UNIVERSITY OF ENERGY AND NATURAL RESOURCES, SUNYANI

SCHOOL OF ENGINEERING

DEPARTMENT OF COMPUTER AND ELECTRICAL ENGINEERING

ELNG 307: ANALOG AND DIGITAL COMMUNICATIONS (202)

Material Version: 2.9.3

Tweneboah-Koduah S. [MEng.]

ELNG 307: ANALOG AND DIGITAL COMMUNICATIONS (202)

Quote: *This mind is not an empty vessel to be filled but a powerhouse to be ignited* – Plutarch

NOTE:
All the contents or materials found in this booklet not limited to images and or figures can be found in the following recommended reading textbooks and others indicated in references.

- Lathi, B.P. and Ding, Zhi, **Modem Digital and Analog Communication Systems**, Oxford University Press, Fourth Edition, 2009.
- Proakis, John G. **Digital Communications**, 4th ed. New York, McGraw-Hill, 2000. ISBN: 9780072321111.
- Proakis, John G. and Masoud Salehi **Communications System Engineering**. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2001. ISBN: 9780130617934.
- S. Haykin, **Communication Systems**, Wiley, 2001 .
- L.W. Couch II, **Digital and Analog Communication Systems**, Prentice-Hall, 2001.

Tweneboah-Koduah, Samuel [MEng.]

ANALOGUE AND DIGITAL COMMUNICATIONS

Contents

***

Part I

Fundamentals of Communication Systems

Chapter 1

Introduction to Electronic Communications

***

Successful and unsuccessful people
do not vary greatly in their abilities.
They vary in their desires to reach their
potential.

– John Maxwell

1.1 Overview

In ancient times, communication was done mainly by means of runners, carrier pigeons, smokes, drum beats, torches, etc. In most part of the world currently, people communicate by means of electronic communication systems, that can transmit signals over long distances at a speed of light. Electronic communication is much reliable and economical in many aspects of our economy. A communication system is system designed for the purpose of transmission and reception of information.

1.2 Analog and Digital Communication Sources

Analog communication information source produces messages that are defined within a continuum. That is analog messages are characterized by data whose values vary over continuous range. For example, the temperature at a location can varies with time and assume infinite number of possible values. A microphone is an example of analog signal source generator. On the other hand, a digital source of information produces a finite set of possible messages. A typewriter is an example. Digital message is transmitted in the binary form (i.e. *1*s and *0*s) for various assigned frequencies. Information from a binary source may be transmitted to the receiver or *sink* by sine wave of 100 Hz as binary *1* and 500Hz as binary *0*. A practical example of digital message is a Morse-coded telegraph, which consist of only mark and space. It is therefore a binary message, employing only two symbols. A digital message constructed with *M* symbols is called an *M*-ary message.

1.3 Transmission Mode

Communication system transmission can be classified as simplex and duplex system.

1.3.1 Simplex Systems

One-way communication in which the receiver only receives without sending back acknowledgment is referred to as simplex communication systems. Example of such systems is remote control, television and radio broadcasting.

1.3.2 Half-duplex

A half-duplex system allows the transmission in either direction but not at the same time. In this system, one has to wait for other to finish sending its data before it can acknowledge or send back data. At a point in time the transmitter becomes the receiver and vice versa. Example of half-duplex communication is talk-back radio.

1.3.3 Full-duplex

Two-way communication requires a transmitter and receiver at each end. A full-duplex system has channel that allows simultaneous transmission in both directions at the same time. In this case, there is no need to switch from transmit to receive mode as in half-duplex system. Examples is Telephone and data services.

1.4 Basic Structure of a Communication System

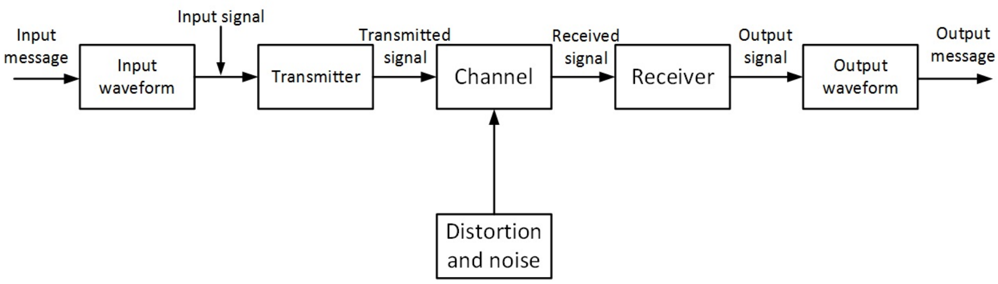


Figure 1.1: Basic structure of communication system

An electronic communication system is designed to send messages or information from a source that generates the messages to one or more destinations. The generated information is typically in the form of voice (speech source), a teletype message or television pictures. The message is then transformed into an electrical waveform known as **baseband signal** which is then modulated or modified and transmitted over a channel such as wire, coaxial cable, radio link or optical fibre. At the sink, the receiver reprocesses the signal from the channel by unwrapping (demodulation) and converting the electric signal to its original form, the message. An essential feature of any source that generates information is that the output is described in a probabilistic term, that is, the output at the sink from the source is not deterministic.

Figure 1.1 show a typical block diagram of a communication system. Every communication system consists of three essential parts namely transmitter, transmission channel and receiver. Both transmitter and receiver consists of a transducer and components needed to processes the message for transmission and reception. A transducer is usually required to convert the output of a source into an electrical signal that is suitable for transmission. For example, a microphone serves as the transducer that converts an acoustic speech signal into an electric signal, and a video camera converts an image into an electric signal. At the destination, a similar transducer is required to convert the electric signals that are received into a form suitable for the user.

1.5 Transmitter

A transmitter is a collection of electronic circuits designed to converts electrical signal into a form suitable for transmission based on the characteristics of the transmission channel. For example, in radio and TV broadcast, the National Communication Authority (NCA) specifies the frequency range for each transmitting station. Hence, the transmitter must translate the information signal to be transmitted into the appropriate frequency range that matches the frequency allocation assigned to the transmitter. Thus, signals transmitted by multiple radio stations do not interfere with each other. The transmitter performs matching of message signal to the channel in a process called **modulation**.

Modulation usually, involves the use of information signal to systematically vary the amplitude, frequency, or phase of a sinusoidal carrier. In carrier modulation such as AM, FM, and PM, modulation performed by the transmitter converts the information signal to a form that matches the characteristics of the channel. The process translates the information signal to frequencies that match the allocation of the channel.
The choice of modulation type is based on several factors such as:

- The amount of bandwidth allocated
- The types of noise and interference that the signal encounters in transmission
- Transmitter / Receiver type

Signal processing involves modulation and sometimes include coding. Most transmitters have built-in amplifier circuits that amplifies incoming signals from the transducer, encoders, etc. before transmission to the receiver.

1.6 Communication Channel

Communication channel is the electrical medium that bridges the source to the destination. It may be a pair of wires, a coaxial cable, or a radio wave or laser beam. Every channel introduces some amount of transmission loss or attenuation (that is signal degradation), so the signal power gradually decreases with increasing distance.

Due to physical limitations, communication channels have only finite bandwidth (B Hz), and the information bearing signal often suffers amplitude and phase distortion as it travels over the channel.

1.6.0.1 Waveguides

Similar to the transmission lines, waveguides are also used to guide electromagnetic waves from one point to other i.e. from source to load. A waveguide is considered as a special case of the transmission line. In general, waveguides are hollow conducting tubes having uniform cross-section. The most commonly used waveguides are rectangular waveguide and circular waveguide. Even though a waveguide is a special case of a transmission line, there are some notable differences between the two. Firstly, the transmission line can support only transverse electromagnetic wave (TEM wave), while a waveguide can support many different possible field configurations. Secondly, a transmission line becomes inefficient at microwave frequencies (3-300 GHz) due to skin effect and dielectric losses. But waveguide is used at that same range of frequencies to achieve larger bandwidth and lower signal attenuation. Lastly a transmission line may operate from DC (frequency 0) to a very high frequency. While a waveguide cannot transmit DC and below microwave frequencies it becomes excessively large.

1.6.0.2 Radio Communication

Radio is the broad general term applied to any form of wireless communication between two points. Radio communication is a wireless communication, requiring no physical wires between transmitter and receiver to carry the signal; on the contrary, the signal is sent through free space or air.

Radio waves are a type of electromagnetic radiation with wavelengths in the electromagnetic spectrum longer than infrared light. They have frequencies from 300 GHz to as low as 3 kHz, and corresponding wavelengths from 1 millimeter to 100 kilometers. Like all other electromagnetic waves, they travel at the speed of light. Radio waves are used for fixed and mobile radio communication, broadcasting, radar and other navigation systems, satellite communication, computer networks and innumerable other applications.

Radio communication essentially requires two antennas, one at the transmitting and the other at the receiving end. Using transmitting antenna, the transmitter transmits signal, over a carrier wave, into the free space. The receiver picks up the signal by means of receiving aerial and separates the signal from the carrier. Radio communication makes it possible to communication over very long distances, even from earth to moon.

1.7 Electromagnetic Wave Propagation

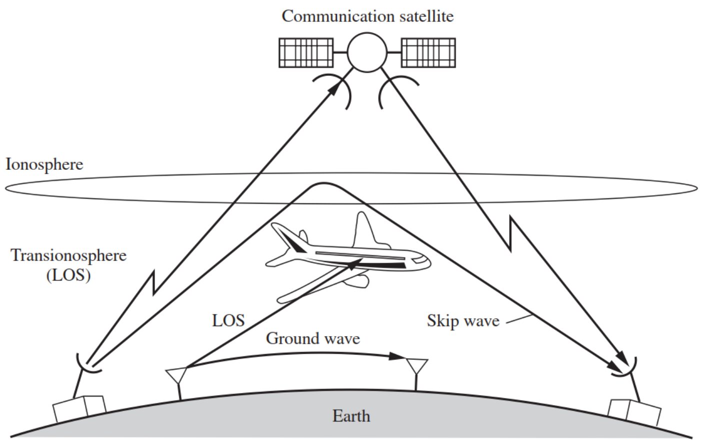


Figure 1.2: Types of electromagnetic wave propagation

In radio communication systems, electromagnetic energy is coupled to the propagation medium by an antenna which serves as the radiator. The physical size and the configuration of the antenna depend primarily on the frequency of operation. To obtain efficient radiation of electromagnetic energy, the antenna must be longer than 1/10 of the wavelength. The mode of propagation of electromagnetic waves in the atmosphere and in free space may be subdivided into three categories, namely, ground-wave propagation, sky-wave propagation, and line-of-sight (LOS) propagation as indicated in *Figure* 1.2.

In the VLF and ELF frequency bands, where the wavelengths exceed 10 km, the earth and the ionosphere act as a waveguide for electromagnetic wave propagation. In these frequency ranges, communication signals practically propagate around the globe. For this reason, these frequency bands are primarily used to provide navigational aids from shore to ships around the world. The channel bandwidth available in these frequency bands are relatively small (usually from 1% to 10% of the centre frequency), and hence, the information that is transmitted through these channels is relatively slow speed and, generally, confined to digital transmission.
A dominant type of noise at these frequencies is generated from thunderstorm activity around the globe, especially in tropical regions. Interference results from the many users of these frequency bands. Ground-wave propagation, as illustrated in *Figures* 1.2 and 1.4, is the dominant mode of propagation for frequencies in the MF band (0.3 to 3 MHz). This is the frequency band used for AM broadcasting and maritime radio broadcasting. In AM broadcasting, the range with ground-wave propagation of even the more powerful radio stations is limited to about 100 miles.

Atmospheric noise, man-made noise, and thermal noise from electronic components at the receiver are dominant disturbances for signal transmission of MF. Sky-wave propagation, *Figure* 1.5, results from transmitted signals being reflected (bent or refracted) from the ionosphere, which consists of several layers of charged particles ranging in altitude from 30 to 250 miles above the surface of the earth. During the daytime hours, the heating of the lower atmosphere by the sun causes the formation of the lower layers at altitudes below 75 miles. These lower layers, especially the *D*-layer serves to absorb frequencies below 2 MHz, thus severely limiting sky-wave propagation of AM radio broadcast. However, during the

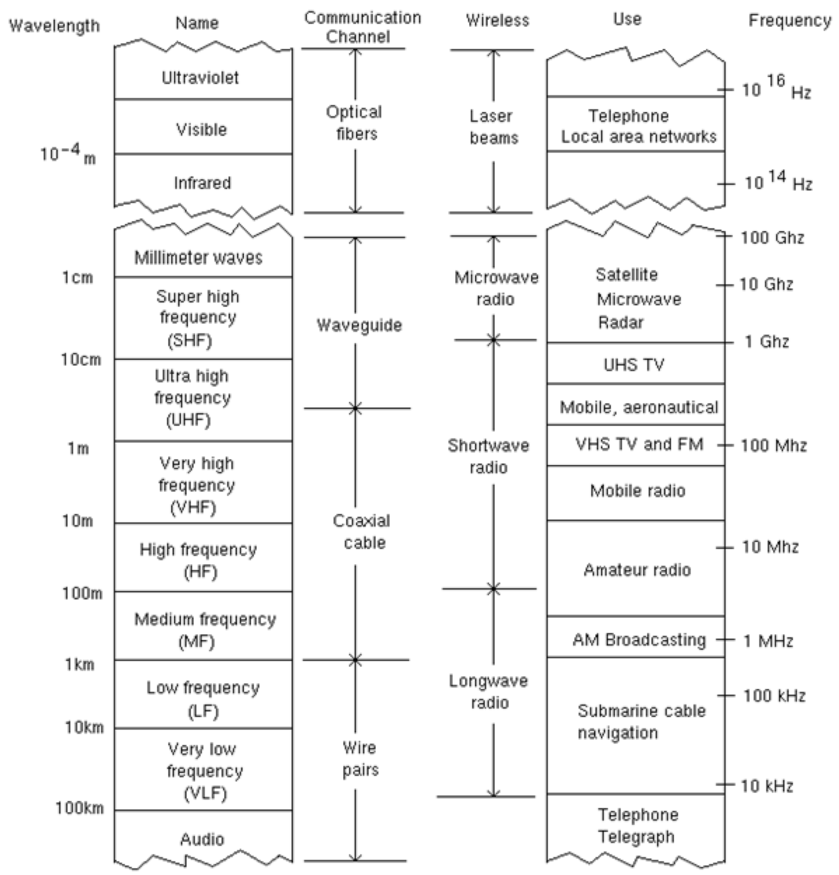


Figure 1.3: Electromagnetic spectrum and its applications

night-time hours, the electron density in the lower layers of the ionosphere drops sharply and the frequency absorption that occurs during the daytime is significantly reduced. As a consequence, powerful AM radio broadcast stations can propagate over large distances via sky wave over the F-layer of the ionosphere, which ranges from 90 miles to 250 miles above the surface of the earth.

A frequently occurring problem with electromagnetic wave propagation via sky wave in the HF frequency range is signal multipath. Signal multipath occurs when the signal multipath generally results in intersymbol interference in a digital communication system.

Moreover, the signal components arriving via different propagation paths may add destructively, resulting in a phenomenon called *signal fading*, which most people have experienced when listening to a distant radio station at night when sky wave is the dominant propagation Mode. Additive noise at HF is a combination of atmospheric noise and thermal noise.

Sky-wave ionospheric propagation ceases to exist at frequencies above approximately 30 MHz, which is the end of the HF band. However, it is possible to have ionospheric scatter propagation at frequencies in the range of 30 MHz to 60MHz, resulting from signal scattering from the lower ionosphere. Troposcatter results from signal scattering due to particles in the atmosphere at altitudes of 10 miles or less. Generally, ionospheric scatter and tropospheric scatter involve large signal propagation losses and require a large amount of transmitter power and relatively

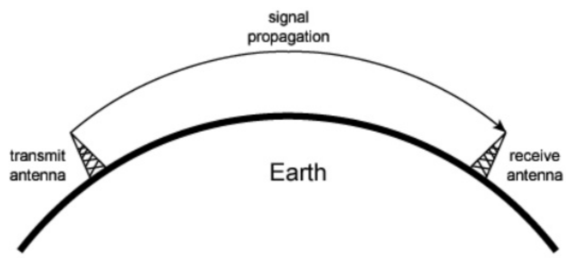


Figure 1.4: Illustration of ground-wave propagation

large antennas.

