ELEC5333M

Wireless Communications Systems Design

25/26

Taught by Dr. Nutapong Somjit

Notes by Jacqueline Walker

Contents

Reading List

- Microwave and Millimetre-wave Design for Wireless Communications, I. Robertson & N. Somjit

Useful Documents

- Unit 1: [0] Module Outline
- Unit 1: [1] Introduction & Link Budget
- Unit 1: [2] Distortion & Nonlinearity
- Unit 1: [3] Noise Analysis
- Unit 1: [4] Simulation & Characterisation
- Unit 1: [5] System Architectures
- Unit 1: [6] Modulation & Demodulation
- Unit 1: [7] RFIC & RF Design
- Unit 1: [8] RFIC & RF Design 2
- Unit 1: [9] Propagation & Antennas

Lecture 1: Introduction

The invention of wireless communication begins with James Clerk Maxwell, who predicted the existence of radio waves, along with various foundational laws useful in the field of Electronic Engineering. His work was followed up by Heinrich Rudolf Hertz, who validated the existence of radio & EM propagation 20 years later and the namesake of the unit Hz. Finally, Guglielmo Marconi who first devised and demonstrated the wireless telegraph. From this point, the spread of wireless communication from 1G-6G and the propagation of Wi-Fi, Bluetooth, etc. became exponential.

1.1: Electromagnetic Spectrum

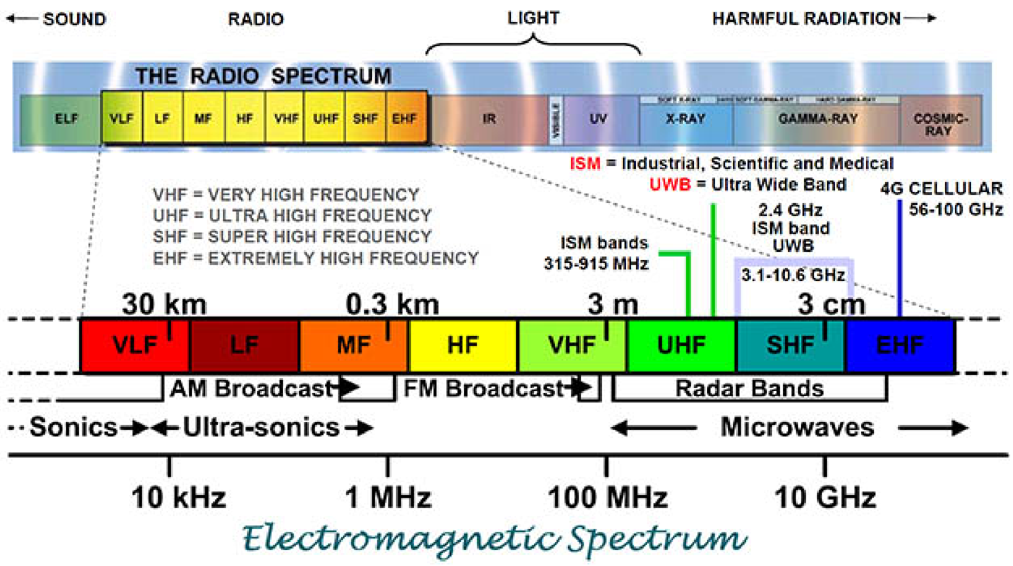


Figure 1: The electromagnetic spectrum, focusing on the radio spectrum within.

The electromagnetic spectrum can be split into multiple segments based on the frequency (and therefore energy & wavelength) of the EM wave. This module will mainly focus on the radio spectrum (3kHz - 3THz). This is because the radio spectrum is the lowest frequency spectrum requiring lower energy to transmit within, and the waves do not have enough energy to cause harm to anything they pass through.

