and therefore $a = (\Delta IM)/2 = IIP - L_{\Delta}$, from which we find the relationship for IIP :

$$IIP = \Delta IM/2 + L_{\Delta} \quad (5.119)$$

In addition, we see that $a + b = 3a = (3b)/2$, which lets us write: $IIP - MDS = (3SFDR)/2$ and so:

$$SFDR = \frac{2}{3}(IIP - MDS) \quad (5.120)$$

In a similar way, we see that $IIP - L_{IM} = SFDR/2$, from which we find the maximum input signal level for which the level of the IM product is equal to the MDS :

$$L_{IM} = (2IIP + MDS)/3 \quad (5.121)$$

An input signal level higher than that given by (5.121) gives an IM product that increases three times faster than the desired signal, and causes an unacceptable distortion.

2nd ORDER INTERCEPT POINT (1/2)

The second-order IM products are given by $\pm f_2 \mp f_1$. For the case in which $f_2 \gg f_1$, the frequency $\Delta f \equiv f_2 - f_1 \approx f_2$ can be near f_2 and thus interfere with the system. In the same way as for the 3rd order IM products, we define an intercept point of the second-order IM products (cf Fig. 5-32).

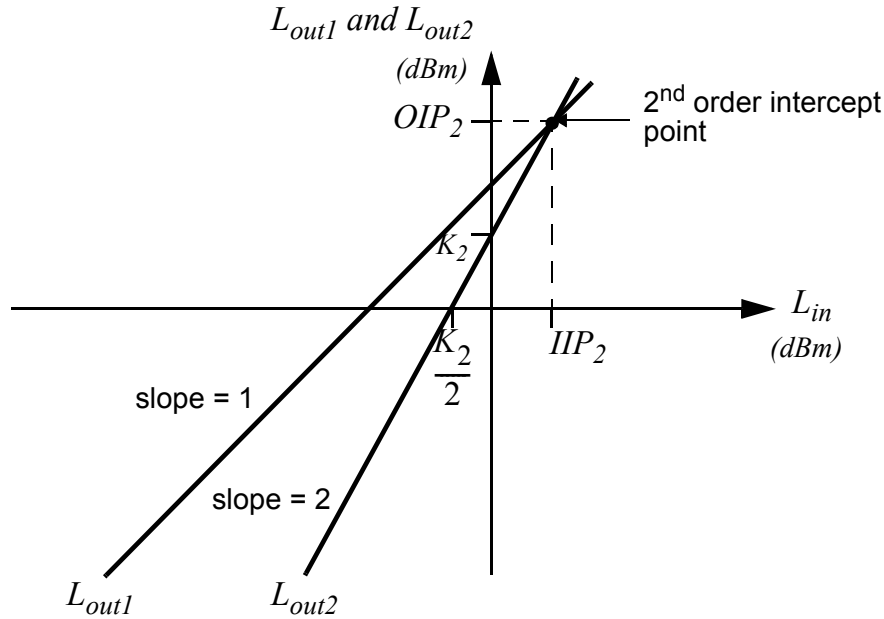


Fig 5-32: 2nd order intercept point.

The amplitudes of the second order IM products are both equal to $k_2 V_{in}^2$. The average power corresponding to the terminals of the load R_L is:

$$P_{out2} = \frac{(k_2 V_{in}^2)^2}{2R_L} \quad (5.122)$$

corresponding to a level L_{out2} :

$$\begin{aligned} L_{out2} &= 10\log\left(\frac{(k_2 V_{in}^2)^2}{2R_L \cdot 1mW}\right) = 10\log\left(\left(\frac{V_{in}^2}{2R_{in} \cdot 1mW}\right)^2 \frac{2R_{in}^2 k_2^2 1mW}{R_L}\right) \\ &= 2L_{in} + 20\log(k_2) + 10\log\left(\frac{R_{in}^2}{R_L}\right) - 27 = 2L_{in} + K_2 \end{aligned} \quad (5.123)$$

where: $K_2 \equiv 20\log(k_2) + 10\log(R_{in}^2/R_L) - 27 \quad [dB] \quad (5.124)$