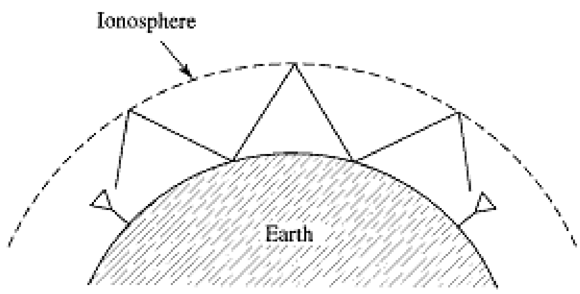


Figure 1.5: Illustration of sky-wave propagation

Frequencies above 30 MHz propagate through the ionosphere with relatively little loss and make satellite and extraterrestrial communications possible. Hence, at frequencies in the VHF band and higher, the dominant mode of electromagnetic propagation is line-of-sight (LOS) propagation. For terrestrial communication systems, the transmitter and receiver antennas must be in direct LOS with relatively little or no obstruction. For this reason, television stations transmitting in the VHF and UHF frequency bands mount their antennas on high towers to achieve a broad coverage area.

In general, the coverage area for LOS propagation is limited by the curvature of the earth. If the transmitting antenna is mounted at a height, h feet above the surface of the earth, the distance to the radio horizon, assuming no physical obstructions such as mountains, is approximately, $CA = \sqrt{2h}$ miles. For example, a TV antenna mounted on a tower of 1000 ft. in height provides coverage of approximately 50 miles. As another example, microwave radio relay systems used extensively for telephone and video transmission at frequencies above 1 GHz have antennas mounted on tall towers or on the top of tall buildings. The dominant noise limiting the performance of communication systems in the VHF and UHF frequency ranges is thermal noise generated in the receiver front end and cosmic noise picked up by the antenna. At frequencies in the SHF band above 10 MHz, atmospheric conditions play a major role in signal propagation. *Figure* 1.6 illustrates the signal attenuation in dB/mile due to precipitation for frequencies in there range of 10 to 100 GHz. We observe that heavy rain introduces extremely high propagation losses that can result in service outages (total breakdown in the communication system).

At frequencies above the EHF band, we have the infrared and visible light regions of the electromagnetic spectrum, which can be used to provide LOS optical communication in free

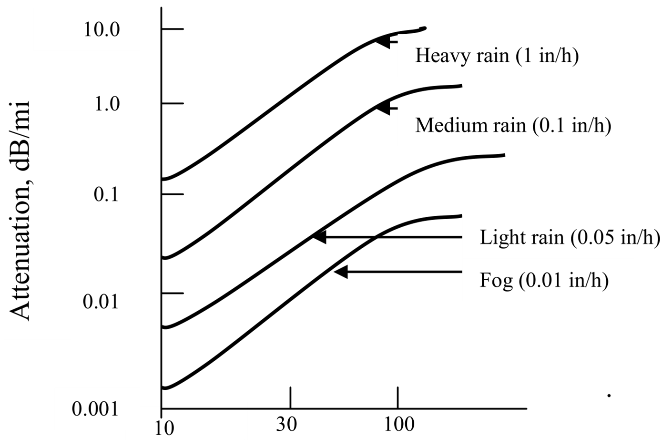


Figure 1.6: Signal attenuation due to precipitation

space. Although the most widely used media are conducting cables and free space (radio), other types of media are used in special communication systems. For example, in sonar, water is used as the medium. Passive sonar listens for under-water sounds with sensitive hydrophones. Active sonar uses an echo-reflecting technique similar to that used in radar for determining how far away objects under water are and in what direction they are moving. The earth itself can be used as a communication medium, because it conducts electricity and can also carry low-frequency sound waves. Alternating-current (AC) power lines, the electrical conductors that carry the power to operate virtually all our electrical and electronic devices, can also be used as communication channels. The signals to be transmitted are simply superimposed on or added to the power line voltage. This is known as carrier current transmission or power line communications (PLC). It is used for some types of remote control of electrical equipment and in some LANs.

1.8 Interference / Noise

Noise is random, undesirable electric energy that enters the communication system via the medium and interferes with the transmitted message. Some noise is also produced in the receiver. Noise can be either natural or man-made. Natural noise includes noise produced in nature, e.g. from lighting during rainy season, or noise due to radiations produced by the sun and the other stars. Man-made noise is the noise produced by electric ignition systems of cars, electric motors, fluorescent lights, etc. Noise is one of the serious problems of electronic communication. It cannot be completely eliminated. However, there are ways to deal with noise, and reduce the possibility of degradation of signal due to noise.

1.9 Receiver

The main function of the receiver is to extract the input message signal from the degraded version of the transmitted signal coming from the channel. It is basically a collection of electronic circuits designed to convert the electrical signal from the channel back to the original information. It consists of amplifier, detector, mixer, oscillator, transducer, tuned circuits, filters and so on. The receiver performs this function through the process of demodulation; the reverse of the transmitter's modulation process. In addition to demodulation, the receiver usually provides decoding, amplification (compensate for transmission loss) and filtering. The output is the original signal, which is then read out or displayed. It may be a voice signal sent to a speaker, a video signal that is fed to an LCD screen for display, or binary data that is received by a computer and then printed out or displayed on a video monitor.

1.10 Mathematical Models of Communication Channels

In the design of communication systems for transmitting information through physical channels, we find it convenient to construct mathematical models that reflect the most important characteristics of the transmission medium. Then, the mathematical model for the channel is used in the design of the channel encoder and modulator at the transmitter and the demodulator and channel decoder at the receiver. Below, we provide a brief description of the channel models that are frequently used to characterize many of the physical channels that we encounter in practice.

1.10.1 Additive Noise Channel Model

The simplest mathematical model for a communication channel is the additive noise channel, illustrated in *Figure* 1.7.

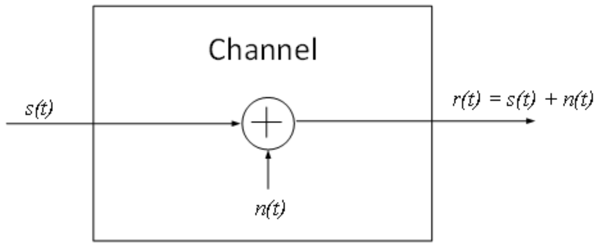


Figure 1.7: Additive noise channel model

In this model, the transmitted signal $s(t)$ is corrupted by an additive random noise process $n(t)$. Physically, the additive noise process may arise from electronic components and amplifiers at the receiver of the communication system, or from interference encountered in transmission as in the case of radio signal transmission. If the noise is introduced primarily by electronic components and amplifiers at the receiver, it may be characterised as thermal noise. This type of noise is characterised statistically as a Gaussian noise process. Hence, the resulting mathematical model applies to a broad class of physical communication channels. This model is know as Additive Gaussian Noise channel when apply to channel.

1.10.2 Linear Filter Channel Model

In some physical channels such as wireline telephone channels, filters are used to ensure that the transmitted signals do not exceed specified bandwidth limitations and thus, do not interfere with one another. Such channels are generally characterized mathematically as linear filter channels with additive noise, *Figure* 1.8. Hence, if the channel input is the signal $s(t)$, the channel output is the signal

$$(t) = s(t) * h(t) + n(t) = \int_{-\infty}^{+\infty} h(\tau)s(t-\tau)d\tau + n(t) \tag{1.10.1}$$

where $h(\tau)$ is the impulse response of the linear filter and denotes convolution.

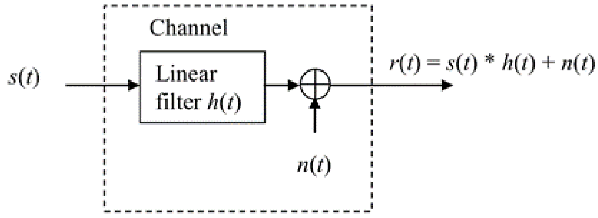


Figure 1.8: Linear filter channel model

When the signal undergoes attenuation in transmission through the channel, the received signal is

$$r(t) = \alpha s(t) + n(t) \tag{1.10.2}$$

where α represents the attenuation factor.

1.10.3 Linear Time-Variant Filter Channel

Physical channels such as underwater acoustic and ionospheric radio channels which result in time-variant multipath propagation of the transmitted signal may be characterized mathematically as time-variant linear filters. Such linear filters are characterized by a time-variant channel impulse response $h(\tau;t)$, where $h(\tau;t)$ is the response of the channel at time t due to an impulse applied at time $t-\tau$. Thus, τ represents the "delay" (elapsed-time) variable. The linear time-variant filter channel with additive noise is illustrated *Figure* 1.9. For an input signal $s(t)$, the channel output signal is

$$\int_{-\infty}^{+\infty} h(\tau;t)s(t-\tau)d\tau + n(t) \tag{1.10.3}$$

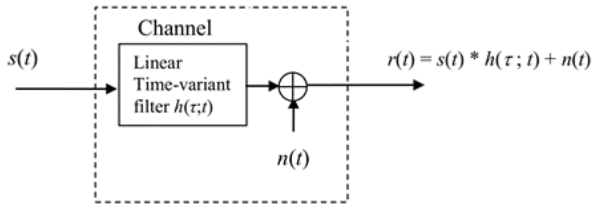


Figure 1.9: Linear time-variant filter channel mode

A good model for multipath signal propagation through physical channels, such as the ionosphere (at frequencies below 30 MHz) and mobile cellular radio channels, is a special case of *Eqn.* 1.10.3 in which the time-variant impulse response has the form

$$h(\tau;t) = \sum_{k=1}^{t} a_k(t)\,\delta(t-\tau_k) \tag{1.10.4}$$

where the $a_k(t)$ represents the possibly time-variant attenuation factors for the L multipath propagation paths. If *Eqn.* 1.10.4 is substituted into *Eqn.* 1.10.3, the received signal has the form

$$h(\tau;t) = \sum_{k=1}^{L} a_k(t)\delta(t-\tau_k) + n(t) \tag{1.10.5}$$

Hence, the received signal consists of L multipath components, where each component is attenuated by $a_k(t)$ and delayed by τ_k. The three mathematical models described above adequately characterize a large majority of physical channels encountered in practice. These three channel models are used in this text for the analysis and design of communication systems.

1.11 Modeling of Information Sources

Information source produces outputs that are of interest to the receiver of the information, who does not know the outputs in advance. The role of a communication system designer is to make sure that this information is transmitted correctly. Since the output of the information source is time-varying unpredictable function, it can be modeled as a random process.

The existence of noise in communication channels causes stochastic dependence between the input and output of the channel. Therefore, a communication system designer designs a system that transmits the output of a random process (information source), to a destination via a random medium (channel) and ensures low distortion. The properties of random process depend on the nature of the source. For example, when modeling speech signals, the resulting random process has all its power in a frequency band of approximately 300 – 4000 Hz.

Quick Pick 1 - 1: Communication Channels

A. **Electrical Conductors**

In its simplest form, the medium may simply be a pair of wires that carry a voice signal from a microphone to a headset. It may be a coaxial cable such as that used to carry cable TV signals. Or it may be a twisted-pair cable used in a local-area network (LAN).

B. **Optical Media**

The communication medium may also be a fiber-optic cable or “light pipe” that carries the message on a light wave. These are widely used today to carry long-distance calls and all Internet communications. The information is converted to digital form that can be used to turn a laser diode off and on at high speeds.

C. **Free Space**

When free space is the medium, the resulting system is known as radio. Radio is used in any wireless communication from one point to another. Radio makes use of the electromagnetic spectrum. Intelligence signals are converted to electric and magnetic fields that propagate nearly instantaneously through space over long distances.

D. **Other Types of Media**

Alternating-current (AC) power lines, the electrical conductors that carry the power can be used as communication channels. The signals to be transmitted are simply superimposed on or added to the power line voltage. This is known as carrier current transmission or power line communications (PLC).

1.12 Review Questions

1. What is the purpose and function of NCA?
2. Name the four main elements of a communication system, and draw a diagram that shows their relationship.
3. List five types of media used for communication, and state which three are the most commonly used.
4. Name the device used to convert an information signal to a signal compatible with the medium over which it is being transmitted.
5. What piece of equipment acquires a signal from a communication medium and recovers the original information signal?
6. What is a transceiver?
7. What are the two ways in which a communication medium can affect a signal? Name three common sources of interference.
8. What is the name given to the original information or intelligence signals that are transmitted directly via a communication medium?
9. What type of electronic signals are continuously varying voice and video signals?
10. What are on/off intelligence signals called?
11. How are voice and video signals transmitted digitally?
12. What terms are often used to refer to original voice, video, or data signals?
13. What technique must sometimes be used to make an information signal compatible with the medium over which it is being transmitted?
14. What is the process of recovering an original signal called?
15. What is a broadband signal?
16. Name the process used to transmit two or more baseband signals simultaneously over a common medium.
17. Name the technique used to extract multiple intelligence signals that have been transmitted simultaneously over a single communication channel.
18. What is the name given to signals that travel through free space for long distances?
19. Which two channels or media do light signals use for electronic communication?
20. Name two methods of transmitting visual data over a telephone network.
21. What is the name given to the signaling of individuals at remote locations by radio?
22. What do you call the systems of interconnections of PCs and other computers in offices or buildings?
23. What is a generic synonym for radio?
24. What are the four main segments of the communication industry? Explain briefly the function of each.
25. Why are standards important?
26. What types of characteristics do communication standards define?

1.13 Textbooks and References

[1] Principles of Communication Systems by Taub and Schilling,2nd Edition. McGraw Hill.

[2] Communication Systems by Siman Haykin,4th Edition, John Wiley and Sons Inc.

[3] Modern digital and analog communication system, by B. P. Lathi, 3rd Edition, Oxford University Press.

[4] Digital and analog communication systems, by L. W. Couch, 6th Edition, Pearson Education, Pvt. Ltd

Chapter 2

Introduction to Signals and Spectra Analysis

***

Some books are to be tasted, others to be swallowed, and some few to be chewed and digested

– Francis Bacon

2.1 Signals and Systems

2.1.1 Properties of Signals and Noise

In communication systems, the received waveform is usually categorized into the desired part containing the information and the extraneous or undesired part. The desired part is called the **signal**, and the undesired part is called **noise**.

The waveform of interest may be voltage or current as a function of time ($v(t)$ or $i(t)$). Often, the same mathematical techniques can be used when one is working with either type of waveform.

A. **Time Average Operator**

Definition: Time average operator is given by

$$\langle[\cdot]\rangle = \lim_{T\to\infty}\int_{-T/2}^{T/2}[\cdot]\,dt \tag{2.1.1}$$

$$\frac{h^2}{L_{o_{pm}}} \tag{2.1.2}$$

For a periodic signal with period, T_o, the time average operator

$$\langle[\bullet]\rangle = \frac{1}{T_o}\int_{-T/2+\alpha}^{T_o/2+\alpha}[\bullet]\,dt \tag{2.1.3}$$

where the period, T_o is the smallest positive number of the waveform and α is an arbitrary real constant that can be set to zero. As these (*Eqn.* 2.1.1 and 2.1.3) integrals are summed, the total area and T are proportionally larger, the resulting value for the time average is the same as just integrating over one period and dividing by the width of that interval, T_o.

B. DC Value of a Signal

The DC value of a waveform $s(t)$ is given by its time average *Eqn.* 2.1.1, that is $\langle S(t)\rangle$. In situation where the DC value of finite interval such as from t_1 to t_2 of a waveform is needed, the value is calculated as

$$S_{dc} = \frac{1}{t_2 - t_1}\int_{-t_1}^{t_2} s(t)dt \tag{2.1.4}$$

C. Decibel

The db (decibel) is a relative unit of measurement commonly used in communications for providing a reference for input and output levels. Power gain or loss. The decibel is a base 10 logarithmic measure of power ratios. For instance, the ratio of the power level at the output of a circuit compared with respect to that at the input is often specified by the decibel gain instead of the actual ratio. That is, it answers the question about “How much of the input power is in the output power?”

Definition: Decibel gain

$$\mathrm{dB} = 10\log\left(\frac{\mathrm{P_{out}}}{\mathrm{P_{in}}}\right) = 20\log\left(\frac{\mathrm{V_{rmsout}}}{\mathrm{V_{rmsin}}}\right) = 20\log\left(\frac{\mathrm{I_{rmsout}}}{\mathrm{I_{rmsin}}}\right) \tag{2.1.5}$$

Definition: Decibel power level with respect to 1mW

$$\begin{aligned}\mathrm{dBm} &= 10\log\left(\frac{\mathrm{actualpowerlevel(W)}}{10^{-3}}\right)\\ &= 30 + 10\log\left[\mathrm{actualpowerlevel(W)}\right]\end{aligned} \tag{2.1.6}$$

where the ”m” in the dBm denotes a milliwatt reference. In terms of voltage

$$\mathrm{dBmV} = 20\log\left(\frac{\mathrm{V_{rms}}}{10^{-3}}\right) \tag{2.1.7}$$

where 0 dBmV corresponds to -48.75 dBm.

