Radio-Frequency Integrated Circuits

Harald Pretl

2025-09-30

Table of contents

1 Introduction

This is the material for an introductory radio-frequency integrated circuits course. The contents are large based on (Razavi 2011) and (Darabi 2020); these two books are an excellent introduction into this topic and are highly recommended! For a generation introduction into RF and microwave (Pozar 2011) is highly recommended.

It is assumed that readers are familiar with the contents of this Analog Circuit Design course.

> ❗ Important
>
> All course material (source code of this document, Jupyter notebooks for calculations, Xschem circuits, etc.) is made publicly available on GitHub (follow this link) and shared under the Apache-2.0 license.
> Please feel free to submit pull requests to fix typos or add content! If you want to discuss something that is not clear, please open an issue.

1.1 Wireless Transmission

In wireless transmission, we usually want to transmit data via a transmitter (TX) and a connected antenna to a receiver (RX). This arrangement is shown in Figure 1.
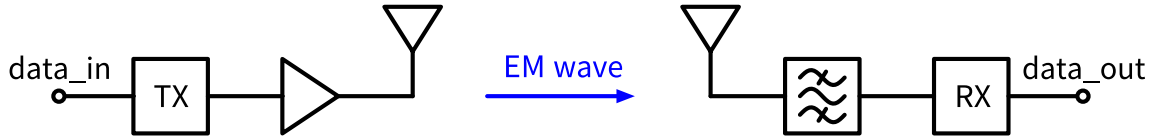

Figure 1: The block diagram of a simple wireless system.

Unfortunately, wireless transmission is hard. The wireless channel, i.e., the usage of electromagnetic waves to transmit information from a transmitter to a receiver, is hard. The wireless channel, while tremendously useful, unfortunately has quite a few undesired features:

- The wireless channels is shared between all users.
- As a consequence, the available bandwidth is shared; this means that bandwidth is a scarce resource.
- The wireless channel has significant losses.
- The channel is time variant, as usually the transmitter and/or the receiver move, and/or the environment changes.

In order to estimate the power P_R of the wireless transmission at the receiver we can use Friis' transmission formula (Pozar 2011):

$$P_\mathrm{R} = P_\mathrm{T} \cdot \frac{A_\mathrm{R} \cdot A_\mathrm{T}}{d^2 \cdot \lambda^2} \tag{1}$$

Here, A_R and A_T are the effective areas of the transmit and the receive antenna, while d is the distance (line of sight) between the two antennas, and λ is the wavelength of the used transmission frequency.

The effective area of an antenna depends on the type and construction, but generally we can say that

$$A \propto \lambda^2$$

For an isotropic antenna (a theoretical construct where the radiation is equal in all directions) $A = \lambda^2/(4\pi)$, while for a $\lambda/2$-dipole $A = 0.13\lambda^2$. Of course, the speed of light c relates frequency f and wavelength λ of an electromagnetic wave by

$$c = \lambda f.$$

Generally speaking, the size of an electromagnetic antenna is proportional to the wavelength of the EM wave use for transmission. For man devices, we seek antenna on the order of a few centimeters, this is way frequencies in the hundreds of MHz to GHz are so popular. Table 1 lists a few typical applications and their frequency and wavelength.

Table 1: Typical RF applications with their operating frequencies and corresponding wavelengths

Application	Frequency	Wavelength
FM Radio	88–108 MHz	2.8-3.4 m
WiFi (lowband)	2.4 GHz	12.5 cm
WiFi (highband)	5 GHz	6 cm
Bluetooth	2.4 GHz	12.5 cm
Cellular	0.6–5 GHz	6-50 cm
GNSS	1.575 GHz	19 cm

As you can see in Table 1 many of these antennas would not fit into the used device form factors, i.e., often we have to use electrically small antennas.

ℹ Note 1: Wavelength Calculation

Let's calculate the wavelength for a Bluetooth signal at 2.4 GHz. Given:

- Frequency $f = 2.4$ GHz $= 2.4 \times 10^9$ Hz

- Speed of light $c = 3 \times 10^8$ m/s

Using the relationship $c = \lambda f$, we can solve for wavelength:

$$\lambda = \frac{c}{f} = \frac{3 \times 10^8 \text{ m/s}}{2.4 \times 10^9 \text{ Hz}} = 0.125 \text{ m} = 12.5 \text{ cm}$$

This means that a quarter-wavelength monopole antenna for 2.4 GHz Bluetooth would be approximately 3.1 cm long, which easily fits into most mobile devices.

In order to get a feeling for the attenuation experienced in wireless communication, we now calculate the following exemplary transmission. We will use the unit of dBm with is often used in RF design and is defined as

$$P|_{\text{dBm}} = 10 \cdot \log_{10} \left(\frac{P|_{\text{W}}}{1\,\text{mW}} \right) \tag{2}$$

> **i** Note 2: Wireless Transmission
>
> We use the following parameters:
>
> - Transmit power $P_{\text{T}} = 1$ W
> - Frequency $f = 2.4$ GHz
> - Communication distance $d = 10$ km
> - Using $\lambda/2$ dipoles on both ends
>
> Using Equation 1 we calculate
>
> $$P_{\text{R}} = P_{\text{T}} \cdot \frac{0.13\lambda^2 \cdot 0.13\lambda^2}{d^2\lambda^2} = P_{\text{T}} \cdot 0.13^2 \left(\frac{\lambda}{d} \right)^2 = 2.64\,\text{pW} = -85.8\,\text{dBm}$$
>
> With the transmit power of 1 W = 30 dBm we have an attenuation of 116 dB! This is a very large number!