2nd ORDER INTERCEPT POINT (2/2)

The projection of the second-order intercept point is given by:

$$IIP_2 = 10\log\left(\frac{k_1^2}{k_2^2 R_{in}}\right) + 27 = 10\log\left(\frac{501k_1^2}{k_2^2 R_{in}}\right) \quad (5.125)$$

The difference between the projections of the 3rd and 2nd order intercept points is given by:

$$IIP_2 - IIP = 10\log\left(\frac{501k_1^2}{k_2^2 R_{in}}\right) - 10\log\left(\frac{66.67k_1}{k_3 R_{in}}\right) = 10\log\left(\frac{k_1 k_3}{1.33k_2^2}\right) \quad (5.126)$$

EXAMPLE (1/2)

Consider an electronic system with a bandwidth of 10kHz, and input and load resistances of 50Ω . The input is matched to two signal generators through the help of an adder for which the attenuation is 6dB. The frequency of the two sine waves applied to the input are $f_1 = 3800$ kHz and $f_2 = 3802$ kHz. When the adder is calibrated, the level of each of the signals before the adder is equal to 4 dBm. For this same input signal, the spectrum analyzer measures a level of each of these two signals at the output equal to 5 dBm and indicates that the 3rd order IM components are located 16 dB lower. We assume that the other IM components are negligible and that there is no compression. The measured noise level when there is no input signal is equal to -36 dBm. We are asked to calculate:

- 1) The corresponding input level L_{in} ;
- 2) The gain in power G_{dB} ;
- 3) The 3rd order intercept point IIP ;
- 4) The minimum detectable signal MDS ;
- 5) The dynamic range $SFDR$ of the system;
- 6) The noise figure NF ;
- 7) The maximum RMS voltage at the system input before the 3rd order IM products become perceptible.

EXAMPLE (2/2)

- 1) The input level is simply given by the difference between the level measured before the adder (+4 dBm) minus the attenuation of the adder (6 dB), so:

$$L_{in} = 4 - 6 = -2 \text{ dBm} \quad (5.127)$$

- 2) The gain in power is given by:

$$G = L_{out} - L_{\Delta} = 5 - (-2) = 7 \text{ dB} \quad (5.128)$$

- 3) From (5.119), with $L_{\Delta} = L_{in}$ and $\Delta IM = 16 \text{ dB}$, we get:

$$IIP = \frac{\Delta IM}{2} + L_{\Delta} = \frac{16}{2} + (-2) = 6 \text{ dBm} \quad (5.129)$$

- 4) In the graph in Fig. 5-31, the distance CD is equal to $36 + 7 + 6 = 49 \text{ dB}$, which also corresponds to the distance AC . So:

$$MDS = IIP - 49 = 6 - 49 = -43 \text{ dBm} \quad (5.130)$$

which corresponds to an RMS voltage at the input of 1.58 mV.

- 5) From (5.120) we get:

$$SFDR = \frac{2}{3}(IIP - MDS) = \frac{2}{3}(6 - (-43)) = 32.67 \text{ dB} \quad (5.131)$$

- 6) From (5.115) we get:

$$\overline{NF} = MDS - 10 \log(B_{kHz}) + 144 = -43 - 10 + 144 = 91 \text{ dB} \quad (5.132)$$

- 7) From (5.121) we get:

$$L_{IM} = \frac{2IIP + MDS}{3} = \frac{2 \cdot 6 - 43}{3} = -10.33 \text{ dBm} \quad (5.133)$$

which corresponds to an RMS voltage at the input of 68.1 mV.