The difference (or ratio) between two signal levels. It is used to describe the effect of system devices on signal strength. For example, a cable has 6 dB signal loss or an amplifier has 15 dB of gain. This is useful since signal strengths vary logarithmically, not linearly. dB scale is a logarithmic measure, it produces in simple numbers for large-scale variations in signals. It is very useful because system gains and losses can be calculated by adding and subtracting whole numbers. Every time power level is doubled (or halved), a 3dB is added (or subtracted) to (or from) the power level. This corresponds to a 50% gain or reduction. 10 dB gain/loss corresponds to a ten-fold increase/decrease in signal level. A 20 dB gain/loss corresponds to a hundred-fold increase/decrease in signal level. In other words, a device (like a cable) that has 20 dB loss through it will lose 99% of its signal by the time it gets to the other side. Thus, big variations in signal levels are easily handled with simple digits.

2.1.2 Signals

A signal under study in a communication system is generally expressed as a function of time or as a function of frequency. When the signal is expressed as a function of time, it gives us an idea of how that instantaneous amplitude of the signal is varying with respect to time. Whereas when the same signal is expressed as function of frequency, it gives us an insight of what are the contributions of different frequencies that make up that particular signal. Basically a signal can be expressed both in time domain and the frequency domain. There are various mathematical tools that aid us to get the frequency domain expression of a signal from the time domain expression and vice-versa. Fourier Series is used when the signal under study is a periodic one, whereas Fourier Transform may be used for both periodic as well as non-periodic signals

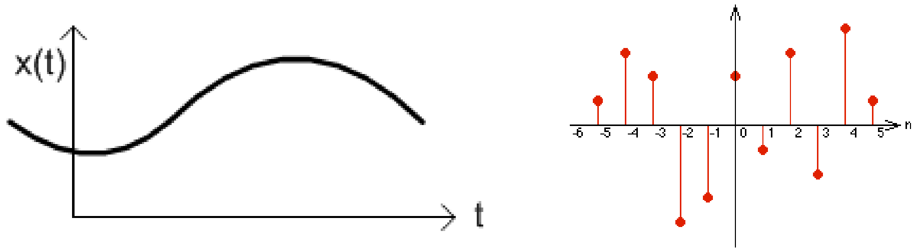
2.1.3 Classification of Signals

A signal, $x(t)$ or $f(t)$, is defined as a function of time ($t \in \Re$). Signals in engineering systems are typically described with the mathematical classifications below:

1. Continuous Time or Discrete Time (and Analog or Digital)
2. Periodic or Aperiodic
3. Deterministic or Random
4. Energy or Power

2.1.3.1 Continuous and Discrete Time Signals

As the names suggest, this classification is determined by whether or not the time axis is discrete (countable) or continuous, *Figure* 2.1a. A continuous-time signal will contain a value for all real numbers along the time axis. In contrast to this, a discrete-time signal, often created by sampling a continuous signal, will only have values at equally spaced intervals along the time axis. The difference between analog and digital is similar to the difference between continuous-time and discrete-time. However, in this case the difference involves the values of the function. Analog corresponds to a continuous set of possible function values, while digital corresponds to a discrete set of possible function values. A common example of a digital signal is a binary sequence, where the values of the function can only be one or zero.



(a) Continuous time signal (b) Discrete time signal

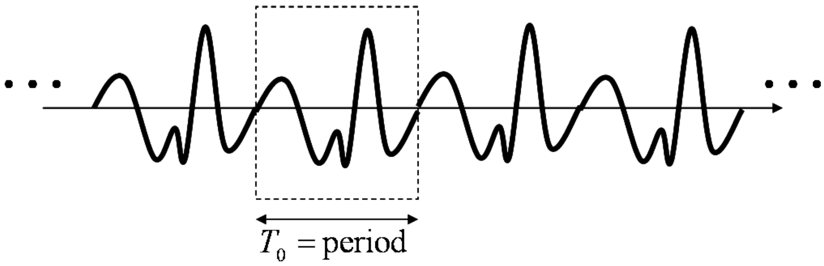
Figure 2.1: Time varying signals

2.1.3.2 Periodic and Aperiodic

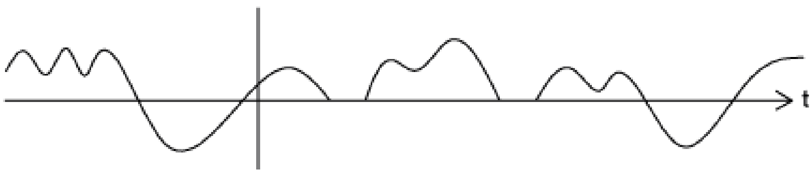
Periodic signals $f(t)$ repeat with some period T, while aperiodic, or non-periodic signals do not *Figure* 2.2b. Periodic function define a periodic can be defined through the following mathematical expression, where t can be any number and T is a positive constant

$$f(t) = f(t+T), \qquad -\infty < t < +\infty \tag{2.1.8}$$

The fundamental period of the function, $f(t)$, is the smallest value of T, i.e. T_o that still allows *Eqn.* 2.1.8 to be true.
By definition, a periodic signal $f(t)$ remains unchanged when time-shifted by one period. This means that a periodic signal must start at $t = -\infty$ because if it starts at some finite instant, say, $t = 0$, the time-shifted signal $f(t+T_o)$ will start at $f = -T_o$ and $f(t+T_o)$ would not be the same as $f(t)$. Therefore, a periodic signal, by definition, must start at $-\infty$ and continue forever, as shown in *Figure* 2.2a.



(a) Periodic signal



(b) Aperiodic signal

Figure 2.2: Periodicity of signals

Note that a periodic signal shifted by an integral multiple of T_o remains unchanged. Therefore, $f(t)$ may be considered a periodic signal with period mT_o, where m is any integer. However, by definition, the period is the smallest interval that satisfies periodicity condition, *Eqn.* (2.1.8). Therefore, T_o is the period. Another important property of a periodic signal $f(t)$ is that $f(t)$

can be generated by periodic extension of any segment of $f(t)$ of duration T_o (the period). This means that $f(t)$ can be generated from any segment of $f(t)$ with a time duration of one period by placing this segment and the reproduction end to end and infinitum on either side.

The fundamental angular frequency, ω_o, of a periodic signal is

$$\omega_o = 2\pi f_o = \frac{2\pi}{T_o} \tag{2.1.9}$$

Example of periodic signal is sinusoidal signals. These represent a class of periodic signals that are commonly used in many analysis techniques. These techniques, such as those involving Fourier series, decompose complicated waveforms into a series of sinusoidal waveforms. A sinusoidal waveform $f(t)$ is usually represented by

$$f(t) = A\cos(\omega t + \theta) \tag{2.1.10}$$

where A is the amplitude of the signal, θ is the phase angle of the signal, and ω is the angular frequency in radians per second $(\omega = 2\pi f)$

2.1.3.3 Phasor and Spectra

A useful periodic signal in system analysis is the signal

$$\tilde{x}(t) = Ae^{j(\omega_o t + \theta)}, \qquad -\infty < t < +\infty \tag{2.1.11}$$

which is characterized by three parameters: amplitude A, phase θ in radians, and frequency ω_o in radians per second or $f_o = \omega_o/2\pi$ Hz. $\tilde{x}(t)$ is referred to as rotating phasor to differentiate it from the phasor $Ae^{j\theta}$, for which $e^{j\omega_o t}$ is implicit. The rotating phasor $Ae^{j(\omega_o t+\theta)}$ can be related to a real, sinusoidal signal $A\cos(\omega t + \theta)$ in two ways.
The first thing to remember is taking its (*Eqn.* 2.1.11) real part

$$\begin{aligned} x(t) = A\cos(\omega_o t + \theta) &= \mathrm{Re}(\tilde{x}(t)) \\ &= \mathrm{Re}\left[Ae^{j(\omega_o t+\theta)}\right] \end{aligned} \tag{2.1.12}$$

and the second is by taking half of the sum of $\tilde{x}(t)$ and its complex conjugate

$$\begin{aligned} A\cos(\omega_o t + \theta) &= \frac{1}{2}\tilde{x}(t) + \frac{1}{2}\tilde{x}^*(t) \\ &= \frac{1}{2}Ae^{j(\omega_o t+\theta)} + \frac{1}{2}Ae^{-j(\omega_o t+\theta)} \end{aligned} \tag{2.1.13}$$

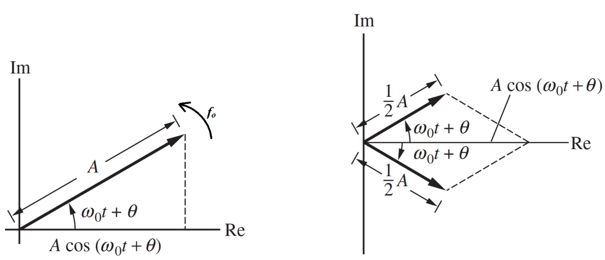


Figure 2.3: Two ways of relating a phasor signal to a sinusoidal signal. (a) Projection of a rotating phasor onto the real axis. (b) Addition of complex conjugate rotating phasors.

Two equivalent representations of $x(t)$ in the frequency domain may be obtained by noting that the rotating phasor signal is completely specified if the parameters A and θ are given for a particular f_o. Thus, plots of the magnitude and angle of $Ae^{j\theta}$ versus frequency give sufficient information to characterize $x(t)$ completely. Since, $\tilde{x}(t)$ exists only at the single frequency f_o, for this case of a single sinusoidal signal, the resulting plots consist of discrete lines called *line spectra*. The resulting plots are referred to as the *amplitude line spectrum* and *the phase line spectrum* for $x(t)$, *Figure* 2.4a. These are frequency-domain representations not only of $\tilde{x}(t)$ of $x(t)$ but by of *Eqn.* 2.1.12.

***NOTE* – Euler's Theorem:**

$$e^{\pm j\theta} = \cos\theta \pm j\sin\theta$$

where $j = \sqrt{-1}$ and θ is arbitrary angle.
Let $\theta = \omega_o t + \phi$, then the sinusoidal part of *Eqn.* 2.1.12 becomes

$$A\cos(\omega_o t + \phi) = \text{Re}\left[Ae^{j(\omega_o t+\phi)}\right] = \text{Re}\left[Ae^{i\phi}e^{j\omega_o t}\right]$$

The length of the phasor is A and it rotates counter-clockwise at a rate of f_o. At $t = 0$, the angle of rotation is with respect to the positive real axis. Remember that phase angles are measured with respect to *cosine* waves. Hence, sine waves must be converted to cosines via the identity

$$\sin\omega t = \cos\left(\omega t - 90^o\right)$$

and

$$\sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{2j} \quad \text{and} \quad \cos = \frac{e^{j\theta} + e^{-j\theta}}{2}$$

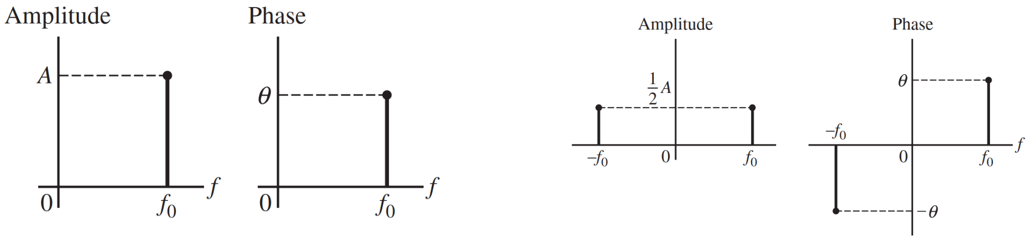
In phasors diagrams, amplitudes are always considered positive quantity so negative values or signs must be absorbed in the phase using

$$-A\cos\omega t = A\cos\left(\omega t + 180^o\right)$$

Note also that

$$\text{Re}\left[Z\right] = \frac{1}{2}\left(Z + Z^*\right)$$

In addition, the plots of *Figure* 2.4a are referred to as the single-sided amplitude and phase spectra of $x(t)$ because they exist only for positive frequencies. For a signal consisting of a sum of sinusoids of differing frequencies, the single-sided spectrum consists of a multiplicity of lines, with one line for each sinusoidal component of the sum. By plotting the amplitude and phase of the complex conjugate phasors of *Eqn.* 2.1.12 versus frequency, one obtains another frequency-domain representation for $x(t)$, referred to as the double-sided amplitude and phase spectra, *Figure* 2.4b.



(a) Single sided (b) Double sided

Figure 2.4: Amplitude and phase spectra for the signal $A\cos\left(\omega_o t + \theta\right)$

Two important observations may be made from *Figure* 2.4b.

i. The lines at the negative frequency $f = -f_o$ exist precisely because it is necessary to add complex conjugate (or oppositely rotating) phasor signals to obtain the real signal, $A\cos(\omega_o t + \theta)$.

ii. The amplitude spectrum has even symmetry and that the phase spectrum has odd symmetry about $f = 0$. This symmetry is again a consequence of $x(t)$ being a real signal.

As in the single-sided case, the two-sided spectrum for a sum of sinusoids consists of a multiplicity of lines, with one pair of lines for each sinusoidal component.

Example 2.1

Find the phase and spectrum representation of the signal in *Figure* 2.5

$$w(t) = 7 - 10\cos(40\pi t - 60^o) + 4\sin 120\pi t$$

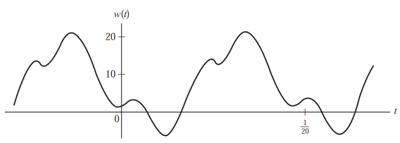


Figure 2.5: Complex waveform

Solution
Convert part of the signal into cosine terms

$$\Rightarrow \qquad w(t) = 7\cos 2\pi 0t + 10\cos(2\pi(20)t + 120^o) + 4\cos(2\pi(60)t - 90^o)$$

The amplitude and phase can be drawn from the result using

$$A\cos(\omega_o t + \phi) = \frac{A}{2}e^{j\phi}e^{j\omega_o t} + \frac{A}{2}e^{-j\phi}e^{-j\omega_o t}$$

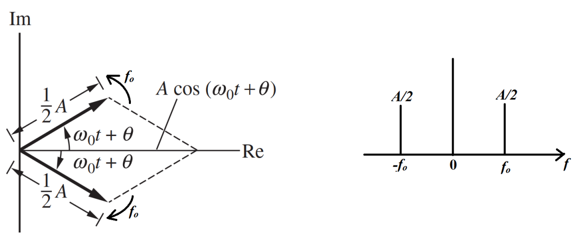


Figure 2.6: Phase diagram and spectrum

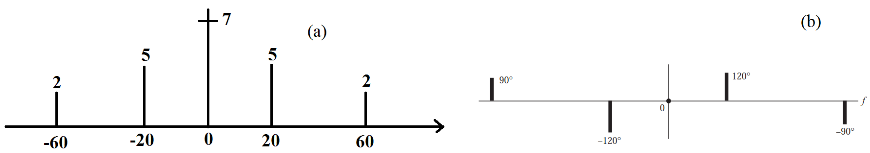


Figure 2.7: Spectrum

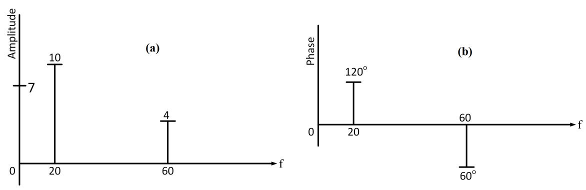


Figure 2.8: Amplitude (a) and phase (b) spectra

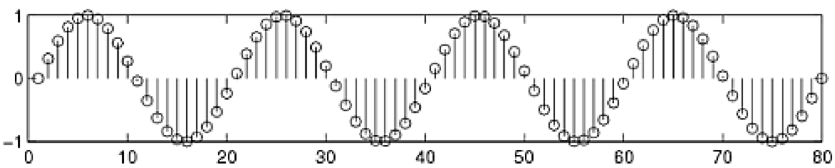
Figure 2.8 is one-sided spectra of $w(t)$. Now, let's obtain phasor and corresponding spectra diagrams from the signal.

2.1.3.4 Deterministic and Random Signals

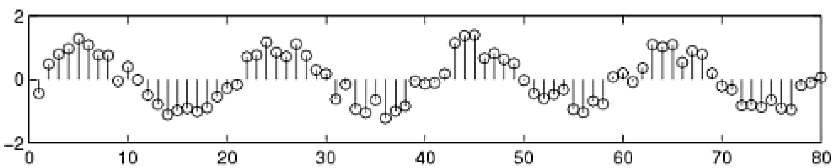
Quick Pick 2 - 1: Random Signals

If a signal is known only in terms of probabilistic description, such as *mean value*, *mean squared value*, and so on, rather than its complete mathematical or graphical description, then, the signal is a **random signal**. Most noise signals encountered in practice are random (in nature) signals.

