This chapter introduces some basic concepts and required theories that will be used in the following chapters. Section 2.1 explains basic concepts in communication systems, including noise, linearity, and dynamic range which will be frequently used in Chaps. 3 and 4. Section 2.2 explains phase modulation basics, which will be used as the guideline to analysis and explain the multiplexing phased-array system in Chap. 4. Section 2.3 discusses the basic theory of phased-array.

### 2.1 Receiver System Basics

Noise and linearity are the most frequently used concepts in receiver designs. Low noise and high linearity are desired and demanded in most communication systems. However, to achieve low noise and high linearity is not always easy.

#### 2.1.1 Noise

The noise performance of the receiver is measured with noise factor \( F \), which is a measure of how much the signal-to-noise ratio is degraded through the system [31]. We note that

\[
F = \frac{SNR_{in}}{SNR_{out}} = \frac{S_{in}/N_{in,source}}{(S_{in} \cdot G)/N_{out,total}} = \frac{N_{out,total}}{N_{out,source}} = 1 + \frac{N_{out,added}}{N_{out,source}} \tag{2.1}
\]

where \( S_{in} \) is the available input signal power, \( G \) is the available power gain, \( N_{out, total} \) is the total noise power at the output, \( N_{out, source} \) is the noise power at the output originating at the source, and \( N_{out, add} \) is the noise power at the output added by the electronic circuitry. This shows that the minimum possible noise factor,
which occurs if the electronics adds no noise, is equal to 1. Noise figure $NF$ is related to noise factor $F$ by

$$NF = 10 \log_{10} F$$ (2.2)

It can be derived that $NF$ is the ratio of the receiver’s signal-to-noise ratio (SNR) at the output to that at the input, which can be expressed in dB format as follows

$$SNR_{out, dB} = SNR_{in, dB} - NF$$ (2.3)

Equation (2.3) indicates that the $NF$ represents the amount of SNR degradation after the signal is processed by the receiver.

In Fig. 2.1, assuming that all stages are matched to the system characteristic impedance, the overall noise factor of the system is determined by the gain and noise factor of each stage as in (2.4), and the overall gain of the system is shown in (2.5)

$$F_{total} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \cdots + \frac{F_n - 1}{G_1 G_2 \cdots G_{n-1}}$$ (2.4)

$$G_{total} = G_1 \cdot G_2 \cdots G_{n-1} \cdot G_n$$ (2.5)

Equation (2.4) is known as Friis’s formula [32], which indicates that the noise factor of the first stage is most critical to the system noise performance because the noise due to each cascade stage is suppressed by the available power gain preceding it.

Figure 2.2 shows the equivalent noise model of a single receiver stage. $N_{eq, in}$ is the input equivalent noise, and $N_{eq, out}$ is the output equivalent noise.

The output equivalent noise can be expressed as

$$N_{eq, out, dBm} = N_{eq, in, dBm} + G_{dB} = N_{floor, dBm} + NF + G_{dB}$$ (2.6)

where $N_{floor}$ represents the noise floor of the stage. In a cascaded system (Fig. 2.1), the output of one stage feeds the input of the next. The total output equivalent noise can be expressed as

$$N_{total, eq, out, dBm} = 10 \log(kT \cdot BW) + NF_{total} + G_{total, dB} = -174 dBm + 10 \log(BW) + NF_{total} + G_{total, dB}$$ (2.7)
where $kT^* BW$ is the receiver input noise floor, and $NF_{total}$ and $G_{total}$ are the system total noise figure and gain, respectively. In (2.7), $k = 1.38*10^{-23} \text{ J/K}$ is the Boltzmann’s constant [33]. $T$ is the temperature. $BW$ is the bandwidth in Hertz. $kT$ corresponds to the minimum equivalent noise per Herz for a receiver at room temperature (290 K), that is $-174 \text{ dBm/Hz}$. $NF_{total}$ is the total noise figure of the system, and it is derived in (2.4). $G_{total}$ is the total available gain (in dB) of the system, and it is derived in (2.5).