Over time, wireless communication systems end up using higher frequency bands due to the crowding of different systems at lower frequencies. Higher frequency bands have many benefits, such as a larger instantaneuos bandwidth for transmitting data, less interference from other systems, faster speeds for transmission & smaller antennas for smaller wavelengths, and difficulty in jamming (doe to the large BW and high energy needed). However, a higher frequency bandwidth has disadvantages. They require more power than lower frequency systems, resulting in more expensive components which produce more heat and less mature designs & technologies used in construction of those components. Along with this, higher frequency EM waves cannot penetrate through obstacles in the atmosphere like lower frequency EM waves can, resulting in larger atmospheric losses.

Another point is that the atmospheric losses change dynamically due to the weather and the atmospheric conditions of the area you are in. There are also specific peaks of losses due to water vapour and other gases in the air, such as a peak 22GHz due to water vapour or a peak 60GHz due to oxygen.

1.2: Applications of Wireless Technology

Communication is not the only application of wireless communication, albeit a large one. Below lists some applications of the technology:

Radar For miltary purposes, collision detection, surveying, etc.

Navigation For GPS, terrain avoidance, auto-pilot & auto-landing in aircrafts, etc.

Remote Sensing For meteorlogy, forecasting, traffic monitoring, surveillance, etc.

RF Identification For security, access control, animal tracking, etc.

Broadcasting For AM/FM/DAB radio, TV, clock synchronising, etc.

1.3: A Simple Wireless System

For a wireless system, there are three components we look into in this module: the transmitter (Tx), the receiver (Rx), and the wireless channel.

1.4: Decibels, Gains, & Losses

Decibels (dB) are a dimensionless unit for logarithmic ratio between two powers, named after Alexander Graham Bell. It aims to increase the fidelity of the low-volume range and decrease the fidelity of the higher-volume range. Below list common dB conversion formula:

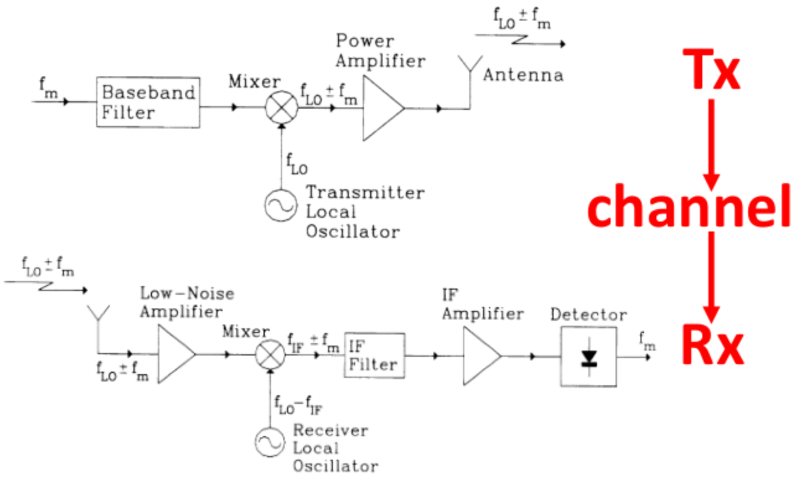


Figure 2: A simple wireless system for transmitting & receiving data.

$$\text{Power(dBm)} = 10\log_{10}(P/P_{\text{ref}}) \qquad \text{(in milliwatts)}$$
$$\text{Voltage(dBV)} = 20\log_{10}(V/V_{\text{ref}}) \qquad \text{(in Volts)}$$

V_{ref} & P_{ref} is the reference voltage/power. For a reference power, either it will be 1 Watt, so the reference power will be 0 dBW, or it will be 1 mW, so the reference power will be 0 dBm. In communications, the usual reference power level is 1mW (so all power gains are in dBm). Ensure that the size of the units match (P & P_{ref} are both in mW).

A trick for dB is that an increase of 3dB is usually a doubling in the linear value and a decrease of 3dB is usually a halving of the linear value.