As dire as the situation of Note 2 already looks, this is not even all factors considered:

- The given attenuation is for line-of-sight paths; often, the attenuation is significantly higher than this due to blockage by buildings, mountains, rain, or foliage.
- In lack of a direct line-of-sight path, the EM wave is redirected by reflections, causing additional attenuation, and the potential destructive interference by multi-path reception.

The consequences of this are (among others):

- The transmitter needs to generate enough **transmit power** to overcome the transmission loss; this has to be done often with high **efficiency**, as the transmit device is battery operated or limited by cooling.
- The receiver has to be able to process **weak signals**, i.e., the **noise** level of the signal processing has to be very low.
- Often, the receive signal is very weak, while there are strong signals at other frequencies (i.e., other wireless transmitters are located close to the receiver). This means the receiver has to be able to process a weak signal while simultaneously tolerate **large interfering signals** (called blockers).
- Since the frequency spectrum is shared among many users and wireless applications, the transmit information has to be packed efficiently into a **small bandwidth**.
- Very often, wireless devices are battery-operated. This means transmit and receive functions have to be implemented using **minimum power consumption**.

As stated in the beginning, designing wireless systems is hard.

2 Fundamentals

In this section, we will discuss a few important concepts which will be instrumental in the further study of RF circuits and systems. As signals in RF circuits and systems are often limited on the top end by linearity, and on the bottom end by noise, we will discuss these two topics in some detail.

2.1 Linearity

As we have already seen in Section 1.1 the transmitter has to process large signals without distorting them, while the receiver has to process small signals in the presence of large signals. Both situations mean we need metrics and models to quantify and discuss linearity properties.

We are going to use a very simple, time-invariant model to study linearity, based on a Taylor polynomial.

> **!** Linearity and Time Invariance in RF Systems
>
> We use time invariance to simplify the mathematics. In practice, many circuits and systems will show time variant behavior which leads to quite a few very interesting and important phenomena!

We model a nonlinear circuit block with the following Taylor polynomial:

$$y(t) = \alpha_0 + \alpha_1 x(t) + \alpha_2 x(t)^2 + \alpha_3 x(t)^3 + ... \tag{3}$$

Usually, the blocks under study will have higher order nonlinear terms but we often stop at 3rd order to keep things simple. For practical work, higher order terms should be included if necessary.

Which $x(t)$ should we use to study wireless systems? Often, the bandwidth f_{BW} of a transmit signal is much smaller than the center frequency f_0, i.e., $f_{\mathrm{BW}} \ll f_0$. In this case using a sinusoidal signal as a mode is both simple to handle and approximately correct.

2.1.1 Single-Tone Linearity

We thus set (with A being the amplitude of the input signal and $\omega = 2\pi f$ the angular frequency)

$$x(t) = A\cos(\omega t)$$

and insert it into Equation 3. After some simple trigonometric manipulations we are at

$$y(t) = \underbrace{\frac{1}{2}\alpha_2 A^2}_{\text{dc component}} + \underbrace{\left(\alpha_1 A + \frac{3}{4}\alpha_3 A^3\right)\cos(\omega t)}_{\text{fundamental}}$$
$$+ \underbrace{\frac{1}{2}\alpha_2 A^2 \cos(2\omega t)}_{\text{2nd harmonic}} + \underbrace{\frac{1}{4}\alpha_3 A^3 \cos(3\omega t)}_{\text{3rd harmonic}} \tag{4}$$

Looking at Equation 4 we can make a few interesting observations:

- Even-order nonlinearity (α_2) creates low-frequency components; it effectively adds frequency components related to the envelope A. If A is a constant then this results in a dc term; if $A(t)$ is time variant it will created a squared version of it at low frequencies.
- The α_1 term is the gain of the circuit block.
- Odd-order nonlinearity (α_3) can impact the gain of the fundamental term passing through the block. Depending on the sign of α_3 this can lead to gain contraction or expansion.
- Even- and odd-order nonlinearities creates additional frequency components, so-called harmonics of the fundamental frequency. These harmonics are often unwanted, as they are far outside the wanted transmission frequency range, and need to be minimized, by either

  1. use a lowpass filter to filter these harmonics, or
  2. increase the linearity, i.e., make the α_2, α_3, etc., small enough.

The created harmonics are illustrated in Figure 2. Note that measuring harmonics to quantify the nonlinearity metrics like α_2 and α_3 is often not very accurate, as these harmonics are often filtered in bandwidth-limited systems.

How can we quantify the nonlinearity with a one-tone test? We can sweep the input signal $x(t)$ in amplitude, and observe the output $y(t)$. If the observed gain drops by 1 dB from the small-signal value we note the input power, and call this point the **1dB compression point** (P_{1dB}). We should always add whether this 1dB compression point is input- or output-referred to avoid ambiguity. The diagram in Figure 3 shows this test ($\alpha_1 = 100$, $\alpha_3 = -0.2$).

