# UNIT 2 SIGNALS AND SIGNAL CONDITIONING

## Structure



## 2.1 INTRODUCTION

In our day-to-day life we use communication for transfer of information from one point to another. The information is sent in the form of some signal wave. You know that there are various types of signals present in the nature, but in modern communication, we deal with mainly audio, video or computer based information. On the basis of their origin and nature, the signals can be classified in various classes. In this unit you will learn about the classification of the signals.

The advent of digital technology has brought in a great improvement in the modern communication systems. For using digital circuits, the signal has to be in the digital form. This can be achieved by sampling the analog signal at appropriate time intervals and then digitising these samples. The digital signals have better noise immunity than the analog signals. By special techniques, the digital signals can be coded to provide a secure data transmission. In this unit we shall discuss the basics of digital signal and its processing.

The signal is carried by a carrier wave or *channel*. Each signal requires a particular bandwidth in this channel. This characteristic of signals is very important in deciding the quantity of signal (information) carried by any channel. The *information theory* is used for quantifying the information handled by the communication system. The power of signal decides the distance or range of transmission in which a signal is received with tolerable attenuation/distortion. So while designing any communication system, we must take into account all these characteristics associated with the signal.

Noise in the communication system is the most important parameter affecting the quality of the signal. There are various causes of noise like inherent noise in the source, noise generated internally by signal handling circuits, external noise picked up by the signal during transmission through medium, etc. Lot of research is going on in developing techniques for reduction of noise. You will be learning some aspects of noise reduction techniques in this unit.

In Sec. 2.2 we shall discuss different classes of signals in communication. Various waveforms used in the communication systems are detailed in Sec. 2.3. Bandwidth is a very important parameter of communication system, which decides the information content of the signal. In Sec. 2.4 we will discuss about the bandwidths of analog as well as digital systems. In Sec. 2.5 you will learn about the Information Theory in brief. Different sources of noise are detailed in Sec. 2.6. In Sec. 2.7, some common methods used for removing or reducing the noise are described.

### Objectives

After learning this unit, you should be able to:

- list various classes of signals;
- state different types of waveforms used in communication;
- define a digital signal and understand its characteristics;
- describe the method of converting analog signals to digital signals;
- explain the coding of digital signals;
- explain the concept of bandwidth for analog and digital signals;
- describe the elements of information theory;
- state causes of noise in communication system; and
- explain hardware and software methods of improving signal to noise ratio in communication systems.

## 2.2 CLASSIFICATION OF SIGNALS

Signal means basically some *information*. Normally we are interested in *time-variation* of information. That is, signal $x$ is a function of time. For example it can be output voltage of an amplifier as a function of time $V_0(t)$. But in general, information may not vary only with time. For example, the photograph on your Identity Card is really the *space variation* of dark and white shades (a function like Shades(space) instead of $V_0(t)$). In this unit we will represent all the signals as $x(t)$, as this is more suitable for modern electronic communication; e.g. when we see a photograph on television, the scanning gives rise to a function Shade($t$).

The signals can be characterised by various parameters. Let us now classify the signals on the basis of these parameters.

a. **Continuous time and discrete time signals**: The signal can be a continuously varying value (amplitude) as a function of time. For example, the intensity of sunlight received at a particular place varies continuously or the value of stock market index keeps on fluctuating during the entire period of business or the speech signal produced by us is also a continuously varying signal.

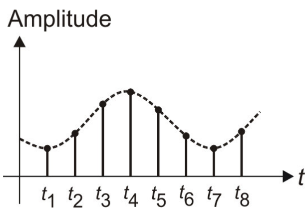


**Fig. 2.1: Discrete Signal**

But in the electronics systems using digital components, we need the signal in **discrete form** which has finite value at certain discrete instances of time and otherwise it has zero amplitude as shown in Fig. 2.1.

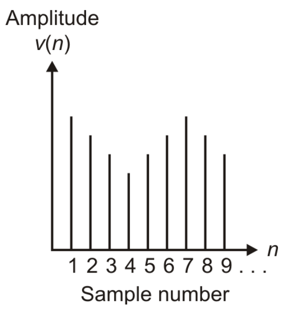


**Fig.2.2: Sampled signal**

Even if the signal is continuous, we may discretise it for our convenience. For example, sunlight is continuous but agriculture department needs *daily average* sunlight received or the stock market index varies continuously during the day, but newspaper reports only the *end of day* index. In both these cases, when we report a discrete value as a sample of the variable value, we carry out discretisation process.

In case of sampling, the sampled signal is a discrete-time signal. To distinguish it from corresponding analogue signal $v(t)$, let us indicate it by $v(n)$, $n$ being integer number (1,2,3,……) indicating sample number (see Fig. 2.2). Note that $t$ is measured in the units of time (s) but $n$ is a count with no units.

**b. Coded and uncoded signals:** If communication were only between human beings and fact-to-face, the coding of signal was not necessary. But when machines were employed for communication coding became necessary. Generally, converting message into some other suitable form is known as **Coding**. In fact, when we created numeric or alphabetic characters, we actually used the coding concept.

Some of the earliest forms of electrical communication used coding to send messages rather than transmitting voice directly. There are many different codes available for use. Most common codes use two signal levels and therefore we refer to them as a binary system.

Coding is the process of transforming messages or signal in accordance with definite rules. The code in the binary form can be represented by two levels *high* and *low*, or '1' and '0', or *mark* and *space* as shown in Fig. 2.3. The signal is in the form of a pulse train, with each pulse having same time period.

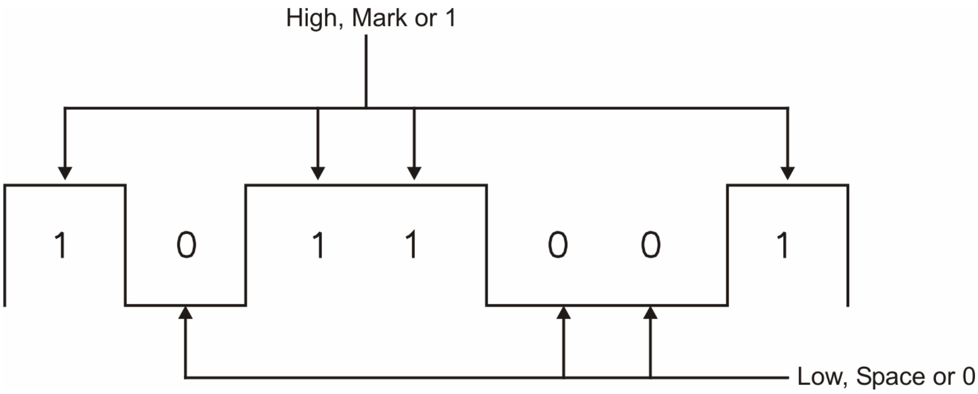


**Fig. 2.3: Binary levels**

The pulses are called *bits*. Fig. 2.3 represents a seven bit code for an event having a sequence 1011001. Since $2^7 = 128$, these 7 bits could be arranged in 128 ways. The number of bits ($M$) and total number of events ($N$), which could be represented by these number of bits are related through the equation:

$$M = \log_2 N \tag{2.1}$$

There are many advantages of using binary coding. These systems are more efficient than other coding systems and have greater noise immunity.
Let us define the efficiency ($\eta$) of a coding system. It is given by:

$$\eta = \frac{\text{number of bits required to represent the total events}}{\text{minimum necessary integer number of bits}} \times 100\% \tag{2.2}$$

If we want to represent 26 alphabets in English language, the number of bits required to represent any one out of these 26 events are

$$M = \log_2 26 = 4.7 \tag{2.3}$$

But the number of bits required must always be an integer. The nearest larger integer is 5. Thus, 5 bits are required. Then the efficiency is

$$\eta = \frac{4.7}{5} \times 100\% = 94\% \tag{2.4}$$

For comparison, consider that the other system of coding is decimal having 10 different levels. Each level is known as **dit** (decimal digit). Now if we are to represent 26 alphabets in decimal system we require

$$M = \log_{10}26 = 1.415 \text{ dits} \tag{2.5}$$

The nearest larger integer number of dits are 2. Then the efficiency is

$$\eta = \frac{1.415}{2} \times 100\% = 71\% \tag{2.6}$$

Comparing the efficiencies in these two coding systems we can see that the binary system of coding is more efficient than the decimal system and therefore it is used almost exclusively.

We can justify the greater noise immunity of binary system by again comparing the binary system with the decimal system of coding.

Consider a binary system where zero volts represent '0' state and 9 volts represent '1'. If the noise level is less than 5 volts, it is considered as '0' level, i.e. no noise, and when the noise level is more than 5 volts, noise is said to be present and denoted by '1' level and may introduce error in the signal.

In a decimal coded system, output '0' is represented by 0V, '1' by 1V, '2' by 2V… '9'by 9V. Here, the 10 discrete levels are 1 volt apart and a noise level of 0.5V can introduce error.

Let us now specify the main difference between the coded and uncoded signals. AM, FM, PM are called uncoded signal systems because there is a direct one to one correspondence between the instantaneous signal amplitude and the corresponding modulated signal form. However, for a coded signal, there is a constant carrier signal amplitude and frequency irrespective of the amplitude of the input signal. You will be learning the details of signal coding in Unit 6.