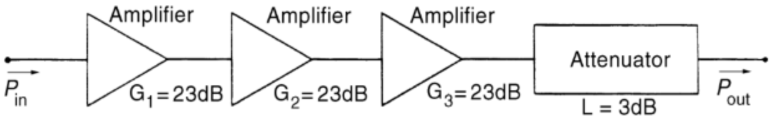
Gains and losses within a control system are usually defined in dB. Gains are an increase in the power (from a reference power), usually via an amplifier. Losses are a decrease in power (from a reference power), usually via an attenuator.

$$\text{Gain}_{\text{linear}} = \frac{\text{output power}}{\text{input power}}$$
$$\text{Loss}_{\text{linear}} = \frac{\text{input power}}{\text{output power}}$$

In decibel scale, gains are added together and losses are subtracted. For linear scale, gains are multiplied together and losses are divided.

Gains and losses within a cascaded circuit (more than one component, where the output of one stage becomes the input of the next stage) may be summed. This is also known as a link budget calculation. In general, this can be seen as:

$$G_T = (G_1 + G_2 + G_3 + \dots) - (L_1 + L_2 + L_3 + \dots)$$



If P_{in} = 1mW = 0dBm

$$\begin{aligned} P_{out} &= P_{in} + G_1 + G_2 + G_3 - L \\ &= 0 \text{ dBm} + 23 \text{ dB} + 23 \text{ dB} + 23 \text{ dB} - 3 \text{ dB} \\ &= 66 \text{ dBm, or } 3981 \text{ W} \end{aligned}$$

P(dBm) = P(dBW) + 30
P(dBW) = P(dBm) – 30

Figure 3: An example closed-loop system with a reference power of 0 dBm.

For a two-port component (one which has one input & one output), the input signal will either pass through to the output, or be reflected back through the input. Insertion Loss (IL) is the loss due to transmission through a component, whereas the Return Loss (RL) is the loss due to reflection from a component (returns back to the previous component). Ideally, the RL should be as low as possible.

$$\begin{aligned} IL &= 10 \log_{10}(P_{in}/P_t) \\ RL &= 10 \log_1 0(P_{in}/P_r) \end{aligned}$$

1.5: Friis' Equation

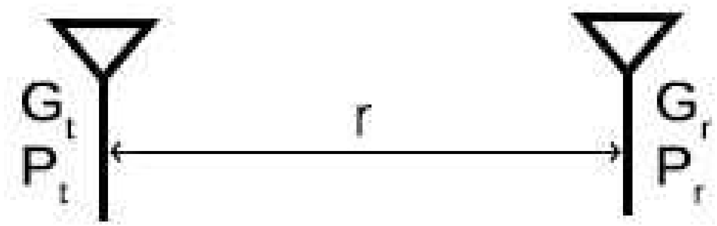


Figure 4: A diagram showing the relevant variables for Friis's Transmission Equation.

Friis' transmission equation is used to calculate the power received after transmission as it relates to the distance between the transmission antenna and the receiver antenna.

$$P_r = P_t \frac{G_t G_r \lambda^2}{(4\pi d)^2}$$

The equation assumes that there is no atmospheric loss, no mismatch in polarisation or impedance, and that there are no obstacles. The antennas must also be reasonably far from each other.

As the signal radiates out from the Tx antenna (we assume it is a perfect antenna), the power density changes depending on how far out from the antenna we are.

For an isotropic Tx antenna, the power density as it reaches the Rx antenna is described as such:

$$S_I = \frac{P_t}{4\pi d^2} \qquad (W/m^2)$$

Instead, if the antenna is directed at the Rx antenna, the signal is not radiated equally and there will be a higher power as it reaches the Rx antenna. This is a gain, but instead of increasing the output power as it does in cascaded circuits, it instead shows how focused the antenna is. The higher the directive-ness of the antenna, the larger the gain will be. For a directive Tx antenna instead, the power density is as such:

$$S_D = \frac{P_t G_t}{4\pi d^2} \qquad (W/m^2)$$

From this equation, we can derive an equation for the (electrical) size of the Rx antenna.

$$P_r = \frac{P_t P_r}{4\pi d^2} A_{er}$$

,where A_{er} is the effective aperture size. We can substitute A_{er} with the Rx gain via:

$$G_r = \frac{4\pi}{\lambda^2} A_{er}$$

$$P_r = P_t \frac{G_t G_r \lambda^2}{(4\pi d^2)}$$

This equation is known as the Friis' power transmission equation. It differs from the previous one as it assumes both antennas are directive.