> **❗ Compressive vs. Expansive Behavior**
>
> Note that for compressive behaviour, α_3 and α_1 have different signs, while for expansive behaviour, they have the same sign.
> At some point, every circuit block will show compressive behavior, as the maximum signal amplitude will be limited by power supply voltages, device breakdown voltages, etc.
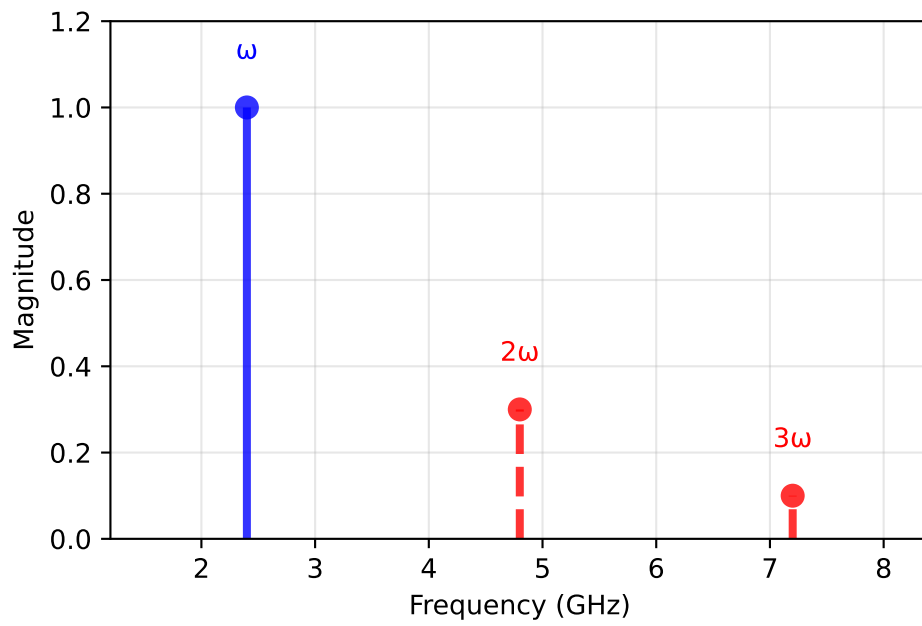
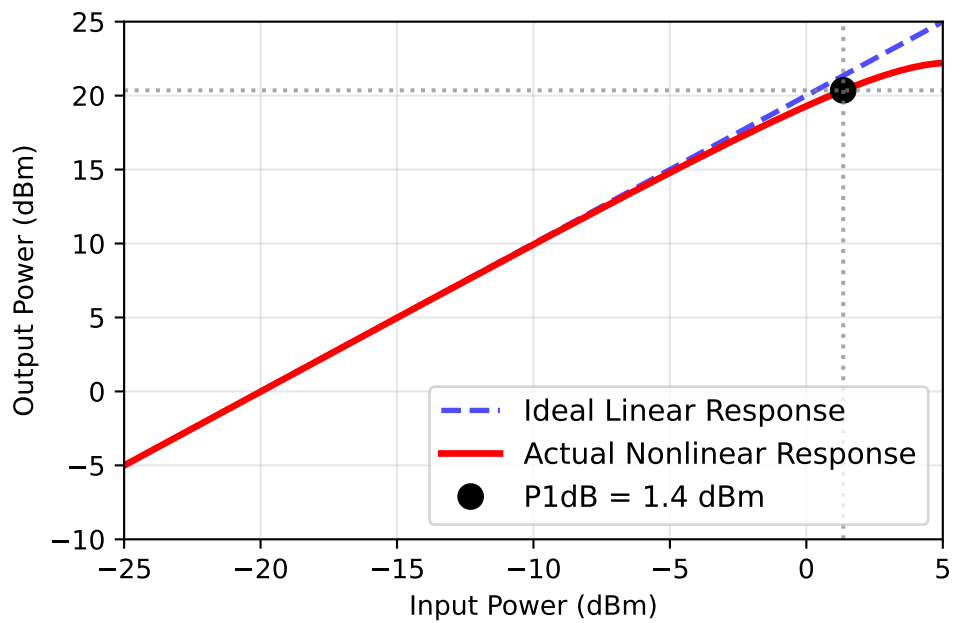Figure 2: Single-tone test showing created harmonics at 2 and 3 .

Figure 3: 1dB compression point test showing input vs output power relationship and the definition of P1dB.

2.1.2 Multi-Tone Linearity

We now elevate our investigations and apply two sinusoids with different frequencies and different amplitudes and see which signals we get at the output of the nonlinear block. The two-tone test and resulting third-order intermodulation products (IM3) are illustrated in Figure 4.

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

We apply the above stimulus to our nonlinear model described by Equation 3 and again, after some trigonometric manipulations, arrive at:

$$y(t) = y'(t) + y''(t) + y'''(t) \tag{5}$$

As many different frequency components are created by this simple two-tone test (and nonlinearity only up to 3rd order) we split the result into different equations and look at the result separately.

First, we start with the fundamental tones:

$$
y'(t) = \left(\underbrace{\alpha_1 A_1 + \frac{3}{4}\alpha_3 A_1^3}_{\text{compression/expansion}} + \underbrace{\frac{3}{2}\alpha_3 A_1 A_2^2}_{\text{cross-modulation/desens}} \right) \cos(\omega_1 t) \\
+ \left(\underbrace{\alpha_1 A_2 + \frac{3}{4}\alpha_3 A_2^3}_{\text{compression/expansion}} + \underbrace{\frac{3}{2}\alpha_3 A_2 A_1^2}_{\text{cross-modulation/desens}} \right) \cos(\omega_2 t)
\tag{6}
$$

As shown in Equation 6, interesting things happen:

- We (again) have the gain compression/expansion effect as already discussed in Section 2.1.1.
- In addition, we have **cross-modulation**, i.e., the envelope of one tone (e.g., $A_2(t)$ of the tone at ω_2) impacts the envelope of the other tone at ω_1. This can lead to unwanted signal distortation, even if there is a large frequency separation between ω_1 and ω_2!
- Further, since the sign of α_3 is usually opposite to α_1, this can also lead to **desensitization** ("desens"). If, for example, $A_2 \gg A_1$, then there would be no compression due to the tone ω_1 itself, however, the large tone at ω_2 will lead to gain compression of the tone at ω_1; this effect is called desense.