*Spend 2 Min.*

**SAQ 1**

Name any two codes that you have come across.

**c. Dimensionality based classification**: In general, the signal can vary not only with time, but with many other independent variables. For the signals handled by electronic communication system in addition to time, there may be three space dimensions ($x$, $y$, $z$ or $r$, $\theta$, $\phi$ or $r$, $\phi$, $z$). The sound is a function of only time where space is not important; except for the stereophonic system. In general, the sound is one dimensional (1-D) signal, photograph is a 2-D signal ($x$, $y$) and dynamic picture (TV) is a 3-D signal ($x$, $y$, $t$). As you have already learnt in the last unit, the sound is also classified as a serial signal, and the picture as a parallel signal.

**d. Periodic and aperiodic signal**: The signal that repeats itself after a fixed period of time is called a **periodic signal**. When a signal does not have a repetitive pattern, it is **aperiodic signal**. In Fig. 2.4a you see a periodic signal, which repeats itself with time period $T_0$. This signal ideally starts at $t = -\infty$ and continues up to $t = +\infty$ in repetitive fashion. The signal shown in Fig. 2.4b is aperiodic. It is random in nature and may continue from $t = -\infty$ to $\infty$. The vowels, while singing, are quite periodic but the consonants are aperiodic.

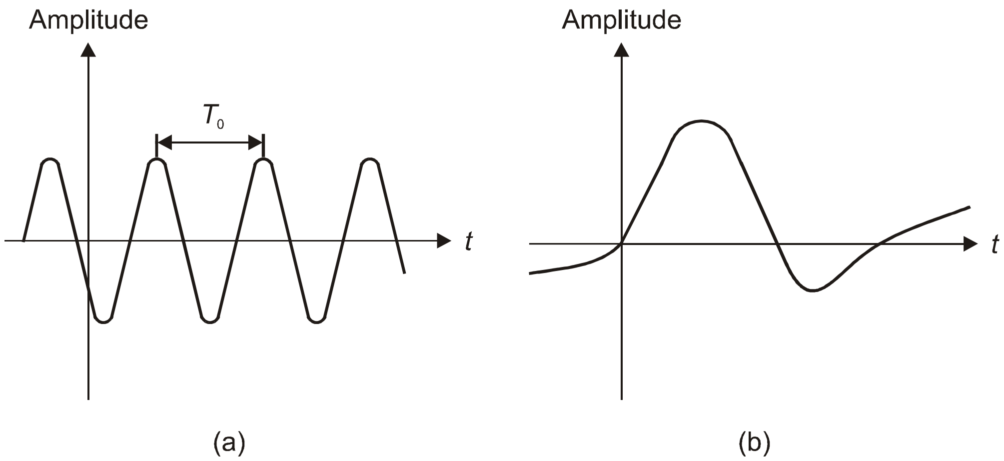


**Fig. 2.4: a) Periodic; and b) Aperiodic signal**

e. **Energy and power signals**: In the discussion so far we have considered the electrical signals as voltage variation with time, i.e. $v(t)$. The radiated signal can also be expressed in terms of field $E$, in the unit of V m$^{-1}$. The most standard model equations for these signals are:

$$v(t) = \hat{v}\sin(\omega t + \phi) \text{ or } E(t) = \hat{E}\sin(\omega t + \phi) \text{ for oscillation}$$

and $$E(t) = \hat{E}\sin(\omega(t - r/c) + \phi) \text{ for wave propagation}$$

Other than voltage or field, the signals can be expressed in terms of power, $P$. For example, output of an oscillator or an amplifier or energy associated with radiated wave are expressed in terms of power.

You know that in a circuit, power is expressed as:

$$P = I\,V = V^2/R = I^2\,R. \qquad (2.7)$$

This means that the power signal is a square of the amplitude signal. For radiation, the intensity or power density is expressed by

$$\mathbf{I} = \mathbf{E} \times \mathbf{H} \qquad (2.8)$$

To quantify the strength of the signal in communication, it is not enough to know its amplitude, but we need to consider its duration as well. The most common way of quantification is to integrate the signal amplitude over the time period from $-\infty$ to $+\infty$. However, in case of a signal having positive and negative amplitudes, it is possible that the areas under the positive half and negative half cancel each other and the signal strength results into zero value. To avoid this, it is a practice to integrate the square of amplitude. The following equation gives the **energy of the signal**:

$$E_f = \int_{-\infty}^{\infty} f^2(t)\,dt. \qquad (2.9)$$

If $f(t)$ is a voltage signal, the energy ($E$) dissipated in resistance $R$ *is*

$$E = \int_{-\infty}^{\infty} \frac{V^2(t)}{R}\,dt = \frac{E_f}{R}$$

For $R = 1\Omega$, the energy $E$ dissipated in $R$ is $E_f$.

If the signal is expressed as a complex quantity, then this equation is modified as:

$$E_f = \int_{-\infty}^{\infty} f(t) f^*(t)\,dt. \qquad (2.10)$$

The signal with finite energy is called an **energy signal**, i.e. it is a signal with

$$0 \le \int_{-\infty}^{\infty} f^2(t)\,dt < \infty.$$

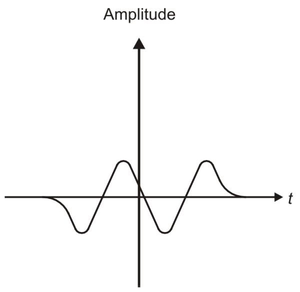


**Fig. 2.5: Energy Signal**

In order to be a meaningful energy signal, the energy of the signal should be finite. This is possible only when the amplitude of signal $f(t) \to 0$ with $t \to \infty$ as shown in Fig. 2.5. However, when the signal is of periodic nature, extending from $t = -\infty$ to $+\infty$, its energy is not finite. In such cases, a time averaging of the energy over the period of the signal ($T_0$ in Fig. 2.4a) is done. This quantity is called **power of the signal**. It is defined as

$$P_f = \lim_{T_0 \to \infty} \frac{1}{T_0} \int_{T_0/2}^{T_0/2} |\,f(t)\,|^2\,dt \qquad (2.11)$$

For a complex signal,

$$P_f = \lim_{T_0 \to \infty} \frac{1}{T_0} \int_{T_0/2}^{T_0/2} f(t)\,f^*(t)\;dt \qquad (2.12)$$

The energy is expressed in the units of Joule; while power is expressed as Watt, which is Joule per second. Hence Eq. (2.9) and (2.11) are dimensionally matching.

You must have noticed that these expressions are the average of the square of signal amplitude. If we take square root of this quantity, we obtain the **root-mean-square (rms)** value of the signal. When the signal is periodic or having statistical regularity and results into finite power ($P_f < \infty$), such signals are called **power signals**.

Hence the conditions for the signal to be energy signal and power signal are:

$$0 \le \int_{-\infty}^{\infty} |\,f(t)\,|^2\,dt < \infty \qquad (2.13)$$

and

$$0 \le \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} |f(t)|^2\,dt < \infty \qquad (2.14)$$

respectively.

Using the classification discussed above you may like to attempt an SAQ.

*Spend 2 Min.*

**SAQ 2**

Classify the signals, shown in Fig. 2.6(a-c) as energy signal, power signal or none.

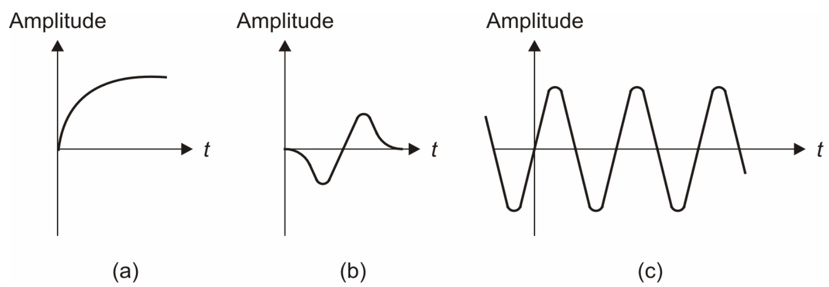


**Fig. 2.6**

**f. Deterministic and random signal**: When the signal is of regular nature, its value can be predicted as the function of time. Such signal is called **deterministic signal**. When the signal cannot be determined or predicted in advance, it is a **random signal**. You will appreciate that the signal we are interested in communications is of random nature, because if the signal is predictable, it is not necessary to send it to the receiver at all.

We must mention here that the above classifications are not exhaustive, but are surely important from the point at view of both Physics and Communication Systems. This classification indirectly gives the parameters/qualities of the communication systems.

For characterising communication systems, standard waveforms are used as test signals. In the following section we shall discuss the important waveforms encountered in Electronic Communication.

## 2.3 WAVEFORMS IN COMMUNICATION

Any waveform used in communication circuits may be viewed as a subclass of only two waveforms, viz. sine and square.