### 2.1.2 Non-Linearity

Any unwanted signal fed into a receiver is called interference and it generally degrades the signal to noise ratio of the wanted signal. Most interference comes from the signals intended for other users or other applications. The interference power can be orders of magnitude higher than the desired signal power and may corrupt the signal as a result of receiving non-linear behavior. Any real receiver is a nonlinear system that responds linearly only if the input signal is sufficiently small. When the input signal increases beyond some extent, the non-linear behavior of the receiver becomes evident.

If a sinusoid is applied to a nonlinear system, the output generally exhibits frequency components that are integer multiples of the input frequency. They are called harmonics of the input frequency.

For simplicity, assuming nonlinear property of the system can be written as Taylor expansion, we limit our analysis to third order, and assume nonlinear terms above the third order are negligible, $y(t)$ in Fig. 2.3 can be derived as

$$y(t) = a_1 A \cos(\omega t) + a_2 A^2 \cos^2(\omega t) + a_3 A^3 \cos^3(\omega t)$$

$$= \frac{a_2 A^2}{2} \cos(\omega t) + \left( a_1 + \frac{3a_3 A^3}{4} \right) \cos(\omega t) + \frac{a_2 A^2}{2} \cos(2\omega t) + \frac{a_3 A^3}{4} \cos(3\omega t)$$

(2.8)

One figure of merit for receiver linearity is the gain compression point. Theoretically, the receiver’s output power increases linearly with the injected input power regardless of the input power level, as shown in Fig. 2.4 [34] by the dashed line. The solid line in Fig. 2.4 depicts a typical input/output transfer function of a real receiver.

It can be seen that at low input power level, the real I/O curve can be approximated with the straight line. As $P_{in}$ increases, $P_{out}$ gradually deviates from the linear curve and is eventually saturated. The point at which $P_{out}$ is 1 dB lower than its
linear theoretical value is called the input 1-dB compression point (ICP1 dB). The importance of this point is that it indicates where the receiver starts to leave the linear region and the saturation becomes a potentially serious problem. The receiver also generates spurs at the harmonics of the signal frequency when the receiver goes into compression.

Figure 2.5 shows two closely spaced interferences at $f_1$ and $f_2$ in the vicinity of signal band, where the strongest interference commonly originates. After passing the nonlinear system, the output signal $y_{two}(t)$ can be derived as

$$y_{two}(t) = x_1x_{two}(t) + x_2^2x_{two}(t) + x_3^3x_{two}(t)$$

$$= x_2A^2 \quad \cdots \quad (DC)$$

$$+ A \left( x_1 + \frac{9}{4} x_3A^2 \right) \left[ \cos \omega_1 t + \cos \omega_2 t \right] \quad \cdots \quad (Fundamental)$$

$$+ \frac{1}{2} x_2A^2 \left[ \cos 2\omega_1 t + \cos 2\omega_2 t \right] \quad \cdots \quad (HD2)$$

$$+ x_2A^2 \left[ \cos (\omega_1 + \omega_2)t + \cos (\omega_1 - \omega_2)t \right] \quad \cdots \quad (IM2)$$

$$+ \frac{1}{4} x_3A^3 \left[ \cos 3\omega_1 t + \cos 3\omega_2 t \right] \quad \cdots \quad (HD3)$$

$$+ \frac{3}{4} x_3A^3 \left\{ \left[ \cos (2\omega_1 + \omega_2)t + \cos (2\omega_1 - \omega_2)t \right] + \left[ \cos (2\omega_2 + \omega_1)t + \cos (2\omega_2 - \omega_1)t \right] \right\} \quad \cdots \quad (IM3)$$

(2.9)

One of the important linearity specifications in (2.9) is the third-order intermodulation point (IM3). When the interference power is high enough, the receiver generates noticeable spurs at $nf_1 \pm mf_2$ due to intermodulation, where $n$ and $m$ are integers including zero. Two of these spurs, located at $2f_1 - f_2$ and $2f_2 - f_1$. 
are particularly threatening to the received signal because they can fall into the signal band and become impossible to eliminate by filtering. The power of the 3rd order distortion increases 3 dB per 1 dB increase of the input power. Figure 2.6 shows the typical curves of the main tone and the third-order intermodulation power as a function of \( P_{in} \).