We now look at the next class of generated tones:

$$y''(t) = \frac{1}{2}\alpha_2 A_1^2 + \frac{1}{2}\alpha_2 A_2^2$$
$$+ \alpha_2 A_1 A_2 \cos[(\omega_1 - \omega_2)t] \tag{7}$$
$$+ \alpha_2 A_1 A_2 \cos[(\omega_1 + \omega_2)t]$$

As we can see in Equation 7 new tones are created (besides the low frequency components we already know from the single-tone test) at the sum and difference of ω_1 and ω_2. These new frequency components are called "**intermodulation products of second order**" (IM2). These tones are created by the even-order nonlinearity (α_2). These IM2 products are far away from the wanted tones, so are often not very problematic in amplifiers (but there can be exceptions!). However, they can be very problematic in frequency conversion blocks like mixers. We will come back to this point when discussing zero-IF receivers.

We now investigate the next couple of tones:

$$y'''(t) = \frac{3}{4}\alpha_3 A_1^2 A_2 \cos[(2\omega_1 + \omega_2)t]$$
$$+ \frac{3}{4}\alpha_3 A_1^2 A_2 \cos[(2\omega_1 - \omega_2)t]$$
$$+ \frac{3}{4}\alpha_3 A_1 A_2^2 \cos[(2\omega_2 + \omega_1)t] \tag{8}$$
$$+ \frac{3}{4}\alpha_3 A_1 A_2^2 \cos[(2\omega_2 - \omega_1)t]$$

The tones shown in Equation 8 are called "**intermodulation products of third order**" (IM3), and are caused by the odd nonlinearities (like α_3). While the IM3 tones located at $2\omega_1 + \omega_2$ and $\omega_1 + 2\omega_2$ are similar to the sum IM2 tone and far away from ω_1 and ω_2, the other two tones are concerning.

Expressing $\Delta\omega = \omega_2 - \omega_1$ (and assuming $\omega_1 < \omega_2$), the building law of $2\omega_1 - \omega_2 = \omega_1 - \Delta\omega$ and $2\omega_2 - \omega_1 = \omega_2 + \Delta\omega$ results in new tones right besides ω_1 and ω_2, with a frequency separation only defined by $\Delta\omega$. This situation is illustrated in Figure 4.

This close localization of the IM3 tones can also be utilized to characterize nonlinear performance. Using gain compression or harmonic generation (H3) it can be very difficult to extract nonlinearity of third order (α_3). However, using a two-tone test, the IM3 tones can be readily measured, even if the measured signal path shows a **bandpass characteristic**! As RF systems frequently employ bandpass filters to suppress out-of-band signals, this is a very important property of the two-tone test.