**a. Transient and pulse waveforms:** In practice, any piece of information signal is a random short-time variation of voltage or *transient*. In Physics, we handle such apparently chaotic situations by modelling them in the form of a simple function. For a transient signal occurring at time $T_0$, the simplest model is a single narrow pulse represented mathematically by a δ function given by

$$V(\text{t}) = \infty, \text{ at } t = T_0 \qquad (2.15a)$$

and $$V(\text{t}) = 0, \text{ for } t \neq T_0 \qquad (2.15b)$$

with a condition that

$$\int_{-\infty}^{\infty} V(t)\,dt = (\text{constant})\ \text{V s}. \qquad (2.15c)$$

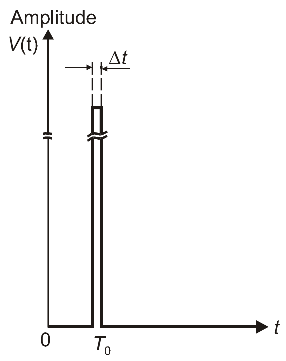


**Fig. 2.7: Transient pulse waveform**

For a unit δ-function, this constant is equal to 1. In real signals as shown in Fig. 2.7, $V$(t) ≠ ∞, but very large and pulse duration $\Delta t$ is very small but never a zero. However, condition (2.15c) will be valid.

A pulse signal has finite width, in contrast to the narrow transient pulse. It is characterised by its duration and amplitude. In Fig. 2.8 these important characteristics are depicted. The pulse signal always has a constant amplitude. For example in case of digital electronics, the TTL signal is defined by 0V for '0' level and 5V for '1' level. The *pulse duration* is defined by the time interval at 50% amplitude. Ideally the pulse should have zero rise and fall times. However, in reality, they have finite values and are defined in the following ways:

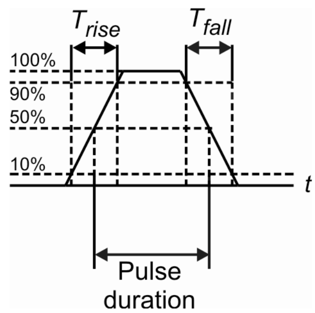


**Fig. 2.8: Pulse signal**

- The *rise time* ($T_{rise}$) is the time required for the signal to reach from 10% to 90% of its amplitude.
- The *fall time* ($T_{fall}$)is the time required for the signal to fall from 90% to 10% of its amplitude.

Normally the rise and fall times are less than 10% of the pulse duration time.

For testing the performance of an electronic system under fast varying input, we give a model transient pulse at the input of the system and observe its output.

Usually, the output signal will also be a transient. For example, voltage pulse given to an *LCR* circuit will result into a transient at the output, which will rise and fall depending on the values of the circuit components. For an ideal *LC* circuit, the output will not be a transient, but a sinusoidal oscillation.

Hence, we can give δ-function like single pulse input to a system and find its transient output response, which will be dependent on the system parameters.

**b. Repeating transient waveform:** A practical problem encountered with single pulse transient is that corresponding output cannot be studied easily. There are two solutions: first is to use a storage CRO which stores the output once and displays it repetitively. But this CRO is not available widely. The second, more practical, solution is to use a properly chosen repetitive pulse as input. This is also called rectangular signal wave. The condition on this repetitive signal is that (i) the pulse duration ($\Delta t$) should be small enough with respect to the rise-time constant of system and (ii) the repetition period ($T_r$) must be large enough with respect to system recovery or fall-time constants as shown in Fig. 2.9.

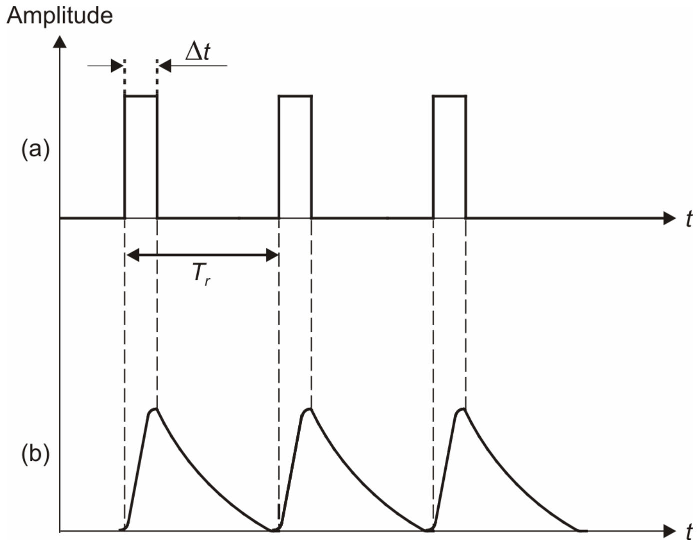


**Fig. 2.9: Repetitive pulse signal: a) input to system and b) output of system**

**c. Step and square wave signal:** One more standard way to study transient response of a system is to apply a step input shown in Fig. 2.10 to the system. It is defined by

$$V(t) = 0 \quad \text{for } t < T_0 \tag{2.16a}$$

$$V(t) = 1 \quad \text{for } t \geq T_0 \tag{2.16b}$$

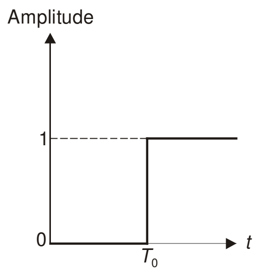


**Fig. 2.10: Step input**

The step input is one of the standard signals used for testing the quick time-domain response of circuits and systems. To get a time-dependent variation in the response, we must have minimum one reactive element in the system. In practice there are two obvious limitations on the use of step function input, viz. (i) finite rise time and (ii) the unity amplitude of the signal after $t = T_0$ cannot remain constant for infinite time. Therefore, we have to use a square wave (Refer Fig. 2.11a) where the on time ($T_{ON}$) and off time ($T_{OFF}$) are equal. The condition on this square wave is that it should have a very small rise time with respect to expected system response, i.e. rise time. The duration of square wave pulse should be sufficiently large with respect to recovery (decay or fall) time of the system depicted in Fig. 2.11b. That is, the limits on rise and fall time of this waveform are governed by the same conditions, specified for the pulse signal.

Now you may like to attempt one numerical example.

**SAQ 3**

*Spend 4 Min.*

What are the maximum allowed values of rise and fall time for a 20 kHz square wave signal?

**d. Ramp and saw-tooth/Triangular waveform:** One more input signal used for circuit performance testing is a ramps input. The corresponding practical (repetitive) input signals are saw-tooth and triangular waves shown in Fig. 2.11c and d. The saw-tooth wave has very short fall time, whereas triangular wave has equal rise and fall times. These waveforms are used for finding the decay-time of the system.

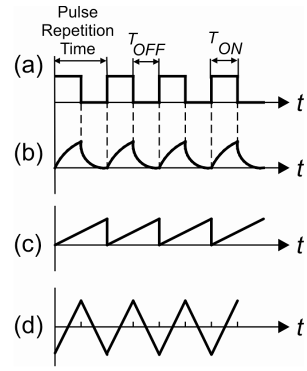


**Fig. 2.11: a) Square wave input; b) System response; c) Saw tooth; and d) Triangular waveform**

**e. Sinusoidal signal (cos or sin):** All the standard signals discussed till now are useful for time domain response and analysis of the system. For frequency domain response the standard input is of (sin $\omega t$) or (cos $\omega t$) form. For mathematical convenience, these signals are usually written as $e^{j\omega t}$. Naturally, the exponential term includes both sin and cos components as:

$$e^{j\omega t} = \cos \omega t + j \sin \omega t \tag{2.17}$$

We will consider here only the real part of the expression. For testing with sinusoidal waveform, a single frequency is given at the input, and the output response of the system is recorded. If the system has no nonlinearity like saturation, this should give the same frequency at the output; the amplitude may differ, as well as there could be a phase change for output with respect to input. The same process is then repeated at many other frequencies over the band of frequencies of interest, *keeping the input amplitude constant*, and the outputs are recorded.

Though $f = \omega/2\pi$. For convenience $f$ and $\omega$ are used synonymously.

Now, the output amplitude is plotted as a function of frequency. This procedure gives the **frequency response** of the system as shown in Fig. 2.12. Obviously this is a time consuming process. Therefore, many times, a quick time domain analysis using the step/square input is done. From a step input response, the frequency response of the system can be obtained by using Fourier transforms.

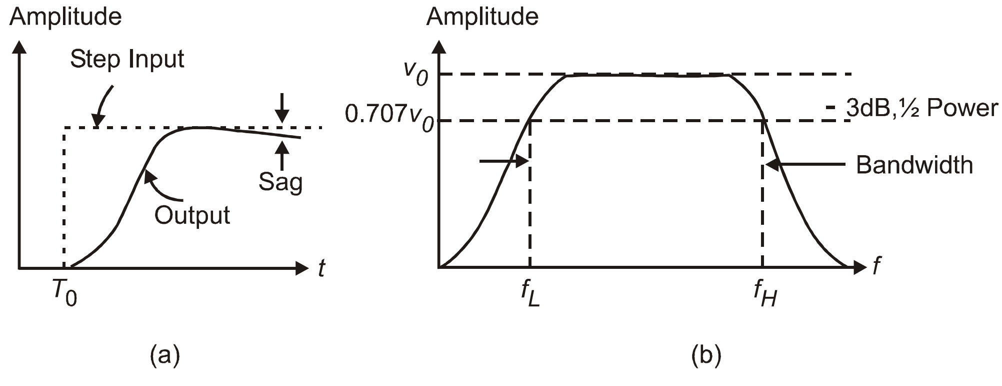


**Fig. 2.12: a) Time domain response of system using step input; and b) frequency response of system using sine wave input**

Note that all these responses are related to reactive elements like capacitance and inductance giving filter action. For example, in *LCR* circuit, $f_H$ is related to shunt capacitance or series inductance, and $f_L$ is related to series capacitance or shunt inductance. Here, shunt and series are with reference to the output load.