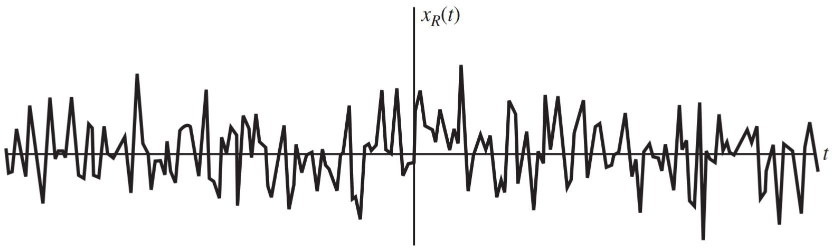
A deterministic signal is a signal in which each value of the signal is fixed and can be determined by a mathematical expression, rule, or table. Because of this the future values of the signal can be calculated from past values with complete confidence. On the other hand, a random signal has a lot of uncertainty about its behavior. The future values of a random signal cannot be accurately predicted and can usually only be guessed based on the averages of sets of signals *Figure* 2.1b.



(a) Deterministic signal



(b) Random signal



(c) Pure random signal

Figure 2.9: Random and deterministic signals

2.1.3.5 Energy and Power Signals

The size of any entity is a number that indicates the largeness or strength of that entity. Generally speaking, the signal amplitude varies with time.

Periodic signals are power signals; non-periodic signals (pulses) are energy signals. When both power and energy are infinite, the signal is neither a power nor an energy signal. As a matter of fact, a true power signal cannot exist in the real world because it would require a power source that operates for an infinite amount of time.

Quick Pick 2 - 2: Channel Capacity

In communication systems, if the received (average) signal power is sufficiently large compared to the (average) noise power, information may be recovered. This concept was demonstrated by the Shannon channel capacity formula

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

Where C is given in bits/sec, B is the channel bandwidth in hertz (Hz) and S/N is the signal-to-noise power ratio at the input of the receiver (normally digital). Channel capacity represents the maximum rate at which nearly errorless data transmission is theoretically possible. The channel capacity C has to be greater than the average information rate R of the source for errorless transmission.

The capacity C represents a theoretical limit, and the practical usable data rate will be much smaller than C. For example, for a typical telephone link with a usable bandwidth of 3 kHz and $S/N = 10^3$, the channel capacity is approximately 30,000 bits/sec. Data rate on such channels typically ranges from 150 to 9600 bits/sec.

Energy signal

Signal are often considered to be a function of varying amplitude through time; it seems to reason that, a good measurement of signal strength would be taking the area under the curve. However, this area may have a negative part. This negative part does not have less strength than a positive signal of the same size. The best way to find the energy of a signal is to either square the signal or by taking its absolute value, then finding the area under its curve. *Signal energy implies the area under the squared signal.* The energy of a signal $x(t)$ is given by

$$E_x = \int_{-\infty}^{\infty} x^2(t)dt \tag{2.1.14a}$$

More generally to consider complex waveforms

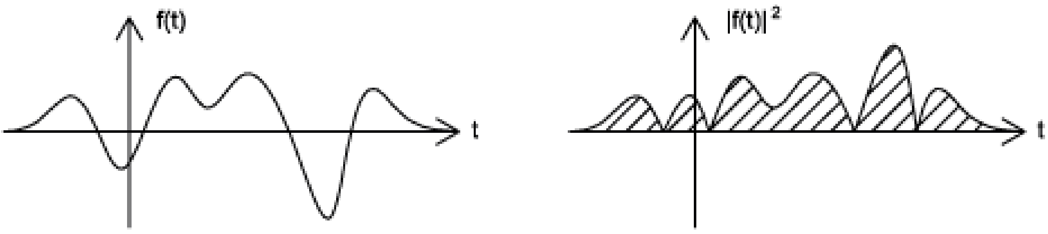
$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt \tag{2.1.14b}$$

The signal energy must be **finite** for it to be a meaningful measure of the signal size. A necessary condition for the energy to be finite is that the signal amplitude, $A_m \to 0$ as $|t| \to \infty$, *Figure* 2.10b. Otherwise the integral in *Eqn.* 2.1.14a will not converge.

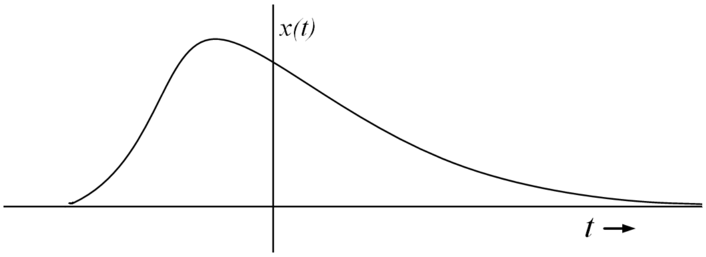
NOTE:	$x(t)$ is an *energy signal* if and only if $0 < E_x < \infty$, so that $P = 0$.

Power Signal

Our definition of energy seems reasonable, and it is. However, what if the signal does not



(a) The energy of the signal is shaded



(b) Finite energy

Figure 2.10: Energy signal

decay? In this case we have infinite energy for any such signal. Does this mean that a sixty hertz sine wave feeding into your headphones is as strong as the sixty hertz sine wave coming out of your outlet? Obviously not. This is what leads us to the idea of signal power. We

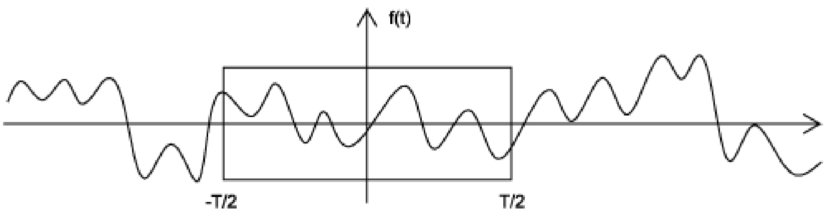


Figure 2.11: Computing power of a signal.

compute the energy per a specific unit of time, then allow that time to go to infinity. Power is a time average of energy (energy per unit time). This is useful when the energy of the signal goes to infinity. The power of a signal $x(t)$ is given by

$$P_x = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} (|x(t)|)^2 dt = \left\langle x^2(t) \right\rangle \tag{2.1.15a}$$

or

$$P_x = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} (|x(t)|)^2 dt = \left\langle x^2(t) \right\rangle \tag{2.1.15b}$$

If the signal energy is infinite, that is, the amplitude of $x(t)$, $A_m \not\to 0$ as $|t| \to \infty$. A more meaningful measure of the signal size in this case would be the time average of the energy (if it exists), which gives the average power P_x defined (for a real signal) by *Eqn.* 2.1.15a.

***NOTE*:** $x(t)$ is classified as *power signal* if and only if $0 < P_x < \infty$, implying that $E_x = \infty$.

For a periodic signal, $x_p(t)$ with period and arbitrary starting time respectively as T_o and t_o, the power is simply given by

$$P_x = \frac{1}{T_o}\int_{t_o}^{t_o+T_o} \left|x_p(t)\right|^2 dt \tag{2.1.16}$$

The average amount of power dissipated by a signal $x(t)$ during an interval of time from t_1 to t_2 is

$$P_x = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} |x(t)|^2 dt \tag{2.1.17}$$

The instantaneous power dissipated by a voltage $v(t)$ in a resistance R is given by

$$p(t) = \frac{|v(t)|^2}{R}$$

For current, the power is

$$p(t) = |i(t)|^2 R$$

It can be seen that the power in each case is proportional to the squared magnitude of the signal. If these signals are applied to a 1 ohm resistor, then both of the equations above assume the same form. In this case, the power value obtained is called **normalized power**. Normalized power is widely used in communication system analysis.

Example 2.2

Consider the periodic signal, $x_p(t) = A\cos(\omega_o t + \theta)$, the average power from *Eqn.* 2.1.16 is computed as follows

$$\begin{aligned}
P &= \frac{1}{T_o}\int_{t_o}^{t_o+T} A^2\cos^2(\omega_o t + \theta)\,dt \\
&= \frac{\omega_o}{2\pi}\int_{t_o}^{t_o+(2\pi/\omega_o)} \frac{A^2}{2} dt + \frac{\omega_o}{2\pi}\int_{t_o}^{t_o+(2\pi/\omega_o)} \frac{A^2}{2}\cos\left[2(\omega_o t + \theta)\right] dt \\
&= \frac{A^2}{2}
\end{aligned}$$

Note:
The identity $\cos^2 U = \frac{1}{2} + \frac{1}{2}\cos(2U)$ has been used, and the second integral goes to zero since the integration is over two complete periods of the intergrand.

Example 2.3

Determine the power and rms value of the following signals

(i) $g(t) = C\cos(\omega_o t + \theta)$

(ii) $f(t) = C_1\cos(\omega_1 t + \theta_1) + C_2\cos(\omega_2 t + \theta_2)$, $\qquad \omega_1 \neq \omega_2$

(iii) $s(t) = De^{j\omega_o t}$

Solution

(i) is a periodic signal with period $T_o = 2\pi/\omega_o$

The suitable measure of this signal is its power. Because it is a periodic signal, we may compute its power by averaging its energy over one period $2\pi/\omega_o$. The power of (i) is computed as

$$
\begin{aligned}
P_g &= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} C^2 \cos^2(\omega_o t + \theta)\, dt \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \frac{C^2}{2} [1 + \cos(2\omega_o t + 2\theta)]\, dt \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} dt + \lim_{T\to\infty} \frac{1}{T}\frac{C^2}{2} \int_{-T/2}^{T/2} \cos(2\omega_o t + 2\theta) dt
\end{aligned}
$$

The first term on the right becomes $C^2/2$ after the integration and limit. However, the second term goes to zero.

Now, the power of the signal is

$$P_g = \frac{C^2}{2}$$

This shows that a sinusoid of amplitude C has a power $C^2/2$ regardless of the value of its frequency $\omega_o\,(\omega_o \neq 0)$ and phase θ. The rms value is $C/\sqrt{2}$.

(ii) In this example, the power is

$$
\begin{aligned}
P_f &= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [C_1 \cos(\omega_1 t + \theta_1) + C_2 \cos(\omega_2 t + \theta_2)]^2 dt \\
&= \lim_{T\to\infty} \tfrac{1}{T} \textstyle\int_{-T/2}^{T/2} C_1^2 \cos^2(\omega_1 t + \theta_1)\, dt + \\
&\quad \lim_{T\to\infty} \tfrac{1}{T} \textstyle\int_{-T/2}^{T/2} C_2^2 \cos^2(\omega_2 t + \theta_2)\, dt + \\
&\quad \lim_{T\to\infty} \tfrac{2C_1C_2}{T} \textstyle\int_{-T/2}^{T/2} \cos(\omega_1 t + \theta_1)\cos(\omega_2 t + \theta_2)\, dt
\end{aligned}
$$

The actual power of **(ii)** is from the first and second integrals on the right-hand as respectively as $C_1^2/2$ and $C_2^2/2$ similar to **(i)**. With proper trigonometry identities, one can show that the third term goes to zero (as $T \to \infty$).

The overall power is

$$P_f = \frac{C_1^2}{2} + \frac{C_2^2}{2}$$

ant the rms value from p_f is

$$\sqrt{\frac{(C_1^2 + C_2^2)}{2}}$$

The results can be extended to sum of any number of sinusoids with distinct frequencies $\omega_n(\omega_n \neq 0)$. Thus, if

$$s(t) = \sum_{n=1}^{\infty} \alpha_n \cos(\omega_n t + \theta_n)$$

where none of the two sinusoids have identical frequencies.
The power of $x(t)$ is found as

$$P_x = \frac{1}{2}\sum_{n=1}^{\infty} \alpha_n^2$$

(iii) is left as exercise to the student.

Example 2.4

Determine the suitable measures of the signals in *Figure* 2.12

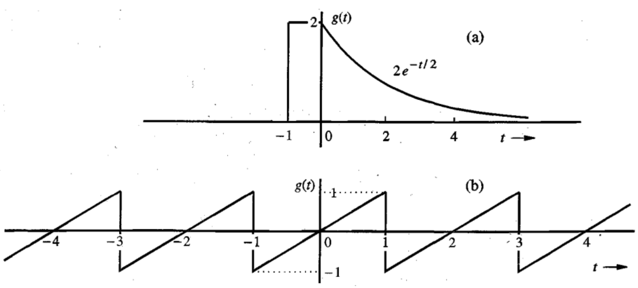


Figure 2.12: Signal for Example 2.4

The amplitude of *Figure* 2.12a $\to 0$ as $|t| \to \infty$. Therefore, the signal is energy signal. The energy is given by

$$\begin{aligned} E_g &= \int_{-\infty}^{\infty} g^2(t)dt \\ &= \int_{-1}^{0} (2)^2 dt + \int_{0}^{\infty} 4e^{-t} dt \\ &= 4 + 4 = 8 \end{aligned}$$

In *Figure* 2.12b, the amplitude does not go zeros ($\nrightarrow 0$) as $|t| \to \infty$. It is also periodic with period of 2, hence, power signal.

The power is calculated as

$$P_g = \frac{1}{2}\int_{-1}^{1} g^2(t)dt = \frac{1}{2}\int_{-1}^{1} t^2 dt = \frac{1}{3}$$

The **rms** value is $P_{rms} = \sqrt{P_g} = \frac{1}{\sqrt{3}}$.

2.2 Signal Operations

There are three basic signal operations: shifting, scaling, and inversion. Since the independent variable in our signal description is time, these operations are discussed as time shifting, time scaling, and time inversion (or folding).

2.2.1 Time Shifting

Consider a signal $x(t)$ *Figure* 2.13a and a type delayed by T seconds of $x(t)$ denoted as $s(t)$. Whatever happens in $x(t)$ *Figure* 2.13a at some instant t also in $s(t)$ *Figure* 2.13b T seconds at the instant $t+T$. Thus,

$$s(t+T) = x(t) \tag{2.2.1}$$

and

$$s(t) = x(t-T) \tag{2.2.2}$$

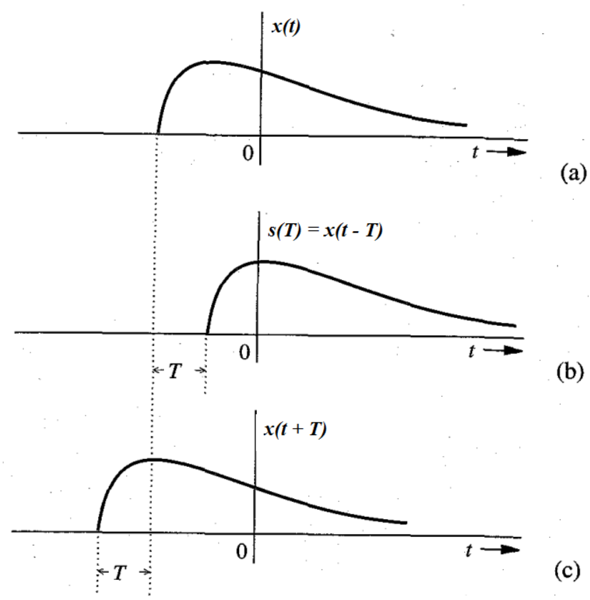


Figure 2.13: Time shifting signal

Therefore, to time-shift a signal by T, we replace t with $t-T$. Thus, $x(t-T)$ represents $x(t)$ time-shifted by T seconds. If T is positive, the shift is to the right (delay). If T is negative, the shift is to the left (advance). Thus, $x(t-2)$ is $x(t)$ delayed (right-shifted) by 2 seconds, and $x(t+2)$ is $x(t)$ advanced (left-shifted) by 2 seconds.

Example 2.5

Consider a triangularly shaped signal having piece wise continuous definition

$$s(t) = \begin{cases} 2t, & 0 \le t \le 1/2 \\ \frac{1}{3}(4-2t), & 1/2 \le t \le 2 \\ 0, & \text{otherwise} \end{cases}$$

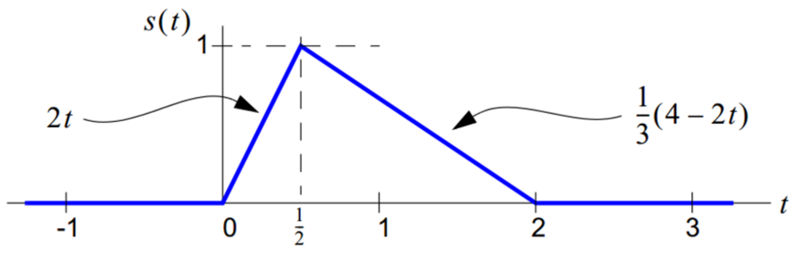


Figure 2.14: Triangular piecewise signal

Find $x_1(t) = s(t-2)$ and draw the corresponding signal.