The resulting test is called a two-tone test yielding the third-order intercept point (IP3). This test is widely used in RF design to characterize the linearity of amplifiers, mixers, and complete transceiver systems. The power relationship between fundamental tones and IM3 products as a function of input power is shown in Figure 5.
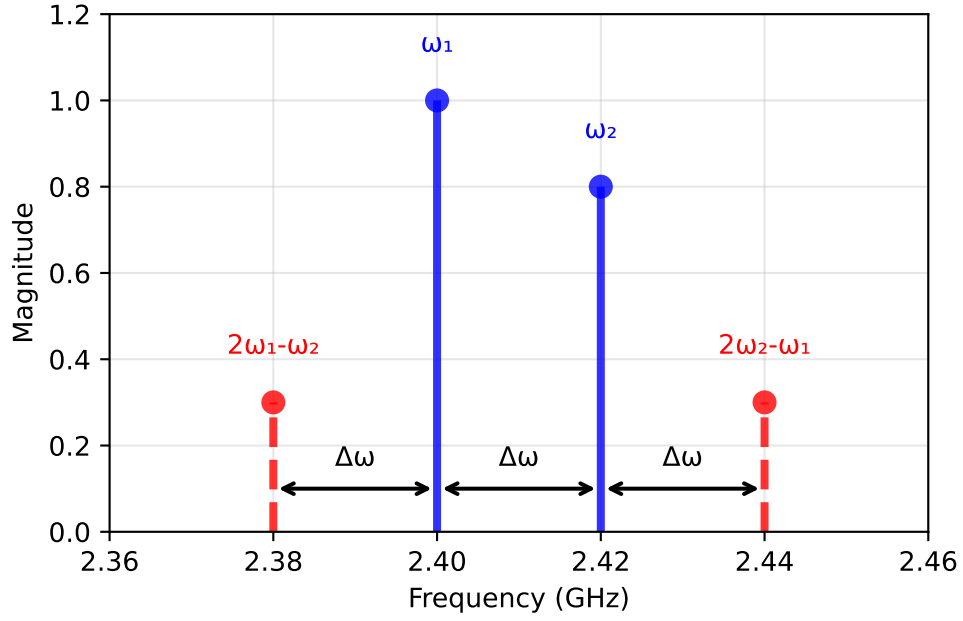
Figure 4: Two-tone test showing fundamental frequencies , and third-order intermodulation products (IM3) at 2 - and 2 - .
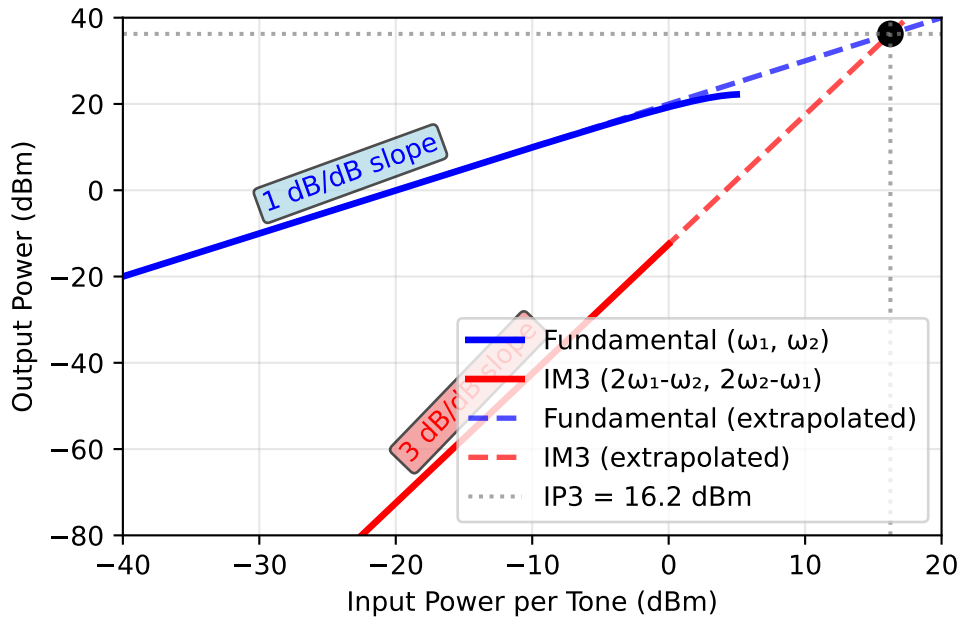

Figure 5: Two-tone IM3 test showing fundamental and IM3 product power vs. input power, with IP3 intercept point definition. Equal input power per tone is assumed.

Note that, as shown in Figure 5, the IM3 products rise with a slope of 3 dB/dB, i.e., if the input power is increased by 1 dB, the IM3 products increase by 3 dB. The fundamental tones rise with a slope of 1 dB/dB (as long as we are in the linear region). The IP3 point is defined as the intersection of the **extrapolated** linear lines of fundamental and IM3 products. As both lines have different slopes, this intersection point is usually far outside the actual operating range of the circuit block under test!

When calculating the IIP3 (input-referred IP3) we can use the following formula, assuming equal input power per tone. It is important to always check the slope of the IM3 products to ensure that we are indeed in the third-order region! If the input power per tone is P_{in} (in dBm) and the input-referred power of one IM3 tone is P_{IM3} (in dBm), then the input-referred IP3 is given by

$$\text{IIP3} = P_{\text{in}} + \frac{P_{\text{in}} - P_{\text{IM3}}}{2} \tag{9}$$

Further, for midly nonlinear systems (i.e., α_3 is dominating), the IIP3 can be approximated from the 1dB compression point as

$$\text{IIP3}|_{\text{dBm}} \approx P_{\text{1dB}}|_{\text{dBm}} + 9.6\,\text{dB}. \tag{10}$$

If we have two blocks which are cascaded, and we know the gain and IIP3 of both blocks, we can calculate the overall IIP3 of the cascade with the following approximation. An exact calculation is very involved, as the nonlinearities of the first block (and the resulting tones) will be processed by the second block, creating even more tones; this process escalates very quickly. However, for practical purposes, the following approximation is often sufficient:

$$\frac{1}{\text{IIP3}_{\text{total}}} \approx \frac{1}{\text{IIP3}_1} + \frac{G_1}{\text{IIP3}_2} + \frac{G_1 G_2}{\text{IIP3}_3} \tag{11}$$

Here G_1 is the linear gain of the first block, and IIP3_1, IIP3_2 are the input-referred IP3 of the first and second block, respectively. Note that all powers have to be in linear units (i.e., Watts) when using Equation 11. An even more simplified version of Equation 11 can be used with all quantities given in dBm and dB, respectively:

$$\text{IIP3}_{\text{total}} \approx \min\{\text{IIP3}_1, \text{IIP3}_2 - G_1, \text{IIP3}_3 - G_1 - G_2\} \tag{12}$$

A typical RF system cascade with multiple blocks and their individual IIP3 contributions is shown in Figure 6.