After taking a brief review of different waveforms used in communication, let us now discuss about *bandwidth*, which is an important parameter determining the performance of a communication system.

## 2.4 BANDWIDTH

The most crucial parameter in communication systems in this age of commercialisation is the signal bandwidth: the range of frequency in which the signal is varying. The subscribers of communication system are charged on the basis of quantity of information they use. Since in digital (computer) communication the information content is directly proportional to the bandwidth of the signal, it is used as a measure to quantify the utility of the system by the subscriber and the billing is done accordingly.

**Table 2.1: Typical audio bandwidths**

| Source | Frequency | Band Width |
|---|---|---|
| Violin | 196-2794Hz | ~2.6 kHz |
| Piano | 27-4186Hz | ~4.2 kHz |
| Guitar | 82-880Hz | ~0.8 kHz |
| **Human speech:** | | |
| Vowels (a,e,i,o,u,y) | 250-500Hz | ~ 4.0 kHz |
| Consonants (b,c,d,...) | 2-4kHz | |

Bandwidth has a different meaning in analog and digital signals. Let us discuss the different aspects of bandwidth with reference to the signals in communication.

### 2.4.1 Analog Bandwidth

Analog bandwidth is defined in terms of frequency in units of cycles per second, i.e. Hertz (Hz). The audio frequency range is 20 kHz. However, typically, the telephone signal is delivered in the range of 200 Hz to 3200 Hz. Hence, the bandwidth of a typical telephone signal is 3 kHz. The frequency ranges of some audio signals and their bandwidths are given in Table 2.1. From this table you can see that human speech has bandwidth of nearly 4 kHz. Since telephone bandwidth is limited to only 3 kHz, sometimes we have a problem in distinguishing consonants like *b*, *d*, *t*, *v* on the telephone. In AM radio transmission the signal bandwidth is about 10 kHz, while in FM transmission it is 15 kHz. The quality of audio signal received from FM broadcast is much better than AM, since larger range of audio frequency is covered in FM and hence there is less distortion of the signal. The compact discs have bandwidth of 20kHz, hence their sound is of better quality than FM.

Another important parameter is the **dynamic range** of the system. It is defined as the difference between highest possible peak signal and the lowest detectable signal. Wider the dynamic range, better is the quality of the signal. Dynamic range for AF signals is expressed in the units of dB (decibels).

In case of audio systems, the bandwidth and dynamic range are closely related. Hence, normally, it is preferred to have large audio bandwidth and dynamic range.

There are some exceptions to these criteria. One example is the radar or sonar. Here, a pulse of very narrow frequency band is sent as a signal to the target and the same frequency signal is received back. However, since the signal gets heavily attenuated for the targets at a longer distance, the amplitude of received signals may be very small. Hence, a wide dynamic range is required to detect very faint echoes, though the frequency range is very narrow.

Another exception is found in Astronomy and Astrophysics. Here, distant stars are detected by catching a signal of wide frequency range, but with very weak signal intensity embedded in large noise. Hence, it is in effect a lower dynamic range and a large bandwidth application.

The bandwidth in case of video signals can be expressed by different parameters like bandwidth of video channel, bandwidth of video signal and number of lines of resolution. A typical television broadcast channel has bandwidth of 6 MHz, which includes the audio and video signal and separation frequency band between two channels. The bandwidth of just video signal is typically 4.2 MHz. Commercially, the video bandwidth is expressed in terms of a number of horizontal lines used in picture rastering (or scanning). This parameter defines the resolution of the video signal. For sending more lines of horizontal scan, more bandwidth is needed and it results into better quality of the picture.



### 2.4.2 Digital Bandwidth

The digital bandwidth gives the quantity of information contained in a digital signal. This is different than the analog bandwidth, which measures the amount (or range) of spectrum each signal occupies. The digital bandwidth is expressed in terms of bits per second (bps). Naturally, the more the bps, more is the information content of the signal. Normally, the digital signal consists of very large number of bits, so the bandwidth is expressed as kilobits per second (kbps) or megabits per second (Mbps) or gigabits per second (Gbps). Please remember that a kilobit of data in digital communication is equal to 1024 ($2^{10}$) bits and not 1000 bits. Similarly, a megabit corresponds to 1,048,576 ($2^{20}$) bits instead of 1,000,000 bits.

Modem allows communication of digital signals over analog telephone lines.

Nowadays the Internet connectivity can be obtained using the telephone line at our homes by connecting a *modem* between the telephone cable and the computer. You will learn about this mode of communication later in this course. Typical modem connections have bandwidth or data transfer rate of 32kbps or 64 kpbs or 128 kbps.

Just the way bandwidth gives a measure of volume information exchange, the information theory quantifies the information content of any communication. Let us now discuss briefly about the *Information Theory*.

## 2.5 THE INFORMATION THEORY

The amount of information carried by any message is inversely proportional to the probability of that message. Let us understand it with a simple example. Suppose, there are two messages:

1) The sun is going to rise tomorrow morning.
2) The sun will be disappearing tomorrow afternoon at 3 pm.

Out of these two, the first message is a common one and you do not gain much information from it. However, the second message surely interests you, since it is a less probable event.

So lesser the probability ($p$), more is the information value ($I$) of the message. Using this fact, we can quantify the information carried in a message as some function of inverse of the probability of its occurrence.

i.e. $$I \propto f\left(\frac{1}{p}\right) \qquad (2.18)$$

Let us consider $L$ messages $m_1$, $m_2$ … $m_L$ with respective probabilities of their occurrence as $p_1, p_2 ... p_L$. When message $m_k$ is transmitted or received, the amount of information transferred will be:

$$I_k = \left[\log_2 \frac{1}{p_k}\right] \qquad (2.19)$$

Here, we assume that there are only $L$ messages. Hence the probabilities $p_i$ ($i$ = 1,...$L$) will add up to 1.

i.e. $$\sum_{i=1}^{L} p_i = 1. \qquad (2.20)$$

$I_k$ is a dimensionless quantity, but by convention, it is assumed that the message uses binary code and hence $I_k$ is expressed in the unit of bit.

Now, we would like you to attempt an SAQ.

*Spend 2 Min.*

**SAQ 4**

Calculate $I$ for alphabets of English language assuming equal probability of their occurrence.

If $L$ messages are equally probable, probability of each message occurring will be $\frac{1}{L}$.

i.e. $$p_1 = p_2 .... = p_L = \frac{1}{L}.$$

If $L = 2^N$ where $N$ is an integer, then the information content in each message will be

$$I = \log_2\left(\frac{1}{1/L}\right) = \log_2 L = \log_2 2^N = N \text{ bits.} \tag{2.21}$$

For example, eight messages of equal probability of occurrence will carry 3-bit information each.

Assume that the $L$ messages have different probabilities $p_1$, $p_{2,......}$ $p_L$, of occurrence. If the source sends $M$ different messages over a sufficiently long period, so that $M >> 1$; then we can easily presume that, out of $M$ messages, $p_1M$ messages will be message $m_1$, $p_2 M$ will be message $m_2$ and so on. The total information carried by this sequence of $M$ messages will be

$$\begin{aligned} I_{total} &= p_1 M \log_2 \frac{1}{p_1} + p_2 M \log_2 \frac{1}{p_2} + ... + p_L M \log_2 \frac{1}{p_L} \\ &= \sum_{i=1}^{L} p_i M \log_2 \frac{1}{p_i}. \end{aligned} \tag{2.22}$$

The average information per message will be obtained by dividing Eq. (2.22) by total number of messages, $M$. This average is called e**ntropy** of the signal and denoted by $H$. Hence, $H$ can be expressed as:

$$H = \frac{1}{M}\sum_{i=1}^{L} p_i M \log_2 \frac{1}{p_i} = \sum_{i=1}^{L} p_i \log_2 \frac{1}{p_i}. \tag{2.23}$$

For $L$ messages of equal probability

$$H = \sum_{i=1}^{L} \frac{1}{L} \log_2 L = \log_2 L = I\ . \tag{2.24}$$

This is maximum limit of entropy, $H_{max}$.

If there is only one message (i.e. $L = 1$), then the message coveys no information, since that is the only message and Eq. (2.22) results into zero. Similarly, on the other extreme, if $L \to \infty$, then $p_i \to 0$ and again, $H$ is zero. It means that if there are infinitely many messages, then the probability of the occurrence of any of them is negligibly small and the information content approaches to zero.

From Eq. (2.21) we can conclude that, $N$ is the minimum number of bits required to express $2^N$ equally probable message.