The third-order interception point is obtained by extrapolating the main-tone output at the slope of 1 dB/1 dB and the third-order distortion curve at 3 dB/1 dB from the low input power level until they intersect with each other, as shown in Fig. 2.6. The x-coordinate of the intersection point is called the input referred third-order interception point (IIP3), and the y-coordinate is called the output referred third-order interception point (OIP3).

In a cascaded system as shown in Fig. 2.1, the overall IIP3 of the system is given by

\[
\frac{1}{IIP_{3, total}} = \frac{1}{IIP_{3,1}} + \frac{G_1}{IIP_{3,2}} + \frac{G_1G_2}{IIP_{3,3}} + \cdots + \frac{G_1G_2G_3\cdots G_{n-1}}{IIP_{3,n}}
\]  

(2.10)

It can be seen from (2.10) that in a cascade system the linearity requirements on the receiver components at the back-end are more stringent because their effects on the overall system are ‘magnified’ by the preceding gain. We should emphasize that (2.10) is merely an approximation. In practice, more precise calculations or simulations must be performed to predict the overall IP3.

### 2.1.3 Dynamic Range

Dynamic range (DR) is defined as the ratio of the maximum input power level that the circuit can tolerate to the minimum input power level that the circuit can properly detect [35]. DR specifies how well the system can handle signals with various power levels.

The lower bound of the dynamic range is set by the receiver sensitivity, defined as the lowest input signal power a receiver can appropriately process. To calculate the receiver sensitivity, one starts from the maximum bit error rate (BER) the data transmission can tolerate in the absence of interference. To achieve this BER, the receiver must provide a minimum \( SNR_{out} \) to the subsequent demodulator. Therefore, a minimum \( SNR_{in} \) must be achieved at the receiver input, which is given by

\[
SNR_{in, min, dB} = SNR_{out, min, dB} + NF_{total}
\]  

(2.11)

Assuming the receiver input is impedance matched to the antenna, the receiver sensitivity can be obtained as
The upper limit of the dynamic range has various definitions that result in different bounds [36], but all are related to the linearity of the receiver. For instance, the most common definition, the spurious-free dynamic range (SFDR), defines the maximum allowed input signal power as the one causing the maximum intermodulation product equal to the output noise power. From Fig. 2.6, this input power level can be solved by using the graphical method, which is given by

\[
P_{in,\text{max,}\text{dBm}} = \frac{2}{3} IIP_{3,\text{total,}\text{dB}} + \frac{1}{3} [NF_{\text{total}} - 174dBm + 10\log(BW)]
\]  

(2.12)

From (2.12) and (2.13), the receiver dynamic range can be found by

\[
DR_{dB} = P_{in,\text{max,}\text{dBm}} - P_{in,\text{min,}\text{dBm}}
\]

\[
= \frac{2}{3} [IIP_{3,\text{total,}\text{dB}} - NF_{\text{total}} + 174dBm - 10\log(BW)] - SNR_{\text{out,\text{min,}\text{dB}}}
\]

(2.14)

### 2.2 Phase Modulation Basics

Modulation is the process of modifying a high-frequency signal (called the carrier signal) with low-frequency information (called the modulating signal). The two most common types of modulation are amplitude modulation (AM) and frequency modulation (FM) [37]. These two forms of modulation modify the carrier’s amplitude or frequency, respectively, according to the instantaneous value of the modulating signal. Phase modulation (PM) is similar to frequency modulation (FM) except that the phase of the carrier waveform is varied, rather than its frequency.
Assume carrier signal $v_c(t)$ and modulating signal $v_m(t)$

$$v_c(t) = V_c \cos(\theta_c(t))$$

$$= V_c \cos(2\pi f_c t + \phi_c)$$

(2.15)

$$v_m(t) = V_m \cos(2\pi f_m t)$$

(2.16)

where $V, f,$ and $\phi$ are the amplitude, frequency and phase, respectively. Combining (2.15) and (2.16), the phase modulated signal in time domain is given by

$$v_{pm}(t) = V_c \cdot \cos[2\pi f_c t + \phi_c + k_p \cdot v_m(t)]$$

(2.17)

The instantaneous phase $\phi_i$ of the carrier is

$$\phi_i = \phi_c + k_p \cdot v_m(t)$$

(2.18)

where $k_p$ is the change in carrier phase per volt of modulating signal, called phase sensitivity (rad/volt). $\phi_c$ is usually 0. Defining $\beta$ as the phase deviation, the max amount by which the carrier phase deviates from its unmodulated value, we get

$$\beta = k_p \cdot |v_m(t)|_{max} = k_p \cdot V_m$$

(2.19)