Solution

As a starting point we note that $s(t)$ is active over just the interval $0 \leq t \leq 2$, so with $t \to (t-2)$ we have $0 \leq (t-2) \leq 2 \Rightarrow 2 \leq t \leq 4$. The piece wise definition of x_1 can be obtained by direct substitution of $t-2$ everywhere t appears.
After that, we get

$$x_1(t) = \begin{cases} 2(t-2), & 0 \leq (t-2) \leq 1/2 \\ \frac{1}{3}(4-2(t-2)), & 1/2 \leq (t-2) \leq 2 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} 2t-4, & 2 \leq t \leq 5/2 \\ \frac{1}{3}(8-2t), & 5/2 \leq t \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

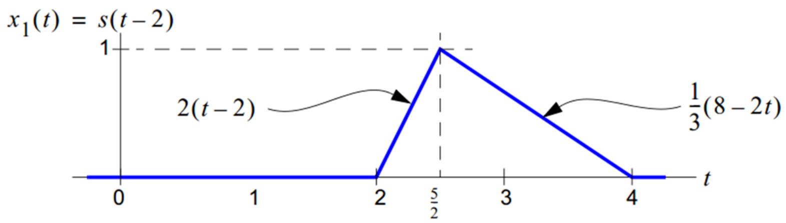


Figure 2.15: Signal after shifting

Note: The original signal $s(t)$ has moved to the right by 2 seconds.

2.2.2 Time Scaling

The compression or expansion of a signal in time is known as **time scaling**. Consider the signal $x(t)$ of *Figure* 2.16a. The signal $s(t)$ in *Figure* 2.16b is $x(t)$ compressed in time by a factor of 2. Therefore, whatever happens in $x(t)$ at some instant t also happens to $x(t)$ at the instant $t/2$, so that

$$s\left(\frac{t}{2}\right) = x(t) \tag{2.2.3}$$

and

$$s(t) = x(2t) \tag{2.2.4}$$

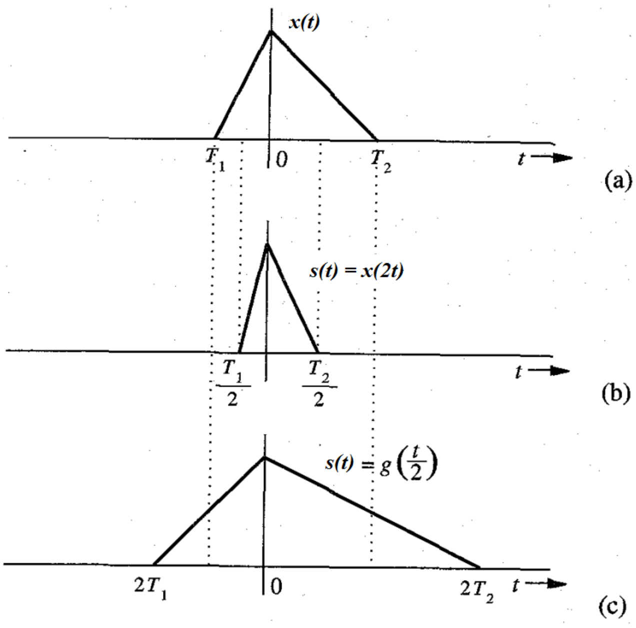


Figure 2.16: Time scaling signal

In general, if $x(t)$ is compressed in time by a factor $a(a > 1)$, the resulting signal $s(t)$ is given by setting the constant 2 in *Eqn.* 2.2.3 and *Eqn.* 2.2.4 to a.

2.2.3 Time Reversal

Time inversion or reversal is a special case of time scaling with $a = -1$ $(s(t) = x(at))$. Therefore, $s(-t) = x(t)$ and $s(t) = x(-t)$. This is done by replacing t with $-t$, thus, the time inversion of $x(t)$ gives $x(-t)$.

2.3 Singularity Function

An important subclass of aperiodic signals is the singularity functions. This function includes the *unit impulse function* $\delta(t)$ (or delta function) and the *unit step function* $u(t)$.

2.3.1 Unit Impulse Function

The unit impulse function is defined in terms of the integral

$$\int_{-\infty}^{\infty} x(t)\delta(t)dt = x(0) \tag{2.3.1}$$

where $x(t)$ is any test function that is continuous at $t = 0$. A change in variables and redefinition of $x(t)$ results in the *shifting property*

$$\int_{-\infty}^{\infty} x(t)\delta(t - t_o)dt = x(t_o) \tag{2.3.2}$$

where $x(t)$ is continuous at $t = t_o$. By considering the special case where $x(t) = 1$ for $t_1 \leq t \leq t_2$ and $x(t) = 0$ for $t < t_1$ and $t > t_2$, the two properties

$$\int_{-t_1}^{t_2} \delta(t - t_o)dt = 1, \qquad t_1 \leq t \leq t_2 \tag{2.3.3}$$

and

$$\delta(t - t_o) = 0, \qquad t \neq t_o \tag{2.3.4}$$

Eqn. 2.3.3 and 2.3.4 provide an alternative definition of the unit impulse.

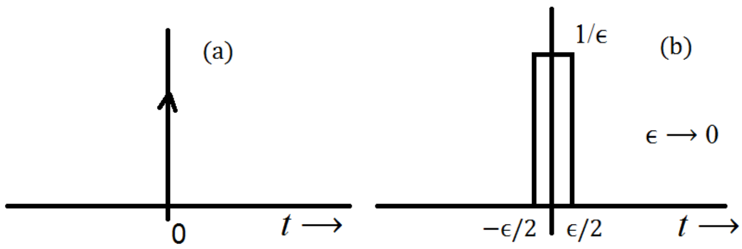


Figure 2.17: Unit impulse and its aproximation

2.3.2 Unit Step Function

Another useful function in system analysis is the *unit step function* $u(t)$, defined by *Figure* 2.18a. The function is given by

$$u(t) = \begin{cases} 1, & t \geq 1 \\ 0 & t < 0 \end{cases} \tag{2.3.5}$$

If we want a signal to start at $t = 0$ (so that it has a value of zero for $t < 0$), we only need to multiply the signal with $u(t)$. A signal that does not start before $t = 0$ is called a causal signal. In other words, $x(t)$ is a causal signal if

$$x(t) = 0, \qquad t < 0 \tag{2.3.6}$$

For example, the signal e^{-at} represents an exponential that starts at $t = -\infty$. If we want this signal to start at $t = 0$ (the causal form), it can be described as $e^{-at}u(t)$ (*Figure* 2.18b). Thus

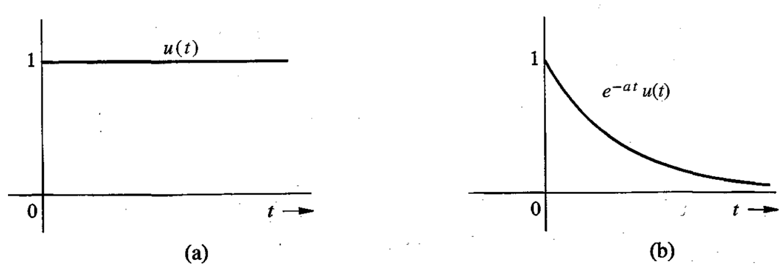


Figure 2.18: (a) Unit step function $u(t)$. (b) Causal exponential $e^{-at}u(t)$

from *Figure* 2.17b, the area from $-\infty$ to t under the limiting form of $\delta(t)$ is zero if $t < 0$ and

unity if $t \geq 1$.
Consequently

$$\int_{-\infty}^{t} \delta(\tau)d\tau = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases} = u(t) \tag{2.3.7}$$

From this result, it follows that

$$\frac{du}{dt} = \delta(t) \tag{2.3.8}$$

2.4 Fourier Series

Quick Pick 2 - 1: Fourier Synthesis

The equation below show various components of Fourier series

$$x(t) = \underbrace{A_o}_{\text{DC Part}} + \underbrace{\underbrace{\sum_{n=1}^{\infty} A_n \cos\left(\frac{2\pi nt}{T_o}\right)}_{\text{Even Part}} + \underbrace{\sum_{n=1}^{\infty} B_n \sin\left(\frac{2\pi nt}{T_o}\right)}_{\text{Odd Part}}}_{\text{T}_o \text{ is a period of all signal (Even and Odd)}}$$

Remember that $\omega_o = 2\pi/T_o$ is the fundamental angular frequency.

A periodic signal $x(t)$ with fundamental period, T_o can be represented as infinite sum of sinusoidal waveforms. This summation is called *Fourier series*. The Fourier series of a signal can be obtained if the following conditions (Dirichlet conditions) are satisfied:

1. $x(t)$ is absolutely integrable over its period

$$\int_{-\infty}^{\infty} |x(t)|\, dt < \infty$$

2. The number of maxima and minima of $x(t)$ in each period is finite
3. The number of discontinuities of $x(t)$ in each period is finite

The expansion of $x(t)$ is

$$x(t) = A_o + \sum_{n=1}^{\infty} A_n \cos\left(\frac{2\pi nt}{T_o}\right) + \sum_{n=1}^{\infty} B_n \sin\left(\frac{2\pi nt}{T_o}\right) \tag{2.4.1}$$

Where A_o is the average value of $x(t)$ given by

$$A_o = \frac{1}{T_o}\int_{-T_o/2}^{T_o/2} x(t)dt \tag{2.4.2}$$

Eqn. 2.4.2 is mostly taken as the DC component of the signal at zero frequency, and $2\pi n$ is the nth harmonic (integer multiple of the fundamental frequency). A_n and B_n are Fourier coefficients. There are described below

$$A_n = \frac{2}{T_o}\int_{-T_o/2}^{T_o/2} x(t)\cos\left(\frac{2\pi nt}{T_o}\right)dt \tag{2.4.3a}$$

$$B_n = \frac{2}{T_o} \int_{-T_o/2}^{T_o/2} x(t) \sin\left(\frac{2\pi nt}{T_o}\right) dt \tag{2.4.3b}$$

Alternative form of Fourier series

$$x(t) = C_o + \sum_{n=1}^{\infty} C_n \cos\left(\frac{2\pi nt}{T_o} - \phi_n\right) \tag{2.4.4}$$

where C_o, C_n, and ϕ_n are related to A_o, A_n and B_n by the equations

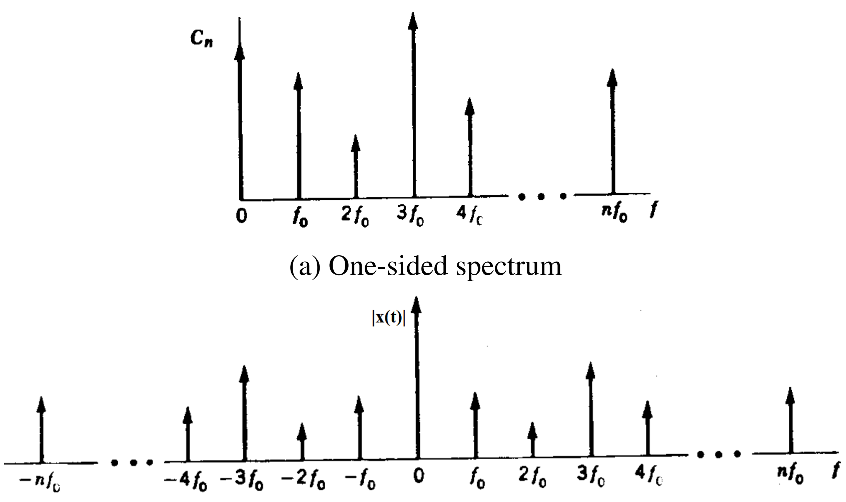
$$C_o = A_o \tag{2.4.5a}$$

$$C_n = \sqrt{A_n^2 + B_n^2} \tag{2.4.5b}$$

$$\phi_n = \tan^{-1}\frac{B_n}{A_n} \tag{2.4.5c}$$

The Fourier series of a periodic function is thus seen to consist of a summation of harmonics of a fundamental frequency $f_o = \frac{1}{T_o}$. The coefficients C_n are called **spectral amplitudes**; that is, C_n is the amplitude of the spectral component $C_n = \cos\left(\frac{2\pi nt}{T_o} - \phi_n\right)$ at frequency nf_o.

A typical amplitude spectrum of a periodic waveform is shown in *Figure* 2.19a. Here, at each harmonic frequency, a vertical line has been drawn having a length equal to the spectral amplitude associated with each harmonic frequency. This spectrum gives one-sided spectral representation of the signal. It is important to note that such an amplitude spectrum, lacking the phase information, does not specify the complete waveform of $x(t)$ as indicated earlier.



(a) One-sided spectrum

(b) Two-sided spectrum

Figure 2.19: Amplitude spectrum of a periodic waveform

2.4.1 Complex Exponential Fourier Series

Given a signal $x(t)$ defined over the interval $\alpha,\ \alpha+T_o$ with frequency definition; $\omega_o = 2\pi f_o = \frac{2\pi}{T_o}$. The complete exponential Fourier series is then given as

$$x(t) = \sum_{n=-\infty}^{\infty} X_n e^{jn\omega_o t}, \qquad \alpha \le t \le \alpha + T_o \tag{2.4.6}$$

where the series coefficients are related to $x(t)$ by

$$X_n = \frac{1}{T_o}\int_{\alpha}^{\alpha+T_o} x(t)e^{-jn\omega_o t}dt \tag{2.4.7}$$

The coefficients are complex quantities in general, that can be expressed in the **polar form**

$$X_n = |X_n|\, e^{j\arg X_n} \tag{2.4.8}$$

The n$^{\text{th}}$ term of this series is

$$X_n = |X_n|\, e^{j\arg X_n} e^{jn\omega_o t} \tag{2.4.9}$$

2.4.2 Parseval's Theorem

A periodic signal $x(t)$ is a power signal, and every term in its Fourier series is also a power signal. The power P_x of $x(t)$ is equal to the power of its Fourier series. Because the Fourier series consists of terms that are mutually orthogonal over one period, the power of the Fourier series is equal to the sum of the powers of its Fourier components. This follows from Parseval's theorem.

In *Example* 2.4, the technique for obtaining power of a signal has been demonstrated for the trigonometric Fourier series. It is also valid for the exponential Fourier series. Thus, for the trigonometric Fourier series

$$g(t) = C_o + \sum_{n=1}^{\infty} C_n \cos\left(n\omega_o t + \theta_n\right) \tag{2.4.10}$$

the power of $g(t)$ is given by

$$P_g = C_o^2 + \frac{1}{2}\sum_{n=1}^{\infty} C_n^2 \tag{2.4.11}$$

For exponential Fourier series

$$g(t) = D_o + \sum_{\substack{n=-\infty \\ (n\neq 0)}}^{\infty} D_n e^{jn\omega_o t} \tag{2.4.12}$$

the power is given by as in *Eqn.* 2.4.13. The power content of a periodic signal $x(t)$ is defined as the mean square value over a period (*Eqn.* 2.1.16). Parseval's theorem for Fourier series states that if $x(t)$ is a periodic signal with period, T_o, then

$$P_x = \frac{1}{T_o}\int_0^{T_o} |x(t)|^2 dt = \sum_{n=-\infty}^{\infty} |X_n|^2 \tag{2.4.13}$$

Note that $X_n(t)$ is given in *Eqn.* 2.4.7.

The theorem basically states that the power of a signal can be calculated using either the time or the frequency domain representation of the signal and the two results are identical.

For real $x(t)$, $|X_{-n}| = |X_n|$, the power from *Eqn.* 2.4.12 is

$$P_g = D_o^2 + 2\sum_{n=1}^{\infty} |D_n|^2 \tag{2.4.14}$$

Note here that Parseval's theorem occurs in many forms.

Example 2.6

Consider the signal

$$x(t) = \cos(\omega_o t) + \sin^2(2\omega_o t)$$

where $\omega_o = 2\pi/T_o$. Find the complex exponential Fourier series.

Solution

Using the appropriate trigonometric identities and Euler's theorem, $x(t)$ becomes

$$\begin{aligned} x(t) &= \cos(\omega_o t) + \frac{1}{2} - \frac{1}{2}\cos(4\omega_o t) \\ &= \frac{1}{2}e^{j\omega_o t} + \frac{1}{2}e^{-j\omega_o t} + \frac{1}{2} - \frac{1}{4}e^{j4\omega_o t} - \frac{1}{4}e^{-j4\omega_o t} \end{aligned}$$

By inspection term by term of the second line with $\sum_{n=-\infty}^{\infty} X_n e^{jn\omega_o t}$, we obtain

$$X_o = \tfrac{1}{2}$$

$$X_1 = \tfrac{1}{2} = X_{-1}$$

$$X_4 = -\tfrac{1}{4} = X_{-4}$$

with all other X_n equal to zero.