Now, the digital channel carrying the data has certain data carrying capacity or **channel capacity**. If you are using a 4-bit system (4 binary pulses per message) and if the sampling frequency $n$ is 8 kHz then the rate at which the information is transferred is (4×8k) bit $s^{-1}$. If this is the maximum rate then it is called the *channel capacity C.*

For an $N$-bit system with sampling frequency $n$, we get

$$C = nN = nH \qquad (2.25)$$

Practically, every data channel is always accompanied by some noise level. We can characterise any channel by its signal-to-noise ratio (SNR).

The pioneer of information theory, C.E. Shenon has shown that, *if the rate of data being generated by the signal source is smaller than the channel capacity, then it is possible to transmit data with negligibly small error, in spite of presence of noise on the channel.*

If $\frac{S}{N}$ is the signal-to-noise ratio for the given data channel then $\sqrt{S+N}$ is total voltage range and $\sqrt{N}$ is the noise-generated step. If $B$ is the bandwidth of the channel then the channel capacity $C$ satisfying following condition ensures that the transmission is error free:

$$C = B \log_2 \left(1 + \frac{S}{N}\right) \qquad (2.26)$$

You will appreciate that, as $B$, the bandwidth increases, we can send more information, since it is possible to make faster changes in the information signal. Also, increase in $S/N$ ratio implies that the noise is less and hence more data can be sent over without any error.

From Eq. (2.26) you may be tempted to conclude that if the bandwidth $B$ increases to infinity, the channel capacity will also become infinity. However, it is not the case. As $B$ increases, the noise associated with the channel also increases and hence $S/N$ decreases, limiting the channel capacity to a finite value.

---

**SAQ 5**

*Spend 2 Min.*

For a bandwidth of 1 Mbps and $\frac{S}{N}$ ratio of 511, calculate the Shannon limit of channel capacity.

---

Now let us address to the most crucial factor affecting the performance of a communication system, viz. *noise*. In the next section we shall take review of different causes of noise in the electronic communication system.

## 2.6 NOISE IN COMMUNICATION

**Noise** is defined as any unwanted signal, which disturbs the communication of useful signal. It can be experienced in the form of a buzz in the telephone or a snow on the television screen. Noise is a random phenomenon and hence, cannot be predicted beforehand. This randomness makes the noise removal very difficult. In a noise, the amplitude, frequency or phase may change randomly. Hence noise level is commonly expressed as root mean square (rms) dc value. Taking rms value takes care of the noise signal taking both positive and negative values. The amplitude of noise signal usually follows *normal* or *Gaussian* statistical distribution. Any of the blocks of communication system like source, detector and the circuits involved can introduce noise into the signal. This is called **internal noise**. There could be some external parameters like atmospheric, ambient conditions that introduce noise during the passage of signal through the medium. This is called **external noise**. There is another cause of external noise in the form of other systems or devices causing disturbance to the functioning. Unintentional electromagnetic signal from a system or device causing undesirable effects or malfunctioning in another system or device is called **Electromagnetic Interference (EMI)**. These undesirable effects can vary between brief annoyance, such as a vacuum cleaner disturbing the television viewing, or to a more serious situation such as a cellular phone interfering with computer controlled instruments.

**Electromagnetic Emission** is the energy radiated to the environment from an electrical or electronic system (or appliance) causing EMI. With the increased use of the electronic technology in our everyday life, electromagnetic interference has become a growing concern. Governments in different countries, through a regulatory authority, decide the amount of acceptable EMI emission produced by any electronic system (or appliance). Electromagnetic interference can be suppressed by optimising the design of electronic circuits and components, by properly shielding connecting cables and by applying housing to the circuits in the system that provides electromagnetic shielding.

Let us now consider some common sources of noise.

If the source is an antenna which feeds to receiver, it is called *antenna temperature*.

**a. Thermal Noise:** The best randomness known in Physics is in the thermal domain. The standard model of thermal emission is the black body radiation. This radiation pattern varies with the source temperature. You know the Planck's law for energy distribution of black body radiation. Two important facts to be remembered in this case are: (i) energy is not same at all the frequencies; and (ii) the radiation intensity changes with frequency as well as temperature. The second fact is very often used to state the strength of the source. If a source radiates energy $E(\lambda_1)$ corresponding to temperature curve $T_1$ shown in Fig. 2.13, then $T_1$ is said to be the **brightness temperature** (i.e. equivalent black body temperature) of source expressed in the units of Kelvin. For the noise source, it is also called the *noise temperature*.

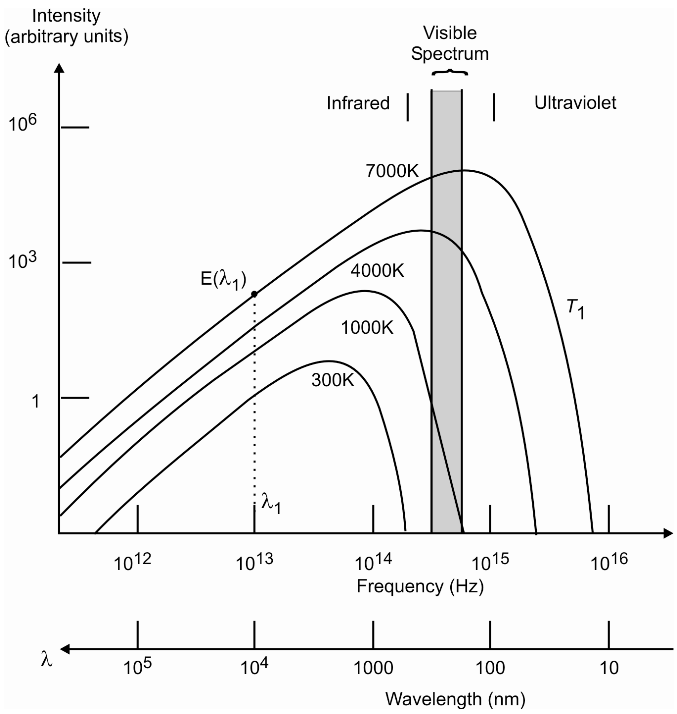


**Fig. 2.13: Black body radiation**

There is another term you may come across: **a white source**. This source is supposed to emit equal energy at all frequencies.

The concept of brightness temperature or equivalent black body temperature is more appropriate for celestial sources. For example, the sun's surface is at ~ 6000 K. This is considered as its black body temperature which corresponds to visible-IR range. But the phenomenon of solar flares gives rise to noise which disturbs RF communication. This effect is cyclic in nature and repeats after every 11 years.

The earth's surface has brightness temperature of 300 K and is a good IR and microwave radiator. Hence, the sun and the earth are the most important thermal noise sources disturbing the communication on earth. As a matter of fact, the whole field of radio astronomy started when it was found that radars are disturbed by some noise source, which finally was found to be the sun's microwave radiation. One more such accidental scientific discovery was of some very regular microwave pulses at sufficiently high black body temperature received from the outer space, as if some extraterrestrial beings are signalling us. They were found to be celestial bodies, initially called *pulsars* and now known as *neutrons stars*.

In an electronic circuit, the electrons in resistor or conductor have kinetic energy due to thermal motion. Due to random velocity distribution of these electrons, there is a small random fluctuation in the potential difference (voltage) across the conductor. This gives rise to the thermal noise in circuit components. This noise is also called a **Johnson noise** named after its discoverer. This noise has an energy spread over a large frequency spectrum. The power generated due to this noise will be dependent on the temperature. It was established by Johnson that the noise power in the frequency range $\Delta\omega_n$ is:

$$P_{nt} = kT\Delta\omega_n \tag{2.27}$$

where $k$ is the Boltzmann constant, $T$ is temperature of the resistor, in Kelvin, and $\Delta\omega_n$ is the frequency range over which the noise is measured. It is called *noise bandwidth*.

From this expression of power, it is possible to calculate the noise voltage $v_{nt}$ generated by the resistor $R$. If we assume that the maximum power is delivered by the noise source, then according to the maximum power transfer theorem, the noise voltage will be equally split between the load and the source itself. Hence, the noise power in Eq. (2.27) can be expressed as

$$P_{nt} = \frac{(v_{nt}/2)^2}{R} = kT\Delta\omega_n \tag{2.27a}$$

Hence, $\dfrac{v_{nt}^2}{4} = kT\Delta\omega_n R$

i.e. $$v_{nt} = \sqrt{4kT\Delta\omega_n R} \tag{2.28}$$

In principle, $\Delta\omega_n$ ranges to infinity and the root mean square noise voltage should be infinite. However, every circuit will have some component of capacitance and inductance (though stray), which would impart a filter effect and limit the noise bandwidth to finite frequency range.

The random thermal motion of electrons rides as a noise on the current flow (drift electrons) in the circuit. Since this noise is dependent on the temperature, cooling of the electronic components (by say liquid nitrogen circulation) reduces it.

Now you will be able to calculate the noise voltage produced by a resistor in the following SAQ.

*Spend 3 Min.*

**SAQ 6**

Calculate the noise voltage produced by a 10 kΩ resistor at 17°C over a 10 MHz bandwidth.