IM DISTORSION AT THE OUTPUT

The 3rd order intercept point (PI3) in relation to the input is given by (5.119), and repeated here:

$$IIP = \Delta IM/2 + L_{\Delta} \quad [dBm] \quad (5.134)$$

The PI3 in relation to the output is obtained simply from (5.134) by adding the gain (in dB):

$$OIP = IIP + G = \Delta IM/2 + L_{\Delta} + G \quad [dBm] \quad (5.135)$$

Let L_{out} be the output level corresponding to L_{Δ} and D_{out} the distortion level at the output corresponding to an input level L_{Δ} . ΔIM is by definition equal to the difference between the output signal level (in dBm) and the corresponding level of the 3rd order IM component (cf Fig. 5-31):

$$L_{out} - D_{out} = \Delta IM \quad [dBm] \quad (5.136)$$

$$\text{In addition: } L_{out} = L_{\Delta} + G \rightarrow L_{\Delta} = L_{out} - G \quad [dBm] \quad (5.137)$$

By plugging the equations (5.136) and (5.137) into (5.135), we get:

$$OIP = \frac{3}{2}L_{out} - \frac{1}{2}D_{out} \quad [dBm] \quad (5.138)$$

If we express OIP not in dBm but directly in Watts we get:

$$OIP = \sqrt[3]{P_{out}^3 / D_{out}} \quad [W] \quad (5.139)$$

The PI3 (in Watts) can thus be calculated from the measurement of the signal power P_{out} and from the IM distortion at the output D_{out} . Knowing OIP (in Watts), we can also express the distortion at the output corresponding to a certain power P_{out} :

$$D_{out} = \frac{P_{out}^3}{OIP^2} \quad [W] \quad (5.140)$$

IM DISTORSION AND THE CORRELATION COEFFICIENT

Fig. 5-33 shows a nonlinear amplifier, subject to a two-tone signal such as that described by (5.89), and producing IM products at its output. The power of the useful signal (in the passband of the system) is P_{in} and the power of the possible IM products falling in the passband and already existing at the input is D_{in} .

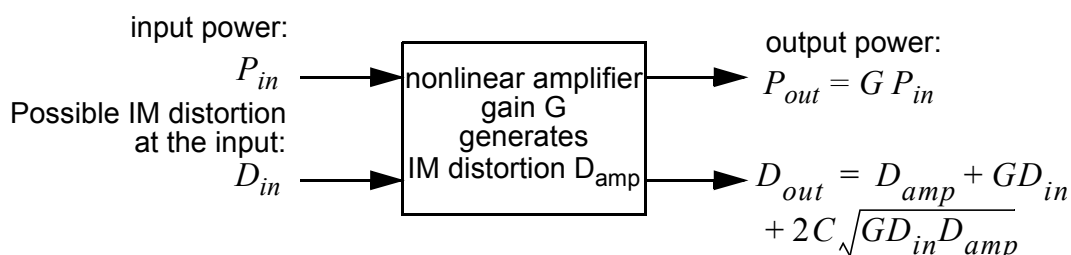


Fig 5-33: *Input and output power of an amplifier subject to a two-tone test.*

The IM distortion at the amplifier output is equal to the distortion produced by the amplifier D_{amp} , plus the distortion already present at the input multiplied by the gain G and an additive term $2C \sqrt{G D_{in} D_{amp}}$ taking into account the correlation existing between the distortion already existing at the input and the distortion generated by the amplifier itself. For the case in which the correlation coefficient C is zero, we have:

$$D_{out} = D_{amp} + G D_{in} \quad [W] \quad (5.141)$$

The two distortion terms add up in power (like independent noise sources). If they are perfectly correlated ($C = 1$), we get:

$$D_{out} = D_{amp} + G D_{in} + 2 \sqrt{G D_{in} D_{amp}} = (\sqrt{D_{amp}} + \sqrt{G D_{in}})^2 \quad [W] \quad (5.142)$$

The two distortion components thus add up in voltage (or in current).

3rd ORDER INTERCEPT POINT FOR AN AMPLIFIER CASCADE (1/2)

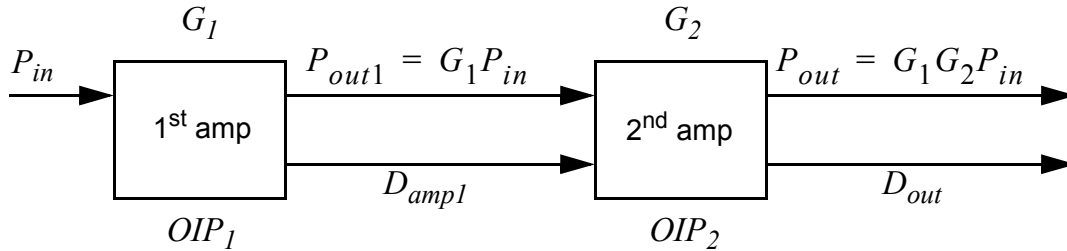


Fig 5-34: Two cascaded amplifiers subject to a two-tone input signal.