It would be good to find the maximum distance we can communicate at. The maximum range ($d_{\max}$) can be found by rearranging the equation for distance and setting the receiver power to the minimum signal ($S_{\mathrm{i,min}}$) that is not lost within noise:

$$d_{\text{max}} = \sqrt{\frac{P_t G_t G_r \lambda^2}{4\pi L_{\text{sys}} S_{\text{i,min}}}}$$

, where L_{sys} is the sum of losses within the system (for any inefficiencies from the system).

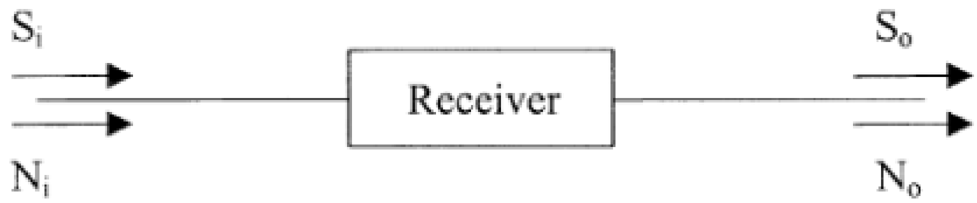


Figure 5: Signal & noise for receiver input & output.

At the receiver, a combination of the noise and the signal are inputted into the receiver. At the input and output, we can derive the signal-to-noise ratio (SNR). We would like our SNR to be as large as possible. The Noise Factor (F) is the change in SNR as it passes through the system (our receiver).

$$F = \frac{\text{SNR}_i}{\text{SNR}_o} = \frac{S_i/N_i}{S_o/N_o}$$

We can then rearrange for S_i to find the minimum power the receiver can detect:

$$S_i = F N_i \frac{S_o}{N_o}$$

The electric noise at the input to the system is the noise from the antenna, which is thermal noise mainly, so $N_i = kTB$ (where the temperature is in Kelvin):

$$S_i = kTBF \frac{S_o}{N_o}$$

Putting this all together, we can derive an equation for the maximum range for a given set of characteristics:

$$d_{\text{max}} = \sqrt{\frac{P_t G_t G_r \lambda^2}{4\pi L_{\text{sys}} kTBF(S_o/N_o)}}$$

We can see that to increase range, we need to increase the power of the transmitters, direct the antennas to be as focused as possible, and keep the noise as minimal as possible by reducing temperature or bandwidth. You should also keep the (major) losses within the system as low as possible.

1.6: Free Space Loss

Free Space Loss (FSL) is the loss due to the spreading of the EM wave as it propagates through free space (aka air). It can be derived from Friis's equation and assumes that the antennas are isotropic (so there are no gains at the antennas). It is commonly rearranged as a ratio:

$$\mathrm{FSL} = \frac{P_t}{P_r} = \left(\frac{4\pi d}{\lambda}\right)^2$$

$$10\log_{10}\left(\frac{P_t}{P_r}\right) = 20\log_{10}\left(\frac{4\pi d}{\lambda}\right)$$

1.7: Link Budget

A link budget is the sum of all gains and losses within a system, leaving only the received signal power in dB. The Friis' equation is a link budget equation for wireless communication between two antennas, only in linear form. Below lists the dB form of Friis' equation:

$$P_r = P_t + G_t + G_r - L_t - L_r - \mathrm{FSL} - L_{\mathrm{sys}}$$

1.8: Measuring System Performance

Effective Isotropic Radiated Power (EIRP) is a measure of how high the energy is radiated at the output of the antenna. It is the same as $P_t G_t$. A larger EIRP means a better system, as you can transmit a high power with great focus (higher gain).