**b. Device Noise:** You know that any charge motion (acceleration and higher derivatives) will give radiation. The devices have charges in motion and any randomness in it will mean corresponding noise. For example, electrons emitted from the filament (or cathode) of vacuum tube devices and CRT have some randomness in their emission. This gives rise to **shot noise**.

In these cases, the charge transport or collection occurs statistically and well-known *Poisson statistics* governs such processes.

If $N$ charges of charge $q$ each drift in time $t$, then the mean current will be

$$I_m = qN/t \tag{2.29}$$

For Poisson distribution, the statistical uncertainty in $N$ events is $\pm\sqrt{N}$ . Hence, the current will have random fluctuation around the mean value. The current will be:

$$I = I_m \pm I_{random} = \frac{qN}{t} \pm \frac{q\sqrt{N}}{t} \tag{2.30}$$

From this equation you can see that the uncertainty in the current ($I_{random}$) is due to shot noise. Let us denote this noise current by $I_{sn}$. To find out the value of $I_{sn}$, we substitute value of $\sqrt{N}$ from Eq. (2.29):

$$I_{sn} = \frac{q\sqrt{N}}{t} = \frac{q\sqrt{\frac{I_m t}{q}}}{t} = \sqrt{\frac{I_m q}{t}} \tag{2.31}$$

In this expression, $t$ is the time interval, over which the shot noise is considered. We can replace it with equivalent bandwidth by considering a pulse of this time duration. In such cases, the pulse will be high for half the time $t$ and zero for half the time $t$. Then, the frequency can be expressed as:

$$\Delta\omega = \frac{2\pi}{2t} = \frac{\pi}{t} \tag{2.32}$$

The shot noise current is then

$$I_{sn} = \sqrt{\frac{I_m q \Delta\omega}{\pi}} \tag{2.33}$$

Since the fluctuations in the shot noise current can occur at any frequency, this is a white noise.

Similarly, in transistor there is randomness in injection of electrons from the emitter to the base and then to the collector, which results into a **transistor noise**.

When circuit current gets divided in two or more parts (e.g. in transistor $I_E = I_B + I_c$ or cathode current of vacuum tube device gets divided in plate and different grids), there is some randomness involved at electron level and this gives rise to **partition noise**. This noise can be reduced only by avoiding the current division.

As diode does not have current division, it is preferred rather than a transistor as a non-linear device to get new frequencies.

Very often we come across another type of noise due to the mains pick-up, called **hum**. This is 50Hz in India and 60 Hz in countries like USA. In some cases like generators or in aeroplanes it can be 120Hz and at times their harmonics. A leaky power supply capacitor may give a noise called *motor boating* due to leaky discharge and recharge.

Similarly, there can be radiation noise from any neighbouring local oscillator, which is not shielded properly. If your circuit is not properly shielded and grounded, it may pick-up signal from local radio stations. There can also be **cross talk** from some other information channel, like telephone lines. A sparking in motors or electrical machines will always radiate its characteristic electrical noise.

All these circuit noises are mostly local because they get attenuated in short distance. This is because normally they are magnetostatic in nature or are surface waves. But sometimes there are some long distance effects. For example, in terrestrial TV transmission, sometimes we get a *ghost* image which is a time-shifted additional image. It is caused by multiple reflections of the signal from a nearby building.

*Ghost image* is a commonly observed phenomenon in the television when terrestrial signal is received by antenna placed on the rooftop. The reflections from nearby buildings give rise to double or multiple pictures on the screen.

**c. Quantisation Noise:** Till now we were mainly concerned with serial analog signals. When we convert the analog signal into digital, the signal can take only discrete amplitude values. If the signal is digitised, you will obviously get *stepped* signal as shown in Fig. 2.14. This means that the analog signal value lying between the two discrete steps will be approximated to the nearest step value. This approximation introduces an error called *quantisation error*. When such a signal is converted back into its analog form, the output is not a smooth analog signal of continuous amplitude, but a stepped one. If the step height is $\Delta V_s$ then the error in amplitude will be of the order of $\pm \frac{\Delta V_s}{2}$. If we assume that the signal is equally likely to be at any point within the range of $\Delta V_s$ then the mean square error noise ($V_n$) is

$$V_n^2 = \frac{1}{\Delta V_s} \int_{-\frac{\Delta V_s}{2}}^{\frac{\Delta V_s}{2}} V^2(q)\,dq = \frac{(\Delta V_s)^2}{12} \qquad (2.34)$$

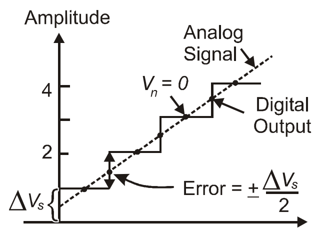


**Fig. 2.14: Quantisation Noise**

This error is effectively a noise introduced by the quantisation process and hence called the **quantisation noise**. It is clear that by reducing the step size $\Delta V_s$ this noise reduces.

---

**SAQ 7**

*Spend 2 Min.*

Calculate the quantisation noise for a digital signal with 10 mV step size.

---

**d. Plasma Generated Noise:** Plasma, as you know, is a collection of electrons and positive ions, often in gaseous ionised form. On the whole, at a macro-scale, the plasma is neutral. In earth's ionosphere, up to *F* layer, the plasma is 0.1% ionised. At micro-scale, the electrons and ions are moving randomly with different rms velocities like in any gas. At intermediate-scale there are local, randomly generated ion clouds due to collection and separation of electrons and ions. The electron cloud, due to attractive forces, starts oscillations through ion cloud at a characteristic plasma frequency. This gives rise to unpolarised, thermal-like radiation which is a white noise.

**e. Fading:** Apart from the noise at transmitter and receiver ends, there are some *on-path* noises. One very important noise-effect of this type is *fading*. On propagation path of RF and microwave signals, the atmospheric conditions may vary due to clouds, rain, air density, moisture etc. Also a variation in ionosphere conditions like ionisation change due to sun rise/set, solar flare, cosmic radiation may occur. This results into attenuation of the signal strength. Such drop in signal during travel is called **fading**.

Sometimes we may get even an enhanced radio wave strength due to *ducts* in the atmosphere, particularly over sea which causes the signal to travel over longer distance. This is due to refractive index inversion caused by the moisture content. This bends the signals like total internal reflection phenomenon observed in the optical fibre. The bent signal experiences multiple reflections betweens earth surface and duct boundary.

In communications, the most important aspect of quality of communication is determined by the noise content in the signal. It is normally expressed as a ratio of signal and noise amplitudes and is termed as **signal to noise ratio (SNR)**.

If $V_a$ is amplitude of signal voltage and $V_n$ is amplitude of noise voltage then SNR is defined as

$$\text{SNR} = \frac{V_a}{V_n} \tag{2.35a}$$

In terms of signal and noise powers, we can define

$$\text{SNR} = \frac{P_a}{P_n} \tag{2.35b}$$

It is usually given in dB.

The larger the SNR, the better is the performance of the communication system.

The SNR defined in Eq. (2.35) identifies the noise content at a specific point in the system. But it is not useful for specifying the noise introduced by any device in the circuit or by any circuit, as a whole, for that matter. To specify this quantity, another parameter called **Noise Figure** is defined as:

$$\text{NF} = 10 \log_{10} \frac{S_i / N_i}{S_o / N_o} \text{ dB.} \tag{2.36}$$

where $S_i / N_i$ is the signal-to-noise power ratio at the device's (or circuit's) input and $S_o/N_o$ is the signal-to-noise power ratio at its output. The ratio $\frac{S_i / N_i}{S_o / N_o}$ is called **noise ratio (NR)**.

You may like to attempt the following SAQ before proceeding further.

*Spend* 3 *Min.*

**SAQ 8**

An amplifier has *S*/*N* power ratio 50 at the input and 20 at output. Calculate NR and NF.

In the next section we shall discuss the methods of improving SNR.

## 2.7 SIGNAL TO NOISE RATIO IMPROVEMENT (SNRI) TECHNIQUES

Lot of research is presently going on in the field of noise reduction techniques for internal and external noises. There are hardware means of noise reduction, where additional circuits, devices or components are used for the noise reduction. Many software tools are also developed for noise reduction by signal processing. Let us now discuss various SNR improvement methods, in brief.

If the noise is very similar to signal, it will be practically impossible to separate the signal from the noise. Hence to improve SNR, we have to first search for the differences in signal and noise; that is, find the difference between *signatures* (or *finger prints*) of the known information signal and the noise present, and then use this difference to remove noise from the signal. You have learnt that any signal has three basic parameters, viz. amplitude, frequency and phase. All these parameters vary with time. For both signal and noise, these parameter variation and their ranges (i.e. their patterns) can be specific, i.e. both have some *specific signature*. The basic principle of SNR improvement is to: *carry out a filtering dependent on the signature difference between signal and noise*.

One situation where we do not have to worry about SNR improvement is when the signal has much higher amplitude than the noise. This is a *brute force* method used in AM broadcast. Here, very high power transmitter is used for signal transmission.

For improving SNR in general when the signal is smaller than the noise, there are two main approaches, i.e. using (i) Hardware and (ii) Software.