Substituting (2.19) into (2.17), the phase modulated signal can be expressed as

$$v_{pm}(t) = V_c \cdot \cos[2\pi f_c t + \beta \cdot \cos(2\pi f_m t)]$$

(2.20)

Expanding the above equation with Fourier analysis, and using the Bessel function [38] to determine the spectrum of a phase modulated signal, we achieve

$$v_{pm}(t) = V_c \cdot J_0(\beta) \cdot \cos(2\pi f_c t)$$

$$+ V_c \cdot J_1(\beta) \cdot \left\{ \cos[2\pi (f_c + f_m) t + \pi/2] + \cos[2\pi (f_c - f_m) t + \pi/2] \right\}$$

$$+ V_c \cdot J_2(\beta) \cdot \left\{ \cos[2\pi (f_c + 2f_m) t + \pi] + \cos[2\pi (f_c - 2f_m) t + \pi] \right\}$$

$$+ V_c \cdot J_3(\beta) \cdot \left\{ \cos[2\pi (f_c + 3f_m) t - \pi/2] + \cos[2\pi (f_c - 3f_m) t - \pi/2] \right\}$$

$$+ V_c \cdot J_4(\beta) \cdot \left\{ \cos[2\pi (f_c + 4f_m) t] + \cos[2\pi (f_c - 4f_m) t] \right\}$$

$$+ V_c \cdot J_5(\beta) \cdot \left\{ \cos[2\pi (f_c + 5f_m) t + \pi/2] + \cos[2\pi (f_c - 5f_m) t + \pi/2] \right\}$$

$$+ \cdots$$

(2.21)

Figure 2.7 shows the Bessel function $J_n(\beta)$ versus $\beta$ for $n = 0$ to $n = 6$. Some properties of the Bessel function can be discovered as follows:

- The higher side frequencies are insignificant in the PM spectrum when $\beta$ is low.
- When $\beta \leq 0.25$, only $J_0(\beta), J_1(\beta)$ have a significant value.
The power in a sinusoidal signal depends only on its amplitude and is independent of frequency and phase. It follows that the power in a PM signal equals the power in the un-modulated carrier

\[ P_{PM} = \frac{1}{2} V_c^2 \]  

The total power in a PM signal is the sum of the power of the sidebands and the carrier power. Hence, for the 1-tone modulation, the total power can also be obtained by summing the power in all spectral components in the PM spectrum

\[ P_{PM} = \frac{1}{2} V_c^2 = \frac{1}{2} V_c^2 \left[ J_0^2(\beta) + 2 \sum_{n=1}^{\infty} J_n^2(\beta) \right] \]  

Obviously, the power in the side frequencies is obtained only at the expense of the carrier power

\[ J_0^2(\beta) + 2 \sum_{n=1}^{\infty} J_n^2(\beta) = 1 \]  

Power contained in the carrier frequency and the first \( N \) pairs of side frequencies is given by
Because the exact spectrum of the phase modulated signals is difficult to evaluate in general, formulas for the approximation of the spectra are very useful. As a rule-of-thumb, when \( N = \beta + 1 \)

\[
\begin{align*}
    r_N &= J_0^2(\beta) + 2 \sum_{n=1}^{N} J_n^2(\beta) \\
    r_{\beta+1} &= J_0^2(\beta) + 2 \sum_{n=1}^{\beta+1} J_n^2(\beta) = 0.9844
\end{align*}
\]  

Equation (2.26) indicates that approximately 98% of the power of a phase modulated signal lies within the bandwidth covered by the first \( N = \beta + 1 \) pairs of side frequencies. It is the minimum number of pairs of side frequencies that along with \( f_c \), account for 98% of the total PM power. Carson’s bandwidth can be defined as

\[
BW_c = 2 \cdot (\beta + 1) \cdot f_m
\]  

This formula gives a rule-of-thumb expression for evaluating the transmission bandwidth of PM signals; it is called Carson’s rule [39]. It gives an easy way to compute the effective bandwidth of PM signals from power perspective. In later chapters, this method will be used to evaluate the effective bandwidth of the time multiplexed receiver.