> **i** Note 3: Simple IIP3 Cascade Calculation
>
> Let's calculate the overall IIP3 of two cascaded blocks. The first block is a low-noise amplifier with an IIP3 of -10 dBm and a gain of 20 dB. The second block is a mixer that has a gain of 10 dB and an IIP3 of 5 dBm. What is the overall IIP3?
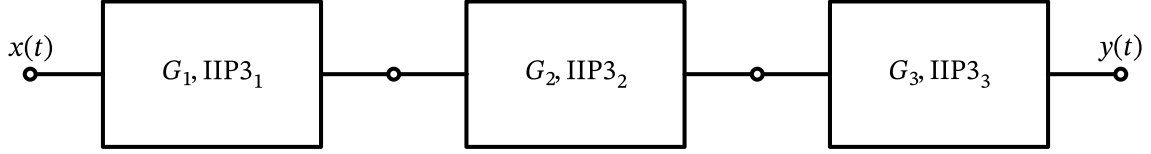
Figure 6: Block cascade for IIP3 calculation showing multiple stages with gains and individual IIP3 values.

Using Equation [12] we can quickly estimate:

$$\text{IIP3}_{\text{total}} \approx \min\{-10\,\text{dBm}, 5\,\text{dBm} - 20\,\text{dB} = -15\,\text{dBm}\} = -15\,\text{dBm}$$

We see that the overall IIP3 is limited by the linearity of the second block, as the first block amplifies all signals (including blockers) by 20 dB before they reach the second block.

2.2 Noise

Just as nonlinearity is a limiting factor for large signals, noise is the limiting factor for small signals. Noise is present in all electronic circuits and systems, and it is impossible to avoid it. However, we can try to minimize its impact on the system performance.

Noise is usually characterized by its power spectral density (PSD) in units of Watts per Hertz (W/Hz). For example, thermal noise at room temperature has a PSD of approximately $kT = 4 \times 10^{-21}$ W/Hz, or –174 dBm/Hz (with the Boltzmann constant $k = 1.38 \times 10^{-23}$ J/K). This means that if we have a bandwidth of 1 MHz, the total thermal noise power would be:

$$P_{\text{thermal}} = \text{PSD} \cdot B = -174\,\text{dBm/Hz} + 10\log_{10}\left(\frac{1\,\text{MHz}}{1\,\text{Hz}}\right) = -114\,\text{dBm}$$

The PSD of noise can be flat vs. frequency (which is called "white noise"), or can decrease with frequency (e.g., "flicker noise" or "1/f noise"). Further, noise can be generated by resistors (thermal noise), semiconductors (shot noise, generation-recombination noise), etc. A detailed discussion of noise sources can be found in (Gray et al. 2009) or (Razavi 2017).

2.2.1 Types of Noise Generation

Resistors generate thermal noise, which is white noise with a PSD of $4kTR$ (in $V^2/$Hz) when looking at the voltage across the resistor, or $4kT/R$ (in $A^2/$Hz) when looking at the current through the resistor. This noise is generated by the random thermal motion of charge carriers in the resistor.

> **!** Thermal Noise
>
> Note that the simple approximation given above is only valid for reasonably high frequencies and typical temperatures, and is known as the Rayleigh-Jeans approximation of Planck's blackbody radiation accounting for quantum effects and is given by (Pozar 2011)
>
> $$\text{PSD} = \frac{hf}{e^{hf/kT} - 1}$$
>
> where h is the Planck constant ($h = 6.626 \times 10^{-34}$ Js) and f is the frequency. The Rayleigh-Jeans approximation is valid for $f \ll kT/h$, which is approximately 6 THz at room temperature (290 K).

MOSFETs generate several types of noise, the most important ones being the thermal noise of the channel and flicker noise.

The thermal noise of the channel can be modeled as a current noise source between drain and source with a PSD of $4kT\gamma g_{d0}$ (in A^2/Hz), where γ is a process-dependent parameter (usually between 2/3 and 2). The parameter g_{d0} is the small-signal output conductance of the MOSFET in triode, i.e., $g_{d0} = g_{ds}$, or equal to $g_{d0} = g_m$ when in saturation.

In saturation, it is often useful to express the thermal noise as a voltage noise source at the gate with a PSD of $4kT\gamma/g_m$ (in V^2/Hz). We can see that we can lower this noise of the MOSFET by increasing the transconductance g_m, which can be achieved by increasing the bias current.

In addition, at high frequencies, the MOSFET also has induced gate-current noise, which is correlated with the channel thermal noise. A detailed discussion of this noise source can be found in (Razavi 2017).

Flicker noise is usually modeled as a voltage noise source at the gate with a PSD of $K_f/(C'_{ox}WLf)$ (in V^2/Hz), where K_f is a process-dependent parameter, C'_{ox} is the oxide capacitance per unit area, L and W are the length and width of the MOSFET, and f is the frequency. Note that we can lower the flicker noise by increasing the area of the MOSFET (WL), however, this increases the parasitic capacitances associated with the MOSFET, and this often prohibitive for RF operation!

In **bipolar junction transistors (BJTs)**, the most important noise source is the shot noise due to the diffusion current in the base-emitter junction. Its PSD can be modeled as a current noise source between collector and emitter with a PSD of $2qI_C$ (in A^2/Hz), where q is the elementary charge ($q = 1.6 \times 10^{-19}$ C) and I_C is the dc collector current.

> **!** Equivalance of Shot and Thermal Noise
>
> Note that it has been shown in (Sarpeshkar, Delbruck, and Mead 1993) that thermal noise and shot noise are actually equivalent, as both are generated by the random, thermally agitated motion of charge carriers!

Ideal **capacitors** and **inductors** do not generate noise, however, real capacitors and inductors have parasitic resistances which generate thermal noise.