### 2.7.1 Hardware Techniques

The noise can be removed only to a limited extent by using hardware means. Some prominent hardware techniques used in SNRI are:

(i) To reduce thermal noise, the main method is to cool the system to lowest possible temperature;

(ii) To reduce the device noise, choose devices like low noise transistors, Low Noise Amplifier (LNA) chip which are specially fabricated. These are run at lower current/voltage ratings for reducing power dissipation in them;

(iii) To avoid partition noise, use diode or devices with least number of internal current divisions;

(iv) To reduce distortion, use the devices (and circuits) in the linear regions of the operating curves as far as possible;

(v) To avoid power supply noises, use highly stable power supplies like battery supply;

(vi) To avoid EMI, pick ups or cross talks use good shielding, grounding etc.; and

(viii) To reduce fading effect, diversity systems are used. For example, operating transmitter and receiver at two different frequencies or placing the receiver antennae at two different elevations can be used for achieving frequency or space diversity. This avoids occurrence of minima in signal due to the interference of the waves (like line-of-sight and ground reflected sky waves).

In general, following strategies are used in the circuit designing for SNR improvement:

**a. Choice of First Stage Low Noise Amplifier:** In the transmitter/receiver amplifier chains, the first stage internal noise is most important. If there are three amplifiers $A_1$, $A_2$ and $A_3$ then the input stage noise gets multiplied 3 times. But if low noise amplifier is used at the first stage, the input signal is amplified by gain of $A_1$, but there is no noise introduction. Hence the signal gets amplified with SNR improvement by a factor of $A_1$.

b. **Low Noise Block:** To avoid any on-path pick-up between actual receiving point (say receiving antenna) and first stage amplifier, the electrical connection between them (transmission line) is kept to lowest length. Therefore, TV booster (amplifier) is mounted on the antenna itself, as a Low Noise Block (LNB).

c. **Negative Feedback:** You know that the negative feedback in amplifier improves all its parameters, at the cost of gain. It increases bandwidth, matches input and output impedances and most importantly, it improves SNR by reducing distortions.

$$\left(\frac{S}{N}\right)_{fb} = \left(\frac{S}{N}\right)_o \times (1 + A\beta) \qquad (2.37)$$

This is normally used in the last power stage of the transmitter and the receiver.

d. **Opto-coupler/Isolation Transformer:** Sometimes noise gets injected from one part of the circuit into another through common ground point. In such a case we need to electrically isolate one part of a circuit from the other part. This can be easily done by a transformer called an isolation transformer or alternatively by use of an opto-coupler. Here the input signal is converted to optical signal by LED; and it is received by a photodiode kept at a short distance and converted back to an electrical signal, thus isolating the input and output electrically, i.e. any grounding related input noise is not transferred to the output.

e. **Clipper for Amplitude Filtering:** For removing high amplitude noise due to the effect of e.m. waves caused by lightning discharge, a biased diode clipper circuit shown in Fig.2.15 is used. The diode does not conduct for normal signal amplitude but when a high voltage appears, the diode conducts (short) and cuts the signal output to maximum limit decided by the reverse bias voltage.

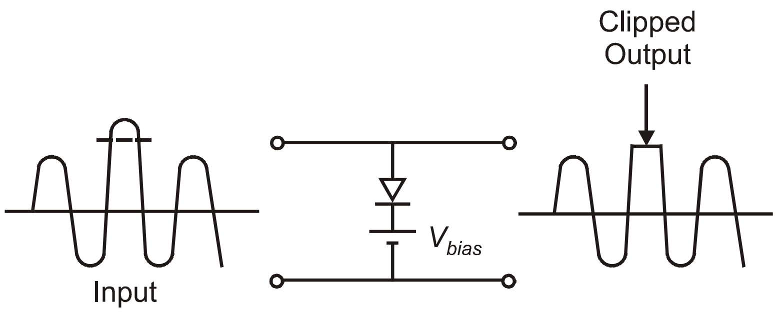


**Fig. 2.15: Biased diode clipper circuit**

A similar principle is used in FM receiver, where information ($\omega_m$) is modulated on carrier frequency ($\omega_c$), as given by Eq. (1.5b). The amplitude of the carrier remains constant in FM. If any noise brings in changes in carrier amplitude they are removed by using a diode clipper. You must have noticed that FM broadcast is better than AM, particularly during the thunder storm. It is because of the clipping of noise amplitude riding on the FM signal.

*Spend* 3 *Min.*

**SAQ 9**

Design a diode clipper circuit to limit the signal amplitude within ± 5V.

f. **Frequency Selective Filtering:** The word *filter* is normally used in electronics with reference to the frequency filtering. It is the most common method used for

SNR improvement. For example, to filter hum due to mains frequency, a band rejection filter is used. It is also used in the special communication circuits like modulators, where non-linearly of the device is used resulting into harmonic frequency generation.

Normally the filter is made of *LCR* circuit. The $\sqrt{LC}$ time constant governs the centre (resonance) frequency and *L/R* and *RC* time constants govern the lower and upper cut off frequencies, i.e. they decide the bandwidth of the filter.

*RC* filters are used along with op-amp to build *active filters* which provide gain. At very high frequencies, particularly at GHz range, the transmission lines themselves are used as resonant circuits. A simple example of this is a λ/2 dipole which is a parallel resonant circuit.

Filtering of frequency is also a filtering of wavelength λ.

Optical filters are used in optical communication systems, if more than one wavelengths are used for signal transmissions.

**g. Integration, Differentiation Circuits:** You have learnt about the circuits used for carrying out integration and differentiation of the input signal. These are actually low-pass and high-pass filters respectively as shown in Fig. 2.16a & b. The outputs of these circuits are given by the following equations:

Integrator:

$$V_o = \frac{1}{RC}\int_0^T V_{in}\,dt \tag{2.38}$$

Differentiator:

$$V_o = RC\frac{dV_{in}}{dt} \tag{2.39}$$

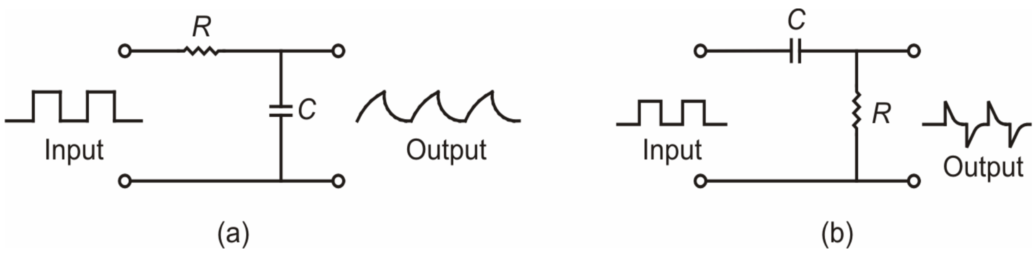


**Fig. 2.16: a) Integrator; and b) Differentiator**

The integration effectively gives the average (dc) of the signal through high frequency noise. The differentiation is used for finding minima and maxima in the signal and particularly to find *hidden small maxima* buried in bigger maxima.

There are various other hardware methods used for noise improvement, like signal chopping, phase sensitive detection, use of comparison circuits like bridges, differential amplifiers, etc. However, since these methods are more commonly used in instrumentation than in communication we will restrict our discussion here. Now we shall go over to take account of software methods used in SNR improvement.

### 2.7.2 Software Techniques

Software methods may be divided into two groups: (i) using purely statistical methods assuming random nature of noise and (ii) using some mathematical treatment on the signal. We will see some examples of these two methods now.

**a. Averaging and Histogram**: the most extensively used basic method for SNRI is the **averaging** method assuming Gaussian (normal) distribution of the noise. Whenever we take measurement of a certain variable, it is very common to call the noise as *an error*. You are told in your practical classes that while taking any measurement, take at least three readings and find their mean. This is for SNR improvement only. As you take more readings, the precision in measuring process increases, i.e. deviation from the mean decreases.

**b. Fourier transforms:** As you know, Fourier transforms convert the time domain signal into frequency domain signal. It means that it resolves the signal in its frequency components. Typically we can express the Fourier transformed signal in the form of frequency spectrum as shown in Fig. 2.17.

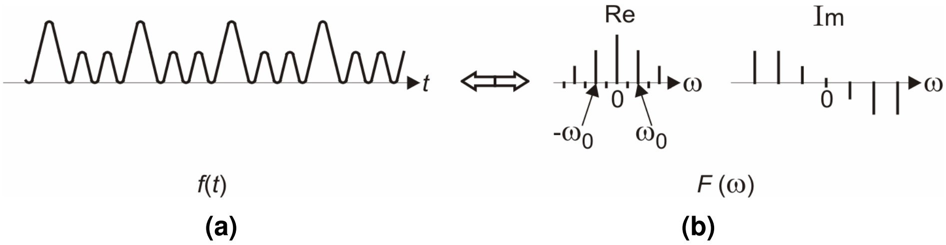


**Fig. 2.17: a) Time; and b) frequency domain signal**

If we know the frequency of noise ($\omega_n$) riding on this signal, then we can remove it by just subtracting the component corresponding to $\omega_n$ from this frequency spectrum. After this, if we calculate the inverse Fourier transform, then the reconstructed signal in the time domain does not contain any contribution of $\omega_n$ component and we get a noise free spectrum.