### 2.3 Phased-Array Basics

Phased-array antenna systems is one of the widely used multiple antenna systems in high frequency applications. In wave theory, a phased-array is a group of antennas in which the relative phases of the respective signals feeding the antennas are varied in such a way that the effective radiation pattern of the array is reinforced in a desired direction and suppressed in undesired directions [40]. Comparing with a conventional single path antenna system, two of the main benefits that a phased-array can provide are signal to noise ratio (SNR) enhancement and interference suppression [41–47] as a result of beam-forming.

A phased-array receiver consists of several signal paths, each connected to a separate antenna. Generally, radiated signal arrives at spatially separated antenna elements at different times. An ideal phased-array compensates for the time-delay difference between the elements and combines the signal coherently to enhance the reception from the desired direction, while rejecting emissions from other directions. The antenna elements of the array can be arranged in different spatial configurations.

Figure 2.8a shows a simplified phased-array system model. For a plane wave, the signal arrives at each antenna element with a progressive time delay \( \Delta t \) as in Fig. 2.8b.
This delay difference between two adjacent elements is related to their distance $d$ and the signal angle of incidence $\theta$ by

$$\Delta t = \frac{d \cdot \sin \theta}{c}$$  \hspace{1cm} (2.28)

where $c$ is the speed of light. While an ideal delay can compensate the arrival time differences at all frequencies, in narrow-band applications it can be approximated via other means. For a narrow-band signal, the amplitude and phase change slowly relatively to the carrier frequency. Therefore, we only need to compensate for the progressive phase difference

$$\phi = \beta \cdot \Delta S = \frac{2\pi}{\lambda} \cdot d \sin \theta$$  \hspace{1cm} (2.29)

Where $\phi$ is the electric phase difference between two adjacent channels; $\beta$ is the phase constant; $\Delta S$ is the distance difference for adjacent channel in the wave propagation direction; $\lambda = c/f$ is the wavelength in the air. Assume that $d = \lambda/2$,

$$\begin{cases} \phi = \pi \cdot \sin \theta \\ \Delta t = \frac{\sin \theta}{2f} \end{cases}$$  \hspace{1cm} (2.30)

For example, the incoming angle of 7.2° corresponds to an electrical phase shift of 22.5°. From the above equation, we can also find the relation between $\phi$ and $\Delta t$ as

$$\phi = 2\pi \cdot \frac{\Delta t}{T_s}$$  \hspace{1cm} (2.31)

where $T_s$ is the period of the propagation wave. Figure 2.8b shows the relation between time and phase.

In a receiver chain, for a given modulation scheme, a maximum acceptable bit error rate ($BER$) translates to a minimum signal-to-noise ratio ($SNR$) at the baseband output of the receiver (input of the demodulator). For a given receiver sensitivity, the output $SNR$ sets an upper limit on the noise figure of the receiver. The noise figure ($NF$) is defined as the ratio of the total output noise power to the output noise power caused only by the source, as shown in (2.11), which cannot be directly applied to multiport systems, such as phased-arrays. Consider the n-path phased-array system, shown in Fig. 2.9. $S_{in}$ is the input signal power; $N_{in}$ is the input noise power; $N_1$ and $N_2$ are the 1st and 2nd stage added noise power, respectively; $G_1$ and $G_2$ are the available power gain of the 1st and 2nd stage,
respectively; $k$ is the antenna number; $K$ is the number of antennas; $\Theta$ is the phase difference between two adjacent channels to compensate the phase difference introduced by angle of incidence $\theta$. We assume here that the noise power $N_{in}$ and $N_1$ are equal for all channels.