Another measure of system performance is G_r/T, which reflects the ability of the system to receive weak signals in noise.

Lecture 2: Distortion & Non-linearity

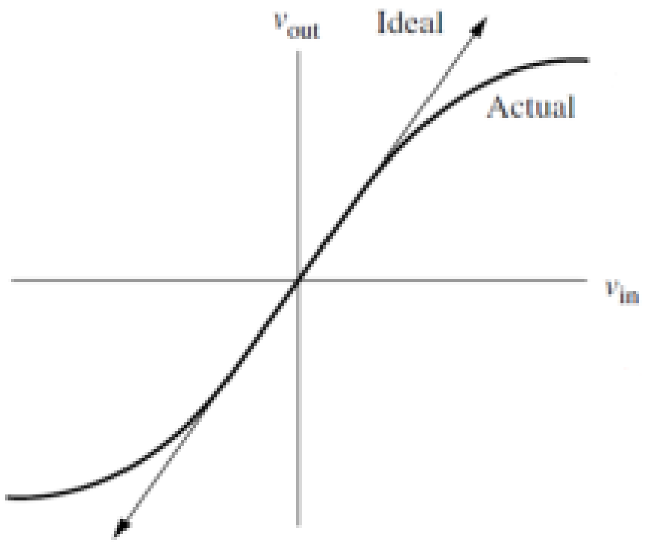


Figure 6: An example of a nonlinear output.

In an ideal world, amplifiers would have a linear relationship between input and output, i.e. $v_o = Gv_i$. This allows for easier prediction of the output voltage for any given input signal, also making system design simpler. However, amplifiers in reality cannot amplify a signal to eternity, as it has a maximum voltage it can amplify to. When the output voltage gets close to the maximum voltage, the output start behaving non-linear. This non-linearity can cause issues such as harmonics, intermodulation points, and intercept points. All active components are non-linear in nature. The non-linearity can be stacked in cascaded circuits, causing further problems.

We can limit the operating conditions for our system (having a minimum/maximum allowed voltage) to cut off any voltages that are non-linear for a pseudo-linear output. Ideally, we'd like to be able to predict at what frequencies these non-linearities occur, and by how much it deviates.

For a nonlinear amplifier, the output is equal to a power series:

$$V_o = a_0 + a_1V_i + a_2V_i^2 + a_3V_i^3 + \cdots + a_nV_i^n$$

The first term, a_0, is a real DC offset. The second term, a_1V_i, is a linear term representing the signal with a voltage gain of a_1. Further terms are unwanted components at different frequencies due to nonlinearity.

In our case, our input signal V_i will be composed of a two-tone signal with a different frequency, amplitude, and phase for both tones. Therefore:

$$V_i = A\cos(\omega_1 t + \theta_1) + B\cos(\omega_2 t + \theta_2)$$

Plugging this into our power series V_o:

$$\begin{aligned} V_o = a_1 &+ a_2(A\cos(\omega_1 t + \theta_1) + B\cos(\omega_2 t + \theta_2)) \\ &+ a_3(A\cos(\omega_1 t + \theta_1) + B\cos(\omega_2 t + \theta_2)) + \ldots \end{aligned}$$

2.1: Second-Order Terms

Expanding out the second order terms gives:

$$\begin{aligned} V_o = a_2 \Bigg[& \frac{A^2 + B^2}{2} + \frac{A^2}{2}\cos(2\omega_1 t + \theta_1) + \frac{B^2}{2}\cos(2\omega_2 t + \theta_2) \\ & + AB(\cos([\omega_1 - \omega_2]t + [\theta_1 - \theta_2]) + \cos([\omega_1 + \omega_2]t + [\theta_1 + \theta_2])) \Bigg] \end{aligned}$$

We can see from visual inspection that we have unwanted frequencies in the second-order terms. The first term is the DC term, the second & third terms are the 'second harmonics' of the signal, and the last terms are the 'intermodulation products'. If we plot the frequency domain of these signals, we can see the unwanted harmonics & intermodulations of the signal.