Note, trig. identity employed: $\sin^2 A = \frac{1}{2} - \frac{1}{2}\cos 2A$

Example 2.7

What are fundamental frequency and Fourier series coefficients of the following signal (Hint: You don't need to calculate any integral!). Illustrate the line spectrum of the signal.

$$s(t) = 10\sin(20\pi t) + 2\cos(40\pi t + \pi/3) + 0.5\sin(80\pi t + \pi/4)$$

Solution

$$s(t) = 10\cos(20\pi t - \frac{\pi}{2}) + 2\cos(40\pi t + \frac{\pi}{3}) + 0.5\cos(80\pi t + \frac{\pi}{4} - \frac{\pi}{2})$$
$$= 0 + 10\cos(2\pi \times 10 \times t - \frac{\pi}{2}) + 2\cos(2\pi \times 20 \times t + \frac{\pi}{3}) + 0.5\cos(2\pi \times 40 \times t - \frac{\pi}{4})$$

The fundamental frequency: $f_o = 10$

Fourier series coefficients (note this solution is in terms of single sided Fourier series representation):

$$A_0 = 0,\ A_1 = 10,\ \varphi_1 = -\pi/2,\ A_2 = 2,$$
$$\varphi_2 = \pi/3,\ A_3 = 0, A_4 = 0.5,\ \varphi_4 = -\pi/4,$$
$$S_k = 0, k >= 4$$

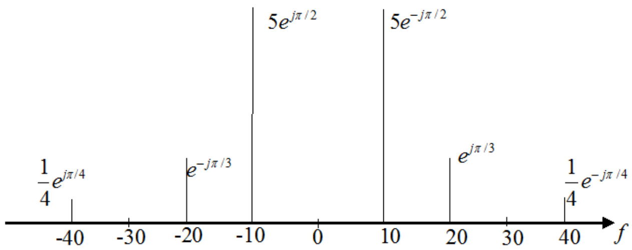
To obtain the two-sided Fourier series representation, we can write, using Euler formula

$$s(t) = 0 + 5e^{-j\pi/2}e^{j2\pi(10)t} + 5e^{j\pi/2}e^{-j2\pi(10)t} + e^{j\pi/3}e^{j2\pi(20)t} +$$
$$e^{-j\pi/3}e^{-j2\pi(20)t} + \frac{1}{4}e^{-j\pi/4}e^{j2\pi(40)t} + \frac{1}{4}e^{j\pi/4}e^{-j2\pi(40)t}$$

Thus the two-sided Fourier series coefficients are:

$$S_0 = 0,\ S_1 = 5e^{-j\pi/2},\ S_{-1} = 5e^{j\pi/2},$$
$$S_2 = e^{j\pi/3},\ S_{-2} = e^{-j\pi/3},$$
$$S_4 = \frac{1}{4}e^{-j\pi/4},\ S_{-4} = \frac{1}{4}e^{j\pi/4}$$

The two-sided line spectrum of the signal is as follows (line height indicates magnitude)



Example 2.8

What are the magnitude, frequency, and phase of the following sinusoidal signals?

a) $10\sin\left(20\pi t + {}^{\pi}/_{2}\right)$

b) $25\cos\left(10\pi t + {}^{\pi}/_{4}\right)$

Solution

	Magnitude (A)	Frequency (f_o)	Phase (ϕ)
a)	10	10	0
b)	25	5	${}^{\pi}/_{4}$

Note:

$$10\sin(20\pi t + \pi/2) = 10\cos(20\pi t + 0)$$

Example 2.9

Compute the first four terms in the Fourier series for a 1 kHz rectangular waveform with a pulse width of 500 μsec and an amplitude of 10 V.

Solution

$T = \text{time} = \frac{1}{1\times 10^3} = 1\times 10^{-3}$

$\tau = \text{pulse width} = 500\times 10^{-6}$

$A = 10$ V

$$\frac{\tau}{T} = \frac{500\times 10^{-6}}{1\times 10^{-3}} = 0.5$$

$$f(t) = [(10)(0.5)] + \left[(2)(10)(0.5)\frac{\sin(0.5\pi)}{0.5\pi}\cos\left(2\pi\times 10^3 t\right)\right]$$
$$+\left[(2)(10)(0.5)\frac{\sin(\pi)}{\pi}\cos\left(4\pi\times 10^3 t\right)\right]$$
$$+(2)(10)(0.5)\left(\frac{\sin(1.5\pi)}{1.5\pi}\right)\cos\left(6\pi\times 10^3 t\right)$$

$$f(t) = [5] + \left[\frac{10n_1}{0.5\pi}\cos 2\pi \times 10^3 t\right] + \left[\frac{10n_2}{\pi}\cos 4\pi \times 10^3 t\right]$$
$$+ \left[\frac{10n_3}{1.5\pi}\cos 6\pi \times 10^3 t\right]$$
$$f(t) = [5] + \left[6.366\cos\left(2\pi \times 10^3 t\right)\right] + [0] + \left[-2.122\cos\left(6\pi \times 10^3 t\right)\right]$$

Because this waveform is a symmetrical square waveform, it has components at (An_o) DC, and (An_1) 1 kHz and (An_3) 3 kHz points, and at odd multiples thereafter.

2.5 Sampling Function

A function frequently encountered in spectral analysis is the sampling function $Sa(x)$ defined by

$$Sa(x) = \frac{\sin x}{x} \tag{2.5.1}$$

A closely related function normally encountered in system analysis is *sinc x* (or sine cardinal of x). *Figure* 2.20 show the plot of Sa and Sinc functions.

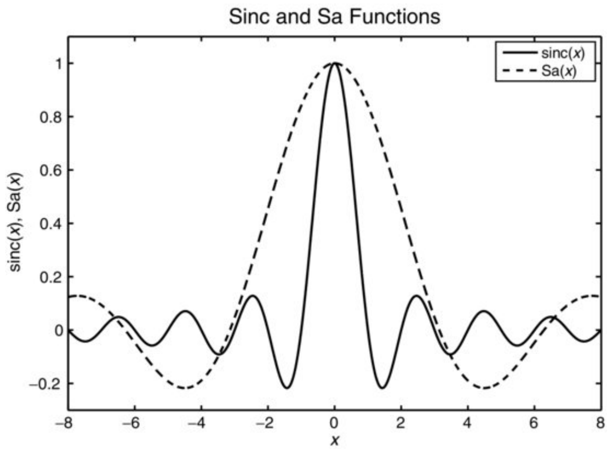


Figure 2.20: Sampling function – Sinc (x) and Sa (x)

2.5.1 The Sinc Function

The sinc(x) function, also known as the "sampling function" is a function that arises frequently in signal processing and the theory of Fourier transforms. The full name of the function is "sine cardinal," but it is commonly referred to by its abbreviation, "sinc." It is defined as

$$\sin c(x) = \begin{cases} 1, & \text{for } x = 0 \\ \frac{\sin x}{x}, & \text{otherwise} \end{cases} \tag{2.5.2}$$

where $\sin x$ is a sine function.

The $sa(x)$ is symmetrical about $x = 0$, and is maximum at this point $sa(x) = 1$. It oscillates with an amplitude that decreases with increasing x. It crosses zero at equal intervals on x at every $x = \pm n\pi$, where n is an non-zero integer.

2.6 Fourier Transform

A periodic waveform may be expressed, as we have seen, as a sum of spectral components. These components have finite amplitudes and are separated by finite frequency intervals $f_o = 1/T_o$. The normalized power of the waveform is finite, as is also the normalized energy of the signal in an interval T_o. Now, suppose we increase without limit the period T_o of the waveform. Thus, say, in *Figure* 2.21 the pulse centered around $t = 0$ remains in place, but all other pulses move outward away from $t = 0$ as $T_o \to \infty$.

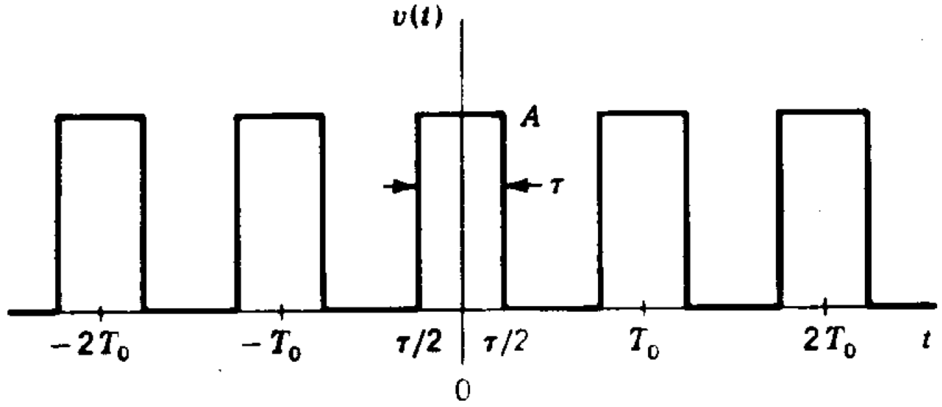


Figure 2.21: A periodic train of pulses duration τ

Then eventually we would be left with a single-pulse non-periodic waveform.

As $T_o \to \infty$. the spacing between spectral components becomes infinitesimal. The frequency of the spectral components, which in the Fourier series was a discontinuous variable with a one-to-one correspondence with the integers, becomes instead a continuous variable. The normalized energy of the non-periodic waveform remains finite, but, since the waveform is not repeated, its normalized power becomes infinitesimal. The spectral amplitudes similarly become infinitesimal. The Fourier series (this time, let's use $v(t)$) for the aperiodic waveform

$$v(t) = \sum_{n=-\infty}^{\infty} V_n e^{j2\pi n f_o t} \tag{2.6.1}$$

becomes

$$v(t) = \int_{-\infty}^{\infty} V(f) e^{j2\pi f t} df \tag{2.6.2}$$

The finite spectral amplitudes V_n are analogous to the infinitesimal spectral amplitude $V(f)df$. The quantity $V(f)$ is called the *amplitude spectral density* or generally **Fourier transform** of $v(t)$.

The Fourier transform is given by

$$V(f) = \int_{-\infty}^{\infty} v(t) e^{-j2\pi f t} dt \tag{2.6.3}$$

therefore V_n is given by

$$V_n = \frac{1}{T_o} \int_{-T_o/2}^{T_o/2} v(t) e^{-j2n\pi f_o t} dt \tag{2.6.4}$$

$$V_n = \frac{1}{T_o}\int_{-T_o/2}^{T_o/2} v(t)e^{-j2n\pi f_o t}dt$$

The inverse Fourier transform of $V(f)$, symbolized by $\mathscr{F}^{-1}$, is defined by

$$v(t) = \mathscr{F}^{-1}[V(f)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} V(f)e^{j2\pi nft}df \tag{2.6.5}$$

Eqn. 2.6.3 and 2.6.5 are often called the *Fourier transform pair* denoted by

$$v(t) \leftrightarrow V(f) \tag{2.6.6}$$

Employing the transfer function $H(f)$ of a network, if the input signal is $v(t)$, then the output $v_o(t)$ is given by

$$v_o(t) = \int_{-\infty}^{\infty} H(f)V(f)e^{j2\pi ft}df \tag{2.6.7}$$

Comparing *Eqn.* 2.6.2 with *Eqn.* 2.6.7, it can be seen that the Fourier transform, $V_o(f) \equiv \mathscr{F}[v_o(t)]$ is related to the transform, $V_i(f)$ of $v_i(t)$ by

$$\mathscr{F}[v_o(t)] = H(f)\mathscr{F}[v_i(t)] \tag{2.6.8}$$

or

$$V_o(f) = H(f)V_i(f) \tag{2.6.9}$$

Note:
The frequency, f in the the Fourier transform analysis can be replaced with ω. This is left to the student to find out how?

2.6.1 Fourier Transform Properties

1. **Linearity (Superposition)**

$$a_1x_1(t) + a_2x_2(t) \Leftrightarrow a_1X_1(\omega) + a_2X_2(\omega) \tag{2.6.10}$$

2. **Time Shifting**

$$x(t - t_o) \Leftrightarrow X(\omega)e^{-j\omega t_o} \tag{2.6.11}$$

Eqn. 2.6.11 shows that the effect of a shift in the time domain is simply to add a linear term $-\omega t_o$ to the original phase spectrum $\theta(\omega)$.

3. **Frequency Shifting**

$$x(\omega)e^{j\omega_o t} \Leftrightarrow X(\omega - \omega_o) \tag{2.6.12}$$

4. **Scaling**

$$x(at) \Leftrightarrow \frac{1}{|a|}X\left(\frac{\omega}{a}\right) \tag{2.6.13}$$

Eqn. 2.6.13 implies that time compression of a signal $(a > 1)$ results in its spectral expansion and that time expansion of the signal $(a < 1)$ results in its spectral compression.

5. **Time-Reversal**

$$x(-t) \Leftrightarrow X(-\omega) \tag{2.6.14}$$

6. **Duality**

$$X(t) \Leftrightarrow 2\pi x(-\omega) \tag{2.6.15}$$

7. **Differentiation**
Time differentiation

$$x'(t) = \frac{d}{dt}x(t) \Leftrightarrow j\omega X(\omega) \tag{2.6.16}$$

Frequency differentiation

$$(-jt)x(t) \Leftrightarrow X'(\omega) = \frac{d}{d\omega}X(\omega) \tag{2.6.17}$$

8. **Integration** If $X(0) = 0$, then

$$\int_{-\infty}^{t} x(\tau)d\tau \Leftrightarrow \frac{1}{j\omega}X(\omega) \tag{2.6.18}$$

This shows the effect of integration in the time domain is the division of $X(\omega)$ by $j\omega$ in the frequency domain, assuming that $X(\omega) = 0$. Note that

$$X(0) = \int_{-\infty}^{\infty} x(t)dt \tag{2.6.19}$$

9. **Convolution** The convolution of two signals $x_1(t)$ and $x_2(t)$, denoted by $x_1 * x_2(t)$, is a new signal $x(t)$ defined as

$$x(t) = x_1(t) * x_2(t) = \int_{-\infty}^{\infty} x_1(\tau)x_2(t-\tau)d\tau \tag{2.6.20}$$

$$x_1(t) * x_2(t) \Leftrightarrow X_1(\omega)X_2(\omega) \tag{2.6.21}$$

Eqn. 2.6.21 is time convolution theorem; it states that the convolution in the time domain becomes multiplication in the frequency domain. Note that the operation of convolution is commutative.

10. **Multiplication**

$$x_1(t)x_2(t) \Leftrightarrow \frac{1}{2\pi}X_1(\omega) * X_2(\omega) \tag{2.6.22}$$

2.6.2 Fourier Transform Pairs

Table 2.1: Table of Fourier Transform

Time Domain $(x(t))$	Frequency Domain $(X(f))$
1	$\delta(f)$
$\delta(t-t_o)$	$e^{-j2\pi f t_o}$
$\cos(2\pi f_o t)$	$\frac{1}{2}\delta(f-f_o)+\frac{1}{2}\delta(f+f_o)$
$\sin(2\pi f_o t)$	$-\frac{1}{2j}\delta(f+f_o)+\frac{1}{2j}\delta(f-f_o)$
$\prod(t)=\begin{cases}1, & \|t\|<\frac{1}{2}\\ \frac{1}{2}, & t=\pm\frac{1}{2}\\ 0, & \text{otherwise}\end{cases}$	$\sin c(f)$
$\sin c(t)$	$\prod(f)$
$\Lambda(t)=\begin{cases}t+1, & -1\leq t<0\\ -t+1, & 0\leq t<0\\ 0, & \text{otherwise}\end{cases}$	$\sin c^2(f)$
$\sin c^2(t)$	$\Lambda(f)$
$e^{-\alpha t}u_{-1}(t), \quad \alpha>0$	$\frac{1}{\alpha+j2\pi f}$
$te^{-\alpha t}u_{-1}(t), \quad \alpha>0$	$\frac{1}{(\alpha+j2\pi f)^2}$
$e^{-a\|\|t\|}$	$\frac{2\alpha}{\alpha^2+(2\pi f)^2}$
$e^{-\pi t^2}$	$e^{-\pi f^2}$
$\text{sgn}(t)=\begin{cases}1, & t>0\\ -1, & t<0\\ 0, & t=0\end{cases}$	$\frac{1}{(j\pi f)}$
$u_{-1}(t)$	$\frac{1}{2}\delta(f)+\frac{1}{j2\pi f}$
$\delta'(t)$	$j2\pi f$
$\delta^{(n)}(t)$	$(j2\pi f)^n$
$\frac{1}{t}$	$-j\pi\text{sgn}(f)$
$\sum_{n=-\infty}^{n=+\infty}\delta(t-nT_o)$	$\frac{1}{T_o}\sum_{n=-\infty}^{n=+\infty}\delta\left(f-\frac{n}{T_o}\right)$

Example 2.10

If $v(t) = \cos\omega_o t$, find $V(f)$

Solution
$v(t)$ is periodic, therefore the Fourier series and transform can be found.
Using the exponential Fourier series form of $v(t)$ is

$$v(t) = \frac{1}{2}e^{+j\omega_o t} + \frac{1}{2}e^{-j\omega_o t}$$

where $\omega_o = \frac{2\pi}{T_o}$, thus

$$V_1 = V_{-1} = \frac{1}{2}$$

and

$$V_n = 0, \qquad\qquad n \neq \pm 1$$

The Fourier transform $V(f)$ is

$$\begin{aligned} V(f) &= \int_{-\infty}^{\infty} \cos\omega_o t e^{-j2\pi f t} dt = \frac{1}{2}\int_{-\infty}^{\infty} e^{-j2\pi(f-f_o)t} + \frac{1}{2}\int_{-\infty}^{\infty} e^{-j2\pi(f+f_o)t} dt \\ &= \frac{1}{2}\delta(f - f_o) + \frac{1}{2}\delta(f + f_o) \end{aligned}$$

Note:

$$\delta(t) = \int_{-\infty}^{\infty} e^{j2\pi f t} df = \int_{-\infty}^{\infty} e^{-j2\pi f t} df$$

The Fourier transform of a sinusoidal signal (or other periodic signal) consists of impulses located at each harmonic frequency of the signal, that is $f_n = \dfrac{n}{nT_o} = nf_o$. The strength of each impulse is equal to the amplitude of the Fourier coefficient of the exponential series.