**c. Digital Signal Coding:** In case of digital data transmission, it is possible that an error occurs in data transmission due to noise interference. This problem can be overcome by processing the digital data. It will be clear from the following examples.

Consider a data transmission of $N$-bit message. It is possible that during the travel of this message from transmitter to receiver, there may be errors in one or more bits of the message, i.e. a bit with value 0 may be received as 1 or vice versa. It is very difficult to find out at the receiver end that error has occurred in the original message.

For example, 8 messages are listed in Table 2.2, and message number 3 (011) is sent by the transmitter. If by the time the message reaches the receiver, there is error in the $2^{nd}$ bit of the message, i.e. it reaches as (001), then the message will be interpreted as message number 1 instead of 3.

**Table 2.2**

| Message Number | Message |
|---|---|
| 1 | 001 |
| 2 | 010 |
| 3 | 011 |
| 4 | 100 |
| 5 | 101 |
| 6 | 110 |
| 7 | 111 |
| 8 | 000 |

To avoid such situations, it is a common practice to add a few more redundant bits in the message. The simplest way is to add a single parity bit whose value is decided by the value of the message bits. If the transmission follows even parity, then the parity bit is equal to 1, if the sum of the value of the message bits is an odd number. If it is even, then parity bit is kept 0.

For example, in the above example, message 3 will have parity bit equal to 0. Hence the transmitted sequence will be 0110. When it is received with error in the second bit, the received message reads as 0010. Since the total of the values of the received message does not add up into an even number, we can detect that an error has occurred during the message transmission.

This procedure of adding redundant bit is called **coding**.

Again, there are many more software ways of noise reduction. However, we will restrict our discussion here.

Let us now summarize the points you have learnt in this unit.

## 2.8 SUMMARY

- Signal can be classified on the basis of various criteria, like continuous and discrete time signals, coding method, periodicity, etc.
- Typically step input and transient input signals are used to check the time dependent response of the system. In practice, these signals are applied in repetitive mode, in the form of square wave and pulse train signals.
- Sinusoidal signals are used to check the frequency response of the system.
- Analog bandwidth refers to range of frequency of base band signal.
- Digital bandwidth refers to rate of data transfer.
- Information content of a message is inversely proportional to the probability of its occurrence. It is defined as

$$I_k = \log_2 \frac{1}{p_k}$$

- Entropy of a signal consisting of $L$ messages is

$$H = \sum_{i=1}^{L} p_i \log_2 \frac{1}{p_i}$$

- Shannon's limit of channel capacity is:

$$C = B \log_2 \left(1 + \frac{S}{N}\right)$$ where $B$ is the channel bandwidth.

- Noise can be generated within the communication system (internal) or get incorporated in the signal during its travel (external).
- White noise source gives equal intensity at all frequencies.
- Thermal noise can be reduced by cooling the circuit components.
- Partition noise can be avoided by using components like diode where there is no current division.
- Quantisation noise can be reduced by reducing step size.
- Atmospheric and ionsopheric conditions affect RF and microwave transmission through free space. This noise cannot be avoided.
- Use of low noise devices, isolation transformers, opto-couplers, filters, clipping circuits, shielding, grounding, earthing etc. are some of the hardware ways of noise reduction.
- Averaging, Fourier transform filtering or coding of digital signal are typical software noise reduction techniques.

## 2.9 TERMINAL QUESTIONS

*Spend* 25 *Minutes*

1. Describe the difference between energy signal and power signal.
2. Why are repeating waveforms used for circuit testing?
3. A noise output of a resistor is amplified by a noiseless amplifier having the gain of 80 and bandwidth 40 kHz. A meter connected to the output of the amplifier reads 1 mV rms.
   (a) What is the value of resistance if it is operated at 47°C?
   (b) If the bandwidth of amplifier is reduced to 10 kHz, its gain remaining constant, what will be the meter reading?

4. What is the difference between internal and external noise? Give two examples of each type.
5. Describe the type of noise to which transistor is prone.
6. What is the use of Fourier analysis?

## 2.10 SOLUTIONS AND ANSWERS

**Self Assessment Questions**

1. ASCII code, Binary code, Hexadecimal code and Morse code are few examples of coded signals.
2. a) None, b) Energy signal and c) Power signal.
3. A 20 kHz signal has the period of 50 μs. For a square wave, the ON time = OFF time = 25 μs. Since the rise and fall time should be less than 10% of pulse duration they should be less than 2.5 μs each.
4. In English language there are 26 alphabets. Assuming equal probability of their occurrence, each will carry information

$$I = \log_2\left(\frac{1}{1/26}\right) = \log_2 26 = 4.7 \text{ bits.}$$

5. $C = B\log_2\left(1+\frac{S}{N}\right) = 10^6 \log_2(1+511) = 10^6 \times 9\text{bps.} = 9\text{Mbps.}$

6. $V_{nt} = \sqrt{4kT\Delta\omega_n R}$

$$= \sqrt{4\times 1.38\times 10^{-23}\,\text{JK}^{-1}\times 290\,\text{K}\times 10\times 10^6\,\text{Hz}\times 10\times 10^3\,\Omega}$$

$= 40 \times 10^{-6}$ V = 40 μV

7. $V_n = \frac{\Delta V_s}{\sqrt{12}}$

Here $\Delta V_s$ = 10 mV, Hence

$$\therefore V_n = \frac{10\times 10^{-3}\,\text{V}}{\sqrt{12}} = 2.89\,\text{mV}.$$

8. NR = $\frac{S_i / N_i}{S_o / N_o} = \frac{50}{20} = 2.5.$

NF = 10 $\log_{10}$ NR = 4 dB.

9. Assuming $V_{fb}$ = 0V for the diode, the circuit for limiting the signal amplitude within ± 5V is given in Fig. 2.20.

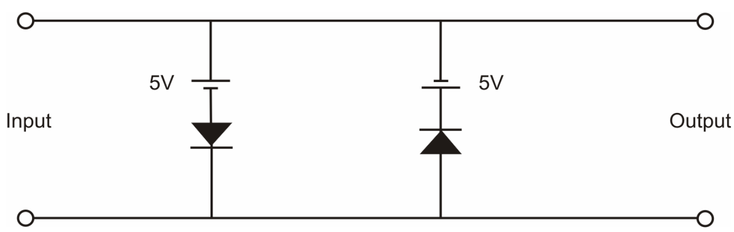


**Fig. 2.18: Diode Clipper**

## Terminal Questions



1. Energy signal is a signal with finite duration. Its energy is calculated by integrating the square of the signal over its entire duration. Power signal has infinitely long duration, but it is repetitive in nature. Its power is calculated by taking the time average of the square of the signal.

2. If a single event waveform (like single pulse or single step) is used, it is necessary to have an appropriate display and recording equipment to catch the single event response of the circuit/system and make it available for longer time to the observer for analysing it. This needs costly equipment like storage oscilloscope, fast response logic analyser etc. Another problem with single event waveform like step input is that, its amplitude (high state) does not remain constant infinitely, but sags down. If repetitive waveforms are used, constant amplitude can be obtained at every repetition.

3. (a) $V_{nt} = \sqrt{4kT\Delta\omega_n R}$

$$R = \frac{V_{nt}^2}{4k\mathrm{T}\Delta\omega_n} = \frac{(1\,\mathrm{mV}/80)^2}{4\times 1.38\times 10^{-23}\,\mathrm{JK}^{-1}\times 320\mathrm{K}\times 40\times 10^3\,\mathrm{Hz}}$$

$$= \frac{1.563\times 10^{-10}}{7.0656\times 10^{-16}} = 221\ \mathrm{k}\Omega$$

(b) For $\Delta\omega_n$ = 10 kHz; $V_{nt}$ reduces to half. Hence the reading on meter will be 0.5 mV.

4. The internal noise is generated by the circuit elements of the communication system, particularly the receiver. The external noise is introduced in the transmitting medium. The internal noise may be generated by thermal interaction between the free electrons and vibrating ions in a conductor. This is called Johnson's noise. The external noise can be introduced by the interference generated by sparking in electrical appliances or lightening in the atmosphere.

5. Johnson noise, transistor noise and partition noise are the main types of noise that are observed in a transistor. Since a transistor is made up of semiconductor material, the temperature variation affects the conductivity and Johnson noise is observed. The randomness of injection of electrons from emitter to base and then to collector governs the random transistor noise. The electrons/holes constituting emitter current get divided into base and collector current, which is a random process at the electronic level. This causes partition noise.

6. Fourier analysis converts the time domain signals into frequency domain. It allows us to find out the different frequency components in a signal and remove unwanted frequency components like the noise signal.

---

## Reference Material:


1. *Communication Systems* by Haykin, Simon; (III Edition) (John Weiley & Sons)
2. *Modern Digital and Analog Communication Systems* by Lathi, B.P.; (III Edition) (Oxford University Press)
3. *Electronic Communication Systems* by Kennedy, George; (III Edition) (Tata McGraw-Hill)