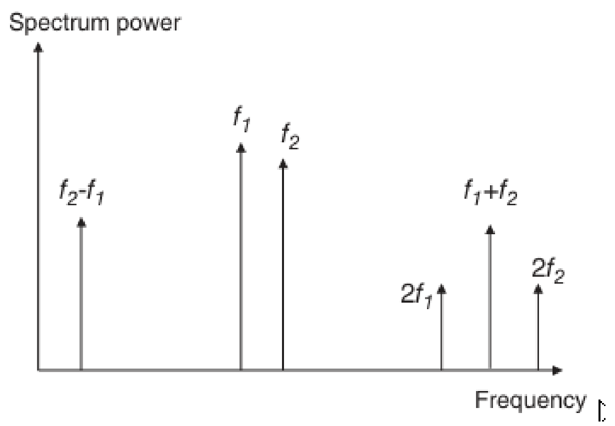


Figure 7: A frequency-domain plot of the second-term signals.

If the two-tones of the input signal has equal amplitudes, then the amplitudes of the intermodulations (IMs) & harmonics of the signal will increase at a greater rate than that of the fundamental frequencies of the signal. This is because the amplitude for those terms increase with the square whereas the fundamental frequencies increase linearly. This is best shown in power-domain.

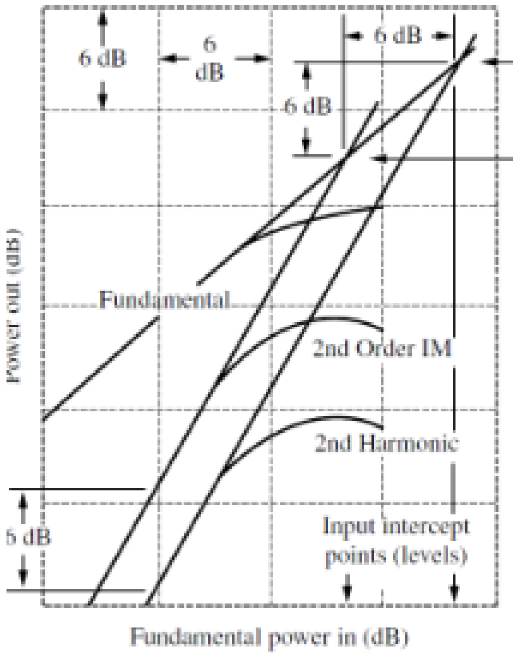


Figure 8: A graph of harmonics & IMs in power domain, where the two input signals have equal amplitude. The straight lines are extraploated from when we assume the output is linear. The curved points are the real output of the harmonics.

As can be seen, the output power for the harmonics & IMs increases twice as fast as the fundamental frequency increases. Assuming that the power increases linearly (ideally with no roll-off), the lines for the harmonics & IMs will intersect with the line for the fundamental frequencies. The point at which the harmonics/IMs intersect with the fundamental frequencies is known as the Intercept Point (IP).

From these intercept points, we can determine the output power for the fundamental frequencies & the harmonics & IMs separately instead of together. If the output power is x dBs below the second-harmonic IP, then the harmonic/IP at that output power will be $2x$ dB below that IP. For example, an IP of 17dB and an output at -8 dBm, then $x = 25$dB and the second-harmonic will be at -33dB. Thus, the signal will be at the midpoint between the IP and the corresponding harmonic or IM level. This then shows that the power of the IMs depends on the IP.