In RF systems additional noise sources can be present. One noteworthy example is the **cosmic microwave background** radiation, which can be modeled as a noise temperature of approximately 3 K. While this is negligible compared to thermal noise at room temperature (approximately 290 K), it can be significant in very low-noise systems, such as radio telescopes pointing to the sky. An other important noise source in RF systems is the **atmospheric noise**, which is generated by natural phenomena like lightning or in the ionosphere.

2.2.2 Noise in Impedance-Matched Systems

We now want to calculate the maximum noise power that can be extracted from a noisy source. We assume the following situation as shown in Figure 7. Note that the voltage source $\overline{V_{n,s}^2}$ models the thermal noise of the source resistor R_s resulting in a Thevenin equivalent circuit.
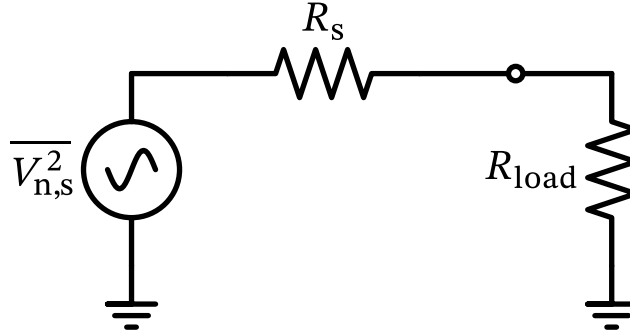

Figure 7: A noise-matched system with source and load impedances.

We know that the noise of the source resistor is given by $\overline{V_{n,s}^2} = 4kTR_s$. We assume the load resistor R_{load} as noiseless and matched to the source resistor, i.e., $R_{load} = R_s$ for maximum power transfer. The noise power spectral density delivered to the load resistor is then given by

$$P_{n,load} = \frac{\overline{V_{n,load}^2}}{R_{load}} = \frac{\overline{V_{n,d}^2}}{4R_s} = kT \tag{13}$$

The calculation of Equation 13 confirms the initial statement that the maximum noise power spectral density that can be extracted from a noisy source is kT (in W/Hz). This result is independent of the actual value of the source resistance R_s.

We can further generalize the thermal noise of any impedance as

$$\overline{V_n^2} = 4kT\Re\{Z\} \tag{14}$$

14

as for example in the complex impedance Z_{ant} of an antenna.

2.2.3 Noise Figure

In RF systems, we often want to quantify the noise performance of a circuit block or a complete system. The most widely used metric is the **noise factor (F)**, which is defined as the ratio of the signal-to-noise ratio (SNR) at the input to the SNR at the output of a circuit block or system. If we express the noise factor in dB, we call it the **noise figure (NF)** (Pozar 2011). The noise factor is given by

$$F = \frac{\text{SNR}_{in}}{\text{SNR}_{out}} = \frac{(P_s/P_n)_{in}}{(P_s/P_n)_{out}} \tag{15}$$

where P_s is the signal power and P_n is the noise power. The noise factor is always larger than or equal to 1 (or 0 dB), as no circuit can improve the SNR!

> **!** SNR Improvement
>
> Note that the SNR can be improved by filtering, as filtering reduces the noise power. If the noise bandwidth is larger than the signal bandwidth, then the SNR can be improved without affecting the signal. However, this is not considered in the noise factor, as the noise factor assumes that both signal and noise pass through the same bandwidth.

Let us look at a simple model of a noise circuit block as shown in Figure 8. The input signal S_{in} is accompanied by noise N_{in}. By definition it is assumed that the input noise power results from a matched resistor at $T_0 = 290\,\text{K}$, so that $N_{in} = kT_0$. The circuit block has a power gain G and adds its own noise N_{dut} to the output signal. For simplicity, we assume that the input and output of the circuit block are impedance matched to avoid reflections.
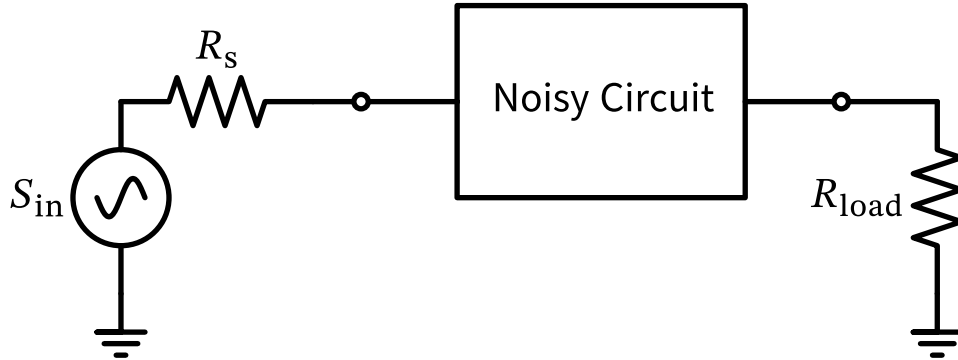

Figure 8: A noise-matched system with source and load impedances and a noisy circuit block.

The output signal and noise powers are then given by

$$S_{\text{out}} = GS_{\text{in}}$$

$$N_{\text{out}} = GN_{\text{in}} + N_{\text{dut}}$$

The resulting noise factor can then be calculated as

$$F = \frac{S_{\text{in}}/N_{\text{in}}}{S_{\text{out}}/N_{\text{out}}} = \frac{1}{G}\frac{GN_{\text{in}} + N_{\text{dut}}}{N_{\text{in}}} = 1 + \frac{N_{\text{dut}}}{GN_{\text{in}}},$$

in other words, the noise factor is 1 plus the ratio of the noise added by the device under test (DUT) to the amplified input noise.