Since the input signals are added coherently, taking into account the weighting factor for each channel when combiners are implemented in analog domain [48], then

$$S_{out} = KG_1G_2S_{in}$$  \hspace{1cm} (2.32)

The antenna’s noise contribution is primarily determined by the temperature of the object(s) at which it is pointed. When antenna noise sources are uncorrelated, such as in indoor environment, and also the front-end noise sources are uncorrelated, the output total noise power is given by

$$N_{out} = (N_{in} + N_1) \cdot G_1G_2 + N_2G_2$$  \hspace{1cm} (2.33)

Thus, compared to the output $SNR$ of a single-path receiver, the output $SNR$ of the array is improved by a factor between $K$ and $K^2$, depending on the noise and gain contribution of different stages. The array noise factor can be expressed as

$$F = \frac{K(N_{in} + N_1)G_1G_2 + N_2G_2}{KN_{in}G_1G_2}$$

$$= K \frac{SNR_{in}}{SNR_{out}}$$  \hspace{1cm} (2.34)
which shows that the SNR at the phased-array output can be even larger than the SNR at the input if $K > F$. For a given $NF$, an n-array receiver improves the sensitivity by $10\log(K)$ in decibels compared to a single-path receiver. For instance, an 8 element phased-array can improve the receiver sensitivity by 9 dB.

Phased-array can enhance the receiving signal power, as shown in Fig. 2.10. Assume each antenna of a phased-array receives $P_0$ power form the main beam direction. After phase shift and combining, assuming no loss in between, the combined power in the main beam direction is $P_0 + 20\log(K)$.

An additional advantage of a phased-array is its ability to significantly attenuate the incident interference power from other directions. In a single-chain receiver, the linearity performance reflects on the third order input intercept point ($IIP_3$). It is in many cases dominated by the interferer instead of the desired signal. A phased-array receiver has the advantage of enhancing the desired signal by adding the path signals in-phase, and reject the unwanted interferer (from another angle) by adding the path signals out-of-phase. This can be expressed as

$$s_{SUM} = \sum_{k=1}^{K} A(t) \cdot e^{j2\pi f_c t} \cdot e^{j(k-1)\varphi} \cdot e^{-j(k-1)\gamma}$$

(2.35)

Where $s_{SUM}$ is the signal at the output; $A(t)$ is the amplitude of the incoming signal and $f_c$ is the carrier frequency; $\varphi$ is the input signal electric phase difference (can be either desired or unwanted signal), and $\gamma$ is the electric phase compensation (for desired signal) on each path and $\gamma = \pi \times \sin \theta$; $k$ is the antenna number; $K$ is the number of antennas. Furthermore, assuming antenna spacing $d = \lambda/2$ ($\lambda$ is the signal wavelength), the space angle $\theta(deg)$ can be transferred to a phase difference by

$$\varphi = \frac{2\pi}{\lambda} \cdot d \cdot \sin \theta = \pi \cdot \sin \theta$$

(2.36)
Combining (2.35) and (2.36), and taking only the absolute amplitude of $s_{SUM}$, the normalized array gain, $A_{SUM}$, can be expressed as (for normalized signal amplitude, $A(t) = 1 \, V$)

$$A_{SUM} = \left| \sum_{k=1}^{K} e^{j(k-1)\pi \sin \theta} \cdot e^{-j(k-1)\pi \sin \phi} \right|$$

(2.37)

When $K = 1$, it is a single antenna receiver without any directivity. Hence, the array gain is unity for all angles of incidence. When $K \neq 1$, multiple antennas produce antenna patterns which are a function of $K$, desired viewing angle $\theta_d$, and un-desired viewing angle $\theta_i$. Assuming $\theta_d = 0^\circ$, adjusting $\phi$ to the desired signal results $\phi = 0^\circ$. $A_{SUM}$ can be expressed in (2.38), and plotted in Fig. 2.11 with $K = 1, 2, 4, 8$ as examples.

$$A_{SUM,\,dB} = 20 \log \left| \sum_{k=1}^{K} e^{j(k-1)\pi \sin \theta} \right|$$

(2.38)

Here, we define a suppression factor $L$ that describes the power rejection for $\theta_i$ relatively to the power at $\theta_d$

$$L = f(n, \, \theta_i, \, \theta_d)$$

(2.39)

For example, assume $K = 4$ and $\theta_i = 35^\circ$ as shown in Fig. 2.11, the suppression from the peak ($K = 4$) is $12 \, dB - (-5 \, dB) = 17 \, dB$, hence $L = f(n = 4, \theta_i=35^\circ, \theta_d=0^\circ) = -17 \, dB$ (note that $L$ in terms of dB is always a negative number, corresponding to a power loss or power gain smaller than one).