Consider the cascade connection of two amplifiers characterized by their gains in power G_i and their 3rd order intercept points OIP_i ($i = 1, 2$). If $C = 1$ we get:

$$D_{out} = (\sqrt{G_2 D_{amp1}} + \sqrt{D_{amp2}})^2 \quad [W] \quad (5.143)$$

From (5.139), the PI3 corresponding to the cascade connection of two amplifiers is:

$$OIP|_{C=1} = \frac{(G_1 G_2 P_{in})^{3/2}}{\sqrt{G_2 D_{amp1}} + \sqrt{D_{amp2}}} \quad [W] \quad (5.144)$$

But from (5.140), we get:

$$D_{amp1} = \frac{(G_1 P_{in})^3}{(OIP_1)^2} \quad \text{and:} \quad D_{amp2} = \frac{(G_1 G_2 P_{in})^3}{(OIP_2)^2} \quad [W] \quad (5.145)$$

and so:

$$OIP|_{C=1} = \left[\frac{1}{G_2 OIP_1} + \frac{1}{OIP_2} \right]^{-1} = \frac{G_2 OIP_1 OIP_2}{G_2 OIP_1 + OIP_2} \quad [W] \quad (5.146)$$

where $G_2 OIP_1$ corresponds to the PI3 of the first amplifier, in relation to the output of the second amplifier. The 3rd order intercept points combine like parallel resistances (for the case in which $C = 1$).

3rd ORDER INTERCEPT POINT FOR AN AMPLIFIER CASCADE (2/2)

If $C = 0$, Eqn. 5.141 gives:

$$D_{out} = D_{amp2} + G_2 D_{amp1} \quad [W] \quad (5.147)$$

and then after (5.139):

$$OIP|_{C=0} = \frac{(G_1 G_2 P_{in})^{3/2}}{\sqrt{G_2 D_{amp1} + D_{amp2}}} \quad [W] \quad (5.148)$$

Plugging in D_{amp1} and D_{amp2} with the help of (5.145), we get:

$$\begin{aligned} OIP|_{C=0} &= \left[\frac{1}{(G_2 OIP_1)^2} + \frac{1}{(OIP_2)^2} \right]^{-1/2} \\ &= \frac{G_2 OIP_1 OIP_2}{\sqrt{(G_2 OIP_1)^2 + (OIP_2)^2}} \quad [W] \end{aligned} \quad (5.149)$$

or brought back to the input:

$$IIP|_{C=0} = \left[\left(\frac{1}{IIP_1} \right)^2 + \left(\frac{G_1}{IIP_2} \right)^2 \right]^{-1/2} \quad (5.150)$$

The denominator of (5.149) being smaller than that of (5.146), the PI3 for $C = 0$ is greater than that obtained for $C = 1$. The real intercept point is usually between two extremes:

$$\frac{G_2 OIP_1 OIP_2}{G_2 OIP_1 + OIP_2} \leq OIP \leq \frac{G_2 OIP_1 OIP_2}{\sqrt{(G_2 OIP_1)^2 + (OIP_2)^2}} \quad (5.151)$$

n^{th} ORDER INTERCEPT POINT FOR AN AMPLIFIER CASCADE

We can generalize (5.150) for the calculation of an n^{th} order intercept point for a chain of amplifiers:

$$\left. \frac{1}{IIP} \right|_{C=0} = \left[\left(\frac{1}{IIP_1} \right)^m + \left(\frac{G_1}{IIP_2} \right)^m + \left(\frac{G_1 G_2}{IIP_3} \right)^m + \dots + \left(\frac{G_1 G_2 \dots G_{r-1}}{IIP_r} \right)^m \right]^{1/m} \quad (5.152)$$

where $m \equiv n - 1$ and IIP (expressed in W) corresponds to the n^{th} order intercept point brought back to the cascade input.