Example 2.11

A signal $m(t)$ is multiplied by a sinusoidal waveform of frequency, f_c. The resulting signal is

$$x(t) = m(t)\cos 2\pi f_c t$$

If the Fourier transform of $m(t)$ is $M(f)$, that is

$$M(f) = \int_{-\infty}^{\infty} m(t)e^{-j2\pi f t} dt$$

find the Fourier transform of $x(t)$.

Solution
Now, since

$$m(t)\cos 2\pi f_c t = \frac{1}{2}m(t)e^{j2\pi f_c t} + \frac{1}{2}m(t)e^{-j2\pi f_c t}$$

then the Fourier transform $X(f)$ is given by

$$X(f) = \frac{1}{2}\int_{-\infty}^{\infty} m(t)e^{-j2\pi(f+f_c)t}dt + \frac{1}{2}\int_{-\infty}^{\infty} m(t)e^{-j2\pi(f-f_c)t}dt$$

Putting these equation together by comparison, the result is readily obtained as

$$X(f) = \frac{1}{2}M(f+f_c) + \frac{1}{2}M(f-f_c)$$

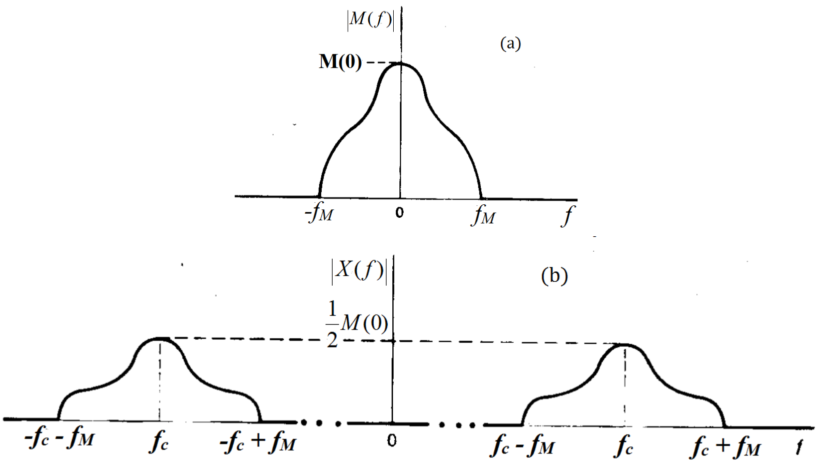


Figure 2.22: Amplitude spectrum of a waveform with no spectra component beyond f_n
The amplitude spectrum of the waveform in (a) multiplied by $cos2\pi f_c t$]

Note:
The relationship of the transform $M(f)$ of $m(t)$ to the transform $V(f)$ of $m(t)\cos 2\pi f_c t$ is illustrated in *Figure* 2.22a. In *Figure* 2.22b the spectral pattern of $M(f)$ is replaced by two patterns of the same form. One is shifted to the right and one to the left, each by amount f_c. Further, the amplitudes of each of these two spectral patterns is one-half the amplitude of the spectral pattern $M(f)$.

Another case of interest is when the waveform $m(t)$ is itself sinusoidal. Thus assume

$$m(t) = A_m \cos 2\pi f_m t$$

where A_m is a constant. $X(f)$ can be found as

$$X(f) = \frac{A_m}{4}\delta(f+f_c+f_m) + \frac{A_m}{4}\delta(f-f_c+f_m) + \frac{A_m}{4}\delta(f-f_c+f_m) + \frac{A_m}{4}\delta(f-f_c-f_m)$$

The amplitude spectrum of the product waveform $m(t) = A_m \cos 2\pi f_m t$ is

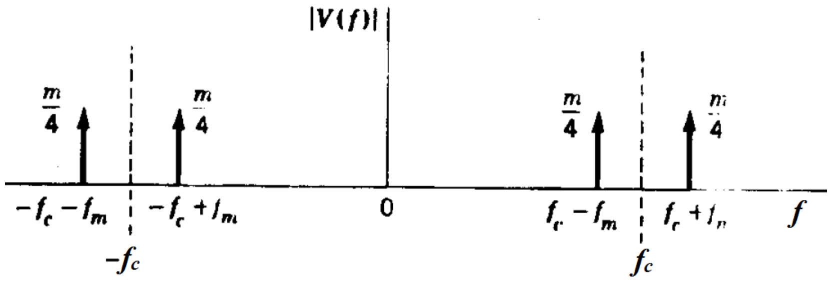


Figure 2.23: Two-sided amplitude spectrum

The waveform is given by

$$\begin{aligned} x(t) &= \frac{A_m}{4}\left[e^{j2\pi(f_c+f_m)t} + e^{-j2\pi(f_c+f_m)t}\right] + \frac{A_m}{4}\left[e^{j2\pi(f_c-f_m)t} + e^{-j2\pi(f_c-f_m)t}\right] \\ &= \frac{A_m}{4}\left[\cos 2\pi\left(f_c+f_m\right)t + \cos 2\pi\left(f_c-f_m\right)t\right] \end{aligned}$$

2.7 Spectral Densities

The spectral density of a signal characterizes the distribution of the signal's energy or power in the frequency domain. This concept is particularly important when considering filtering in communication systems.

2.7.1 Energy Spectral Density

The total energy of energy signal $x(t)$ has been defined in *Section* 2.1.3.5 over the interval $(-\infty,+\infty)$ (*Eqn.* 2.1.14). With the help of Parseval's theorem, (*Subsection* 2.4.2), the energy of the signal expressed in time domain can be related to the energy expressed in frequency domain (Using Fourier transform) as

$$E_x = \int_{-\infty}^{\infty} x^2(t)dt = \int_{-\infty}^{\infty} |X(f)|^2 df \tag{2.7.1}$$

where $X(f)$ is the Fourier transform of the signal $x(t)$, in this case it is non-periodic.
Let ψ_{x} be the squared magnitude spectrum as

$$\psi_{\mathrm{x}} = |X(f)|^2 \tag{2.7.2}$$

ψ_{x} is the **energy spectral density (ESD)** of $x(t)$. Therefore the total energy of $x(t)$ by way of integration with respect to frequency is

$$E_x = \int_{-\infty}^{\infty} \psi_{\mathrm{x}}(f)df \tag{2.7.3}$$

This equation states that the energy of a signal is equal to the area under the curve versus frequency curve. Energy spectral density describes the signal energy per unit bandwidth measured in joules/hertz. There are equal energy contributions from both positive and negative frequency components, since for a real signal, $x(t)$, $|X(f)|$ is an even function of frequency. Therefore, the energy spectral density is symmetrical in frequency about the origin, and thus, the total energy of the signal $x(t)$ is expressed as

$$E_x = 2\int_{0}^{\infty} \psi_{\mathrm{x}}(f)df \tag{2.7.4}$$

2.7.2 Power Spectral Density

Preceding from *Section* 2.1.3.5 *Eqn.* 2.1.15 and *Section* 2.4.2, the **power spectral density**, $S_x(f)$ of a periodic signal $x(t)$ is given in frequency domain as

$$S_x(f) = \sum_{n=-\infty}^{\infty} |C_n|^2 \delta(f - nf_o) \tag{2.7.5}$$

Eqn. 2.7.5 describes the PSD of periodic (power) signals only.
Now, the average power of a signal is

$$P_x = \int_{-\infty}^{\infty} S_x(f) df = 2\int_{0}^{\infty} S_x(f) df \tag{2.7.6}$$

The average normalized power using Parseval's theorem becomes

$$P_x = \lim_{T\to\infty} \frac{1}{T} \int_{-\infty}^{\infty} |X_T(f)|^2 df = \int_{-\infty}^{\infty} \left(\lim_{T\to\infty} \frac{|X_T(f)|^2}{T} \right) df \tag{2.7.7}$$

where $X_T(f) = \mathscr{F}[X_T(f)]$. The unit of the RHS is watts/hertz (or volts^2/hertz or amperes^2/hertz, as appropriate). PSD of non-periodic finite energy signal, $x(t)$ within the interval $(-T/2, T/2)$ is

$$S_x(f) = \lim_{T\to\infty} \frac{1}{T} |X_T(f)|^2 \tag{2.7.8}$$

where $X_T(f)$ is the Fourier transform of the signal in the interval taken. $S_x(f)$ unit is W/Hz.

2.8 Correlation of Signals

The correlation between signals is a measure of the similarity or relatedness between waveforms or signals. Suppose that we have two waveforms $x_1(t)$ and $x_2(t)$, not necessarily periodic nor confined to a finite time interval. Then correlation between them or more precisely the average cross correlation between $x_1(t)$ and $x_2(t)$ is $R_{12}(\tau)$ is defined as

$$R_{12}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_1(t) x_2(t+\tau) dt \tag{2.8.1}$$

If $x_1(t)$ and $x_2(t)$ are periodic with the same fundamental period T_o, then the average cross correlation is

$$R_{12}(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} x_1(t) x_2(t+\tau) dt \tag{2.8.2}$$

If $x_1(t)$ and $x_2(t)$ are signals of finite energy (for example, non-periodic pulse type waveforms), then the cross correlation is defined as

$$R_{12}(\tau) = \int_{-\infty}^{\infty} x_1(t) x_2(t+\tau) dt \tag{2.8.3}$$

The need for introducing the parameter τ in the definition of cross correlation may be seen by the example illustrated in *Figure* 2.24. Here the two waveforms. while different, are obviously related. The waveforms have the same period and are nearly of the same form. However, the integral of the product $x_1(t) \cdot x_2(t)$ is zero since at all time, one or the other function is zero. The function $x_2(t+\tau)$ is the function $x_2(t)$ shifted to the left by amount τ.

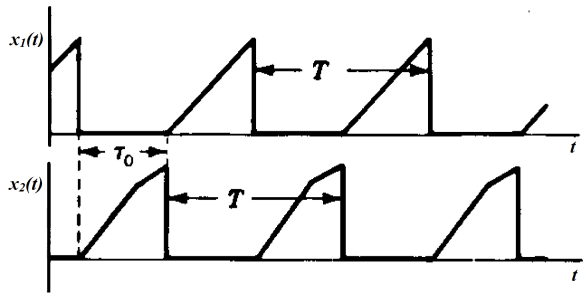


Figure 2.24: Two related signal

It is clear from *Figure* 2.24 that, while $R_{12}(0) = 0$, $R_{12}(\tau)$ will increase as τ increases from zero, becoming a maximum when $\tau = \tau_o$. Functions for which $R_{12}(0) = 0$ for all τ are classified as uncorrelated or *non-coherent*.

2.8.1 Autocorrelation

As described earlier, correlation is a matching process; **autocorrelation** refers to the matching of a signal with a delayed version of itself.

Autocorrelation of Energy Signal
The autocorrelation function of a real-valued energy signal, $x(t)$ is defined as

$$R_x(\tau) = \int_{-\infty}^{\infty} x(t)x(t+\tau)dt, \qquad \text{for} -\infty < \tau < \infty \tag{2.8.4}$$

The autocorrelation function $R_x(\tau)$ provides a measure of how closely the signal matches a copy of itself as the copy is shifted t units in time. The variable τ plays the role of a scanning or searching parameter. $R_x(\tau)$ is not a function of time; it is only a function of the time difference t between the waveform and its shifted copy.

The auto-correlative function of a real-valued **energy signal** has the following properties:

1. $R_x(\tau) = R_x(-\tau)$	symmetrical in τ about zero
2. $R_x(\tau) \leq R_x(0)$ for all τ	maximum value occurs at the origin
3. $R_x(\tau) \leftrightarrow \psi_x(f)$	autocorrelation and ESD from Fourier transform pair
4. $R_x(0) = \int_{-\infty}^{\infty} x^2(t)dt$	value at origin is equal to the energy of the signal

If point 1 through 3 are satisfied, $R_x(\tau)$ satisfies the properties of an autocorrelation function. Point 4 gives the energy content of $x(t)$ by setting $\tau = 0$. Using that the energy of a signal in frequency domain is

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df \tag{2.8.5}$$

$\psi_x(f) = \mathscr{F}[R_x(\tau)] = |X(f)|^2$ is called *energy spectral density* of $x(t)$, and represents the amount of energy per hertz of bandwidth present in the signal at various frequencies.

Autocorrelation of Power Signal
The autocorrelation function of $x(t)$ is defined as

$$R_x(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)x(t+\tau)dt, \qquad \text{for} -\infty < \tau < \infty \tag{2.8.6}$$

When the power signal $x(t)$ is periodic with period T_o, the time average in *Eqn.* 2.8.6 can be taken over a single period T_o, and the autocorrelation function is expressed as

$$R_x(\tau) = \frac{1}{T}\int_{-T_o/2}^{T_o/2} x(t)x(t+\tau)dt, \qquad \text{for} -\infty < \tau < \infty \tag{2.8.7}$$

The autocorrelation function of a real-valued periodic signal has properties similar to those of an energy signal:

1. $R_x(\tau) = R_x(-\tau)$	symmetrical in τ about zero
2. $R_x(\tau) \leq R_x(0)$ for all τ	maximum value occurs at the origin
3. $R_x(\tau) \leftrightarrow S_x(f)$	autocorrelation and PSD from Fourier transform pair
4. $R_x(0) = \frac{1}{T_o}\int_{-T_o/2}^{T_o/2} x^2(t)dt$	value at origin is equal to the average power of the signal

In summary, the PSD can be evaluated by either of the following two methods:

I. *Direct method*, by using Eqn 2.7.8

II. *Indirect method*, by first evaluating the autocorrelation function and then taking the Fourier transform: $S_x(f) = \mathscr{F}[R_x(\tau)]$.

Finally, the total average normalized power for the waveform $x(t)$ can be evaluated by

$$P_x = \langle x^2(t)\rangle = X_{rms}^2 = \int_{-\infty}^{\infty} S_x(f)df = R_x(0) \tag{2.8.8}$$

We define $S_x(f)$, the *power spectral density* of the signal $x(t)$ to be the Fourier transform of the time average autocorrelation function

$$S_x(f) = \mathscr{F}[R_x(\tau)] \tag{2.8.9}$$

Example 2.12

Find the PSD of $w(t) = A\sin\omega_o t$.

Solution
The PSD will be calculated using the indirect method. The autocorrelation is

$$\begin{aligned} R_w(\tau) &= \langle w(t)w(t+\tau)\rangle \\ &= \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} A^2 \sin\omega_o t \sin\omega_o(t+\tau)dt \end{aligned}$$

By trigonometric identity

$$R_w(\tau) = \frac{A^2}{2}\cos\omega_o\tau \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} dt - \frac{A^2}{2}\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\cos(2\omega_o t + \omega_o\tau)\,dt$$

This reduces to after the integration

$$R_w(\tau) = \frac{A^2}{2} \cos \omega_o \tau$$

The PSD is then

$$P_w(f) = \mathscr{F}\left[\frac{A^2}{2} \cos \omega_o \tau\right] = \frac{A^2}{4}\left[\delta(f - f_o) + \delta(f + f_o)\right]$$

A plot of the PSD gives

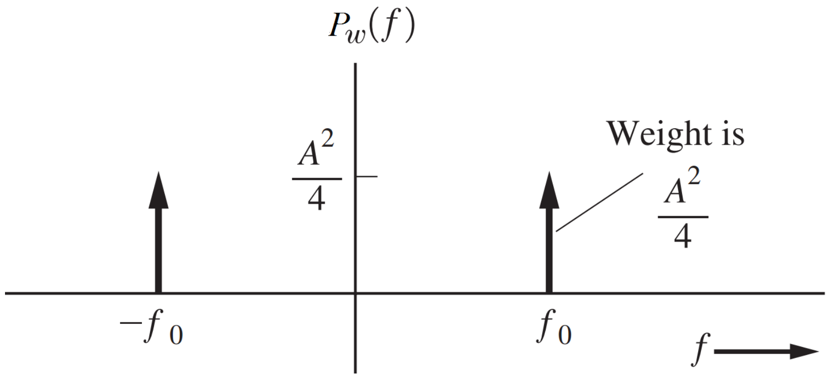


Figure 2.25: Power spectrum of a sinusoid

The average normalized power is obtained as

$$P_w = \int_{-\infty}^{\infty} \frac{A^2}{4}\left[\delta(f - f_o) + \delta(f + f_o)\right] df = \frac{A^2}{2}$$

2.9 Review Questions

1. A pulse of amplitude A extends from $t = -\tau/2$ to $t = +\tau/2$.
   Find its Fourier transform $V(f)$. Consider also the Fourier series for a periodic sequence of such pulses separated by intervals T_o. Compare the Fourier series coefficients V_n with the transform in the limit as $T_o \to \infty$.