Note that a noiseless block ($N_{\text{dut}} = 0$) has a noise factor of $F = 1$. A passive block with loss factor L (and impedance matched at input and output) has a noise factor of $F = L$ (in linear units), as it attenuates the signal and $N_{\text{out}} = N_{\text{in}} = kT$ if everything is in thermal equilibrium.
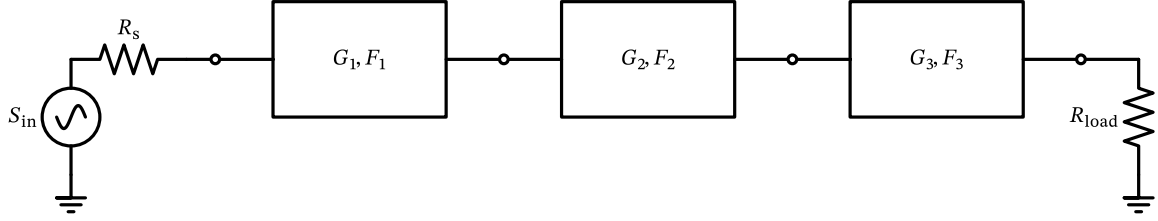

Figure 9: Block cascade for noise factor calculation showing multiple stages with gains and individual noise factors.

If we have a cascade of multiple blocks, as shown in Figure 9, we can calculate the overall noise factor with the **Friis formula** (Pozar 2011)

$$F_{\text{total}} = 1 + (F_1 - 1) + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} \tag{16}$$

where F_i and G_i are the noise factor and power gain of the i-th block, respectively. Note that all gains have to be in linear units (not dB) when using Equation 16. We can interpret Equation 16 as follows:

- The overall noise factor F_{total} is always larger than or equal to the noise factor of the first block (F_1).
- The noise factor of the first block is the most important one, as the noise factors of the following blocks are reduced by the gain of all preceding blocks. This is especially important in RF receivers, where the first block is usually a low-noise amplifier (LNA) with a very low noise figure (e.g., 1 dB or less) and a high gain (e.g., 10 dB or more). This ensures that the noise of the following blocks is negligible.

- The noise factor of the last block is reduced by the gain of all preceding blocks, so it is usually not very important.

Here we also see a trade-off between noise and linearity, as shown by Equation 11 and Equation 16. For low noise, we should try to maximize G_1, however, this will affect linearity (IIP3) in a negative way. As in many other situation in RF design, we have to find a good compromise between conflicting requirements.

2.2.4 Sensitivity

In RF receivers, we often want to know the minimum input signal power that can be detected with a certain SNR. This minimum input signal power is called the **sensitivity** of the receiver. The sensitivity can be calculated as

$$P_{\text{in,min}} = P_{\text{n}} \cdot \text{SNR}_{\text{min}} \cdot F \tag{17}$$

where P_{n} is the noise power at the input, SNR_{min} is the minimum detectable SNR, and F is the noise factor of the receiver. The input noise power can be calculated as

$$P_{\text{n}} = kTB$$

where k is the Boltzmann constant, T is the temperature in Kelvin, and B is the bandwidth of the receiver. Expressing Equation 17 in dBm we get the following formula:

$$P_{\text{in,min}}|_{\text{dBm}} = -174\,\text{dBm/Hz} + \text{NF} + 10\log_{10}(B/\text{Hz}) + \text{SNR}_{\text{min}}|_{\text{dB}} \tag{18}$$

where -174 dBm/Hz is the thermal noise PSD at room temperature (290 K). We can see that the sensitivity improves with lower noise figure, smaller bandwidth, and lower minimum detectable SNR.

> **i** Note 4: Sensitivity Calculation for WiFi
>
> Let's calculate the sensitivity of a WiFi receiver operating at 5 GHz with a bandwidth of $B = 80\,\text{MHz}$, a noise figure of $NF = 7\,\text{dB}$, and a minimum detectable SNR of 25 dB. This high SNR means that a high-order modulation scheme (like 64-QAM) is used for high data rates.
> Using Equation 18 we get:
>
> $$P_{\text{in,min}} = -174\,\text{dBm/Hz} + 7\,\text{dB} + 10\log_{10}(80 \times 10^6) + 25\,\text{dB} \approx -63\,\text{dBm}$$
>
> This means that the minimum input signal power that can be detected by the WiFi receiver is approximately -63 dBm.

Darabi, Hooman. 2020. *Radio Frequency Integrated Circuits and Systems*. 2nd edition. Cambridge University Press.

Gray, Paul R., Paul J. Hurst, Stephen H. Lewis, and Robert G. Meyer. 2009. *Analysis and Design of Analog Integrated Circuits*. Fifth. Wiley.

Pozar, David M. 2011. *Microwave Engineering*. Wiley.

Razavi, Behzad. 2011. *RF Microelectronics*. 2nd edition. Pearson.

———. 2017. *Design of Analog CMOS Integrated Circuits*. McGraw-Hill.

Sarpeshkar, R., T. Delbruck, and C. A. Mead. 1993. "White noise in MOS transistors and resistors." *IEEE Circuits and Devices Magazine* 9 (6): 23–29. https://doi.org/10.1109/101.261888.