2.2: Third-Order Terms

The third-order terms gives :

$$V_o = a_3 a_2 \left[\frac{A^2 + B^2}{2} + \frac{A^2}{2}\cos(2\omega_1 t) + \frac{B^2}{2}\cos(2\omega_1 t) + AB[\cos((\omega_1 - \omega_2)t) + \cos((\omega_1 + \omega_2)t)] \right] (A\cos(\omega_1 t) + B\cos(\omega_2 t))$$

When expanded out, we get two terms of the fundamental frequency with non-linear amplitudes, four terms representing the third-order IMs, and two terms that are the third-order harmonics. Looking in frequency-domain, a pair of the IMs are very close to the fundamental frequencies and are hard to be cleanly filtered.

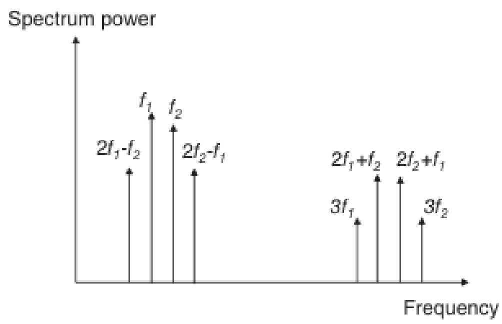


Figure 9: The third-order frequency spectrum.

The power domain shows that the slopes for the harmonics & IMs are steeper than for second-order harmonics. As with the second-order, the harmonics are $3x$ below the IP. Thus, the signal level is 1/3 between the IP and the corresponding harmonic or IM level (closer to the IP than the harmonic/IM).

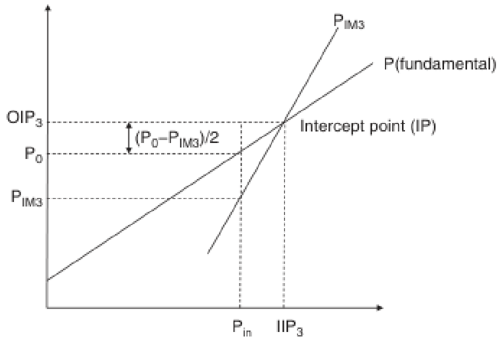


Figure 10: The third-order power graph.

2.3: Cascaded Circuits

For a component, nonlinearities are present within the component. For our purposes, we model this as an addition of nonlinearity to the expected output of the circuit. The nonlinearity is amplified by any future gains in the circuit.

For calculating the IM intercept point in cascaded circuits, we must multiply each stage by the gaim it is affected by. For a two-module cascaded circuit with second-order IM with random phases, the output power for the cascaded intercept point is:

$$\frac{1}{OP_{IM2,tot}} = \frac{1}{g_2 OP_{IM2,1}} + \frac{1}{OP_{IM2,2}}$$

, where the subscript denotes that it is the IM at the second order.

If the phases are coherent, i.e. the IMs add constructively, the output power is:

$$\frac{1}{OP^{1/2}_{IM2,tot}} = \frac{1}{g_2 OP^{1/2}_{IM2,1}} + \frac{1}{OP^{1/2}_{IM2,2}}$$

It can be seen that coherent phases is the worst-case scenario, as the output power of the IMs is squared. For third-order and higher IM, the coherent phase output is the same as the random phase, so no square root.

2.4: 1-dB Compression

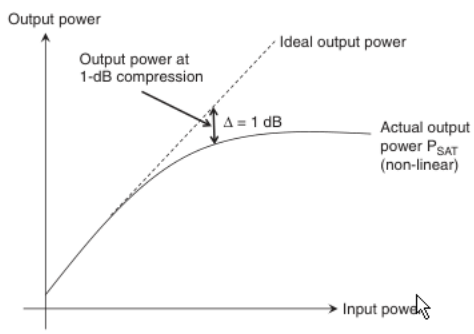


Figure 11: The 1-dB compression point.

The 1-dB compression is where the real ouput power is 1-dB below the expected linear power level. It indicates when the real power level is saturated and starts to behave non-linearly.

2.5: Measuring the Non-Linearity of Components

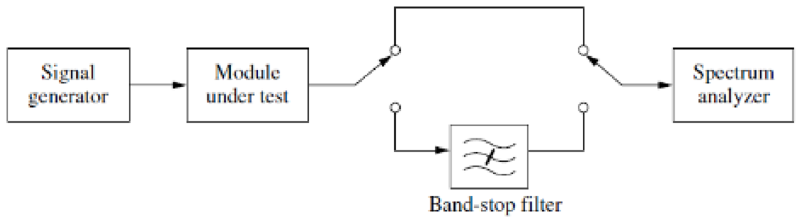


Figure 12: A circuit for measuring IMs of a component.

The IM of a component may be measured via the above circuit. First, the power of the component is measured using a spectrum analyser. The signal generator power is increased until the IPs can be found. The bandstop filter is used to filter out the IMs & harmonics of the signal generator & other components in the circuit.

Lecture 3: Noise

Understanding the noise added to a signal is critical when designing wireless communication systems. An important measurement is the signal-to-noise ratio.

The noise factor is the ratio between the SNR at the input against the output:

$$F = \frac{SNR_i}{SNR_o}$$

The noise figure is the decibel version of noise factor:

$$NF = 10\log_{10}(f)$$

The input noise to a system is usually thermal noise (kTB, where $T = 290K$). The noise temperature is equal to Boltzmann's constant and the temperature (kT). It is useful for categorising the shape of noise from a system.

3.1: Cascaded Circuits

For a cascaded circuit, we want to know the noise factor from the perspective at the input. This is because the noise from the first stages are amplified the most and therefore have a greater bearing on the final result.

For a cascaded circuit composed of 3 modules:

$$F_{\text{casc}} = F_1 + \frac{F_2 - 1}{g_1} + \frac{F_3 - 1}{g_1 g_2}$$

It is also useful to determine the temperature of the system, as each stage in a cascaded circuit will contribute to the temperature. Like noise, temperature is from the perspective at the input. For a 3 module cascaded circuit:

$$T_{sys} = T_s + \frac{T_1}{g_1} + \frac{T_2}{g_1 g_2} + \frac{T_3}{g_1 g_2 g_3}$$

3.2: Noise Colours & Types

Noise in electronic circuits comes from a variety of sources. Below lists three common types:

Thermal noise Generated by random motion of charge due to temperature. Approximately white in colour.

Shot noise Random fluctuations of charge when moving in free space between energy states (capacitators).

Flicker noise Noise from a variety of effects near DC frequency.

Noise is usually described in terms of colours, to depict how the power of the noise changes over frequency.

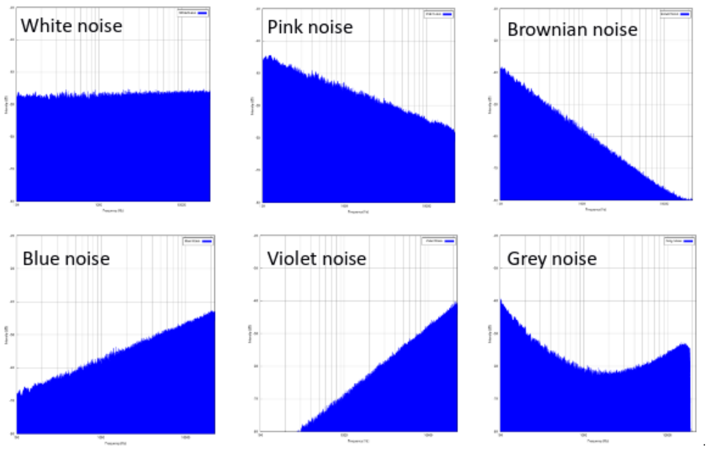


Figure 13: Colours of noise.

Lecture 4: Simulation & Characterisation

Lecture 5: System Architectures