2. (i) Find the Fourier transform of $\delta(t)$, an impulse of unit strength.

   (ii) Given a network whose transfer function is $H(f)$. An impulse $\delta(t)$ is applied at the input. Show that the response $v_o(t) = h(t)$ at the output is the inverse transform of $H(f)$, that is, show that $h(t) = \mathscr{F}^{-1}[H(f)]$.

3. Find the Fourier transform of

   $$s(t) = e^{\alpha t} u(t)$$

   for some real number α.

4. For a sinusoidal waveform with a peak value of A and frequency of f_o, use the time average operator to show that the RMS value for the waveform is $A/\sqrt{2}$

5. The voltage across a 50 Ω resistive load is the positive portion of the cosine wave. That is,

$$v(t) = \begin{cases} 10\cos\omega_o t, & |t - nT_o| < T_o/4 \\ 0, & t \text{ elsewhere} \end{cases}$$

where n is any integer.

   (a) Sketch the voltage and current waveforms
   (b) Evaluate the DC values for the voltage and current
   (c) Find the RMS values for the voltage and current
   (d) Find the total average power dissipated in the load

6. Show that if $x_T(t)$ denotes the truncated signal corresponding to the power-type signal $x(t)$; that is

$$x_T(t) = \begin{cases} x(t), & -\frac{T}{2} < t \leq \frac{T}{2} \\ 0, & \text{otherwise} \end{cases}$$

and if $S_{x_T}(f)$ denotes the energy spectral density of $x_T(f)$, then $S_x(f)$, the power spectral density of $x(t)$, can be expressed as

$$S_x(f) = \lim_{T\to\infty} \frac{S_{x_T}(f)}{T}$$

7. An average-reading power meter is connected to the output circuit of a transmitter. The transmitter output is fed into a 50 Ω resistive load, and the watt-meter reads 25 W.

   a. What is the power in dBm units?
   b. What is the power in dBk units?
   c. What is the value in dBmV units?

8. A cable company provides a 0 dBmV signal at 843 MHz via a 75 Ω coaxial cable to a cable modem which is located in a residence. This modem uses this downstream signal to produce the downloaded data signal for the Internet subscriber's computer via an Ethernet connection. (The up-link signal that is sent by the modem is on 33.5 MHz.) What is the power of the 843 MHz signal that is supplied to the modem if the modem acts as a 75 Ω resistive load to this downstream signal?

2.10 Textbooks and References

[1] Principles of Communication Systems by Taub and Schilling,2nd Edition. McGraw Hill.

[2] Communication Systems by Siman Haykin,4th Edition, John Wiley and Sons Inc.

[3] Modern digital and analog communication system, by B. P. Lathi, 3rd Edition, Oxford University Press.

[4] Digital and analog communication systems, by L. W. Couch, 6th Edition, Pearson Education, Pvt. Ltd

[5] Introduction to Analog and Digital Communications, Second Edition, by Simon Haykin and Michael Moher

Part II

Analog Signal Transmission and Reception

Chapter 3

Fundamentals of Modulation Schemes

***

> It's always helpful to learn from your mistakes because then your mistakes seem worthwhile.
>
> – Garry Marshall

3.1 Modulation Overview

Before an information-bearing signal is transmitted through a communication channel, some type of modulation process is typically utilized to produce a signal that can easily be accommodated by the channel. The modulation process commonly translates an information-bearing signal, usually referred to as the message signal, to a new spectral location depending upon the intended frequency for transmission.

If more than one signal utilizes a channel, modulation allows translation of different signals to different spectral locations, thus allowing the receiver to select the desired signal. Multiplexing allows two or more message signals to be transmitted by a single transmitter and received by a single receiver simultaneously. The logical choice of a modulation technique for a specific application is influenced by the characteristics of the message signal, the channel, the performance desired from the overall communication system, the use of the transmitted data, and the economic factors that are always important in practical applications.

3.1.1 Baseband Signals

Although digital transmission can be made up of signal that originated in digital form, such as computer data; analog signals can be converted into digital form and then transmitted. Regardless of whether the original information signals are analog or digital, they are all referred to as "baseband signals". In a communication system, the original information signals (baseband signals) may be transmitted over the medium. Putting the original signal directly into the medium is referred to as baseband transmission. The common example is telephony, especially for the local calls. Here the voice signal, converted into electrical form, is placed on the wires and transmitted over some distance to the receiver. In some computer networks, the digital signals are applied directly to coaxial cables for transmission to another computer.

3.1.2 Limitations of Baseband Transmission

There arc many instances when the baseband signals arc incompatible for direct transmission over the medium. For example, voice signals cannot travel longer distances in air, the signal gets attenuated rapidly. Hence for transmission of baseband signals by radio, modulation technique has to be used.

3.1.3 Modulation Techniques

In the modulation process baseband signal (such as voice, video, etc.) modifies another higher-frequency signal called the carrier. The carrier is usually a sine wave that is higher in frequency than the highest baseband signal frequency. The baseband signal modifies the amplitude or frequency or phase of the carrier in the modulation process.
The message signal is called a **modulating signal**. Modulation converts message signal $m(t)$ from lowpass to bandpass, in the neighborhood of the carrier (centre) frequency f_c. For a sinusoidal carrier, a general modulated carrier can be represented as

$$c(t) = A(t)cos[2\pi f_c t + \phi(t)] \tag{3.1.1}$$

where f_c is the carrier frequency. Sometimes $c(t)$ is replaced with $x_c(t)$.
The component of the carrier signal that can be varied in the modulation process is the instantaneous amplitude $A(t)$ and Phase deviation $\phi(t)$. When the amplitude $A(t)$ is linearly related to the modulating signal, the result is **linear modulation**. E.g. Amplitude Modulation. Now, letting $\phi(t)$ or the time derivative of $\phi(t)$ be linearly related to the modulating signal yields phase or frequency modulation, respectively. Collectively, phase and frequency modulation are referred to as angle modulation, since the instantaneous phase angle of the modulated carrier conveys the information.

Modulation techniques are classified according which parameter of the carrier is changed as described below:

- **Amplitude modulation**: tn amplitude modulation, the amplitude of the carrier is varied according to the baseband signal keeping its frequency and phase constant.
- **Frequency modulation**: In frequency modulation, the frequency of the carrier is varied according to the baseband signal.
- **Phase modulation**: in phase modulation, the phase of the carrier is varied according to the baseband signal

3.1.4 Importance of Modulation

We have seen that baseband signals are incompatible for direct transmission over the medium and therefore we have to use modulation technique for the communication of baseband signal. The advantages of using modulation technique are as given below:

3.1.4.1 Reduces antenna height

The height of an antenna required for transmission and reception of radio waves in radio transmission is a function of wavelength of the frequency used. The minimum height of the antenna is given as $\lambda/4$. The wavelength λ is given as

$$\lambda = \frac{c}{f} \tag{3.1.2}$$

where c is the velocity of light, f is the frequency

From the above equation it can be noticed easily that, at low frequencies wavelength is very high and hence the antenna height. For example, consider the baseband signal with $f = 15$ kHz, Then

$$\begin{aligned} \text{Heigth of antenna } &= \frac{\lambda}{4} = \frac{c}{f \times 4} \\ &= \frac{3 \times 10^8}{15 \times 10^6 \times 4} = 5 \text{ m} \end{aligned}$$

This height of antenna is practical and such antenna can be installed.

3.1.4.2 Avoid mixing of signals

All sound signals are concentrated within the range from 20 Hz to 20 kHz. The transmission of baseband signals from various sources causes the mixing of signal and then it is difficult to separate at the receiver end. In order to separate the various signals, it is necessary to translate them all to different portions of the electromagnetic spectrum (channel); each must be given its own bandwidth commonly known as channel bandwidth. This can be achieved by taking different carrier frequency for different signal source as shown in the *Figure* 3.1. Once the

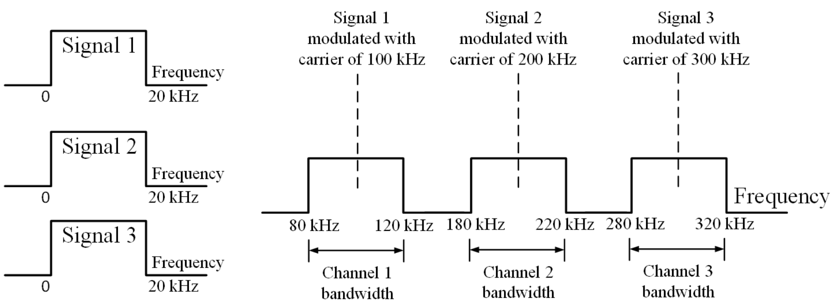
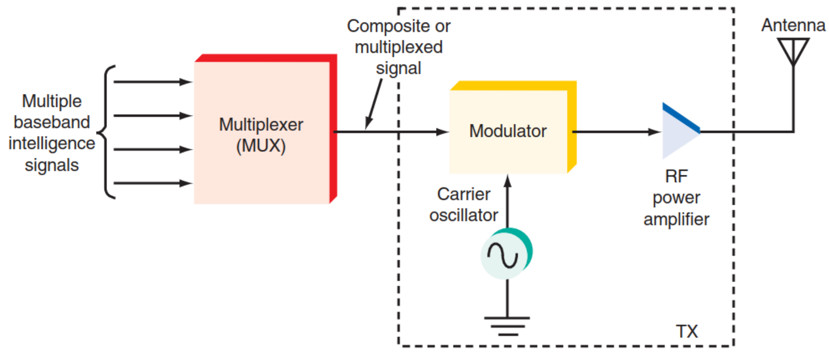


Figure 3.1: Modulation avoid mixing of signals

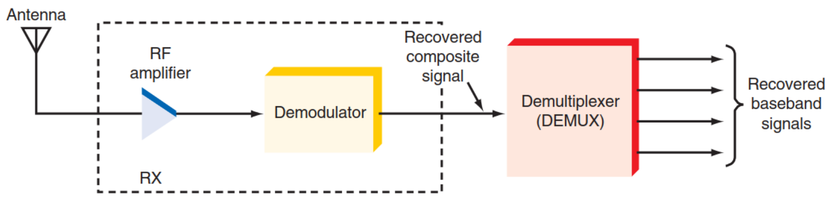
signals have been transmitted, a tuned circuit at the receiver end selects the portion of the electromagnetic spectrum it is tuned for. Therefore modulating different signal sources by different carrier frequencies avoid mixing of signals

3.1.4.3 Allows multiplexing of signals

The modulation permits multiplexing to be used. Multiplexing means transmission of two or more signals simultaneously over the same channel, *Figure* 3.2. The common examples of multiplexing are the number of television channels operating simultaneously or number of radio stations broadcasting the signal in MW and SW band, simultaneously.



(a) Multiplexing at the transmitter



(b) De-multiplexing at the receiver

Figure 3.2: Multiplexing of signals

Signals from different stations can be separated in the receiver since the carrier frequencies for these signals are different. It is commonly known as tuning the receiver to the desired station. By tuning process, the desired signal is selected and at the same time, other unwanted signals are rejected.

There are three basic types of multiplexing:

I. Frequency-division multiplexing the intelligence signals modulate sub-carriers on different frequencies that are then added together, and the composite signal is used to modulate the carrier. In optical networking, wavelength division multiplexing (WDM) is equivalent to frequency-division multiplexing for optical signal

II. Time-division multiplexing he multiple intelligence signals are sequentially sampled, and a small piece of each is used to modulate the carrier. If the information signals are sampled fast enough, sufficient details are transmitted that at the receiving end the signal can be reconstructed with great accuracy.

III. Code-division multiplexing the signals to be transmitted are converted to digital data that is then uniquely coded with a faster binary code. The signals modulate a carrier on the same frequency. All use the same communications channel simultaneously. The unique coding is used at the receiver to select the desired signal.

Example of multiplexing systems are:

- Number of TV channels operating simultaneously
- Number of radio stations broadcasting the signals in MW and SW band simultaneously

3.1.4.4 Increases the range of communication

At low frequencies radiation is poor and signal gets highly attenuated. Therefore baseband signals cannot be transmitted directly over long distance. Modulation effectively increases the frequency of the signal to be radiated and thus, increases the distance over which signals can be transmitted faithfully.

3.1.4.5 Allows adjustment in the bandwidth

Bandwidth of a modulated signal may be made smaller or larger than the original signal. Signal to noise ratio in the receiver which is a function of the signal bandwidth can thus be improved by proper control of the bandwidth at the modulating stage.

3.1.4.6 Improves quality of reception

The signal communication using modulation techniques such a frequency modulation, pulse code modulation reduce the effect of noise to great extent. Reduction in noise improves the quality of reception

3.2 Basic Techniques

Lets denote the analog signal as $m(t)$, that is assumed to be lowpass signal and power-type signal of bandwidth W, in other words, $M(f) \equiv 0$, for $|f| > W$. The power of $m(t)$, P_m is given by

$$P_m = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |m(t)|^2 dt$$

as before.

The message signal is transmitted through the communication channel by impressing it on a carrier signal given by

$$c(t) = A_c \cos(2\pi f_c t + \phi_c) \tag{3.2.1}$$

where A_c is the carrier amplitude, f_c is the carrier frequency, and ϕ_c is the carrier phase.
We say that the message signal $m(t)$ modulates the carrier signal $c(t)$ in either amplitude, frequency, or phase, if after modulation, the amplitude, frequency, or phase of the signal become functions of the message signal. In effect, modulation converts the message signal $m(t)$ from lowpass to bandpass, in the neighborhood of the center frequency f_c.

The next chapters deals with all the basic modulation techniques.

3.3 Frequency Translation

Suppose that a signal is bandlimited to the frequency range extending from a frequency f_1 to f_2. Frequency translation is used to replace the signal with a new signal whose spectral range extends from f'_1 to f'_2 that has a recoverable form of the same information as was within the original.

A signal may be translated to a new spectral position by multiplying it by an auxiliary sinusoidal signal. Let's take a signal

$$v_m(t) = A_m \cos \omega_m t = A_m \cos(2\pi f_m t) \tag{3.3.1}$$

Using Euler's theorem

$$v_m(t) = \frac{A_m}{2}\left(e^{j\omega_m t} + e^{-j\omega_m t}\right) = \frac{A_m}{2}\left(e^{j2\pi f_m t} + e^{-j2\pi f_m t}\right) \tag{3.3.2}$$

The two sided spectral amplitude pattern is as shown *Figure* 3.3a

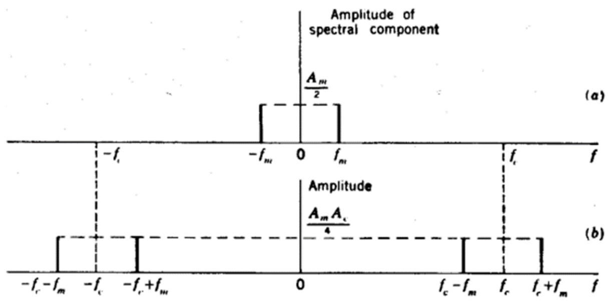


Figure 3.3: Two-sided spectral amplitude

Let's multiply $v_m(t)$ with an auxiliary signal $v_c(t)$

$$\begin{aligned} v_m(t)v_c(t) &= \frac{A_m A_c}{2}\left[\cos(\omega_c + \omega_m)t + \cos(\omega_c - \omega_m)t\right] \\ &= \frac{A_m A_c}{4}\left[e^{j(\omega_m+\omega_c)t} + e^{-j(\omega_m+\omega_c)t} + e^{j(\omega_c-\omega_m)t} + e^{-j(\omega_c-\omega_m)t}\right] \end{aligned} \tag{3.3.3}$$

The new spectral amplitude pattern is shown in *Figure* 3.3(b). Observe that the two original spectral lines have been translated, both in the positive-frequency direction by amount f_c and also in the negative-frequency direction by the same amount. There are now four spectral components resulting in two sinusoidal waveforms, one of frequency $f_c + f_m$ and the other of frequency $f_c - f_m$ with amplitude $\frac{A_m A_c}{2}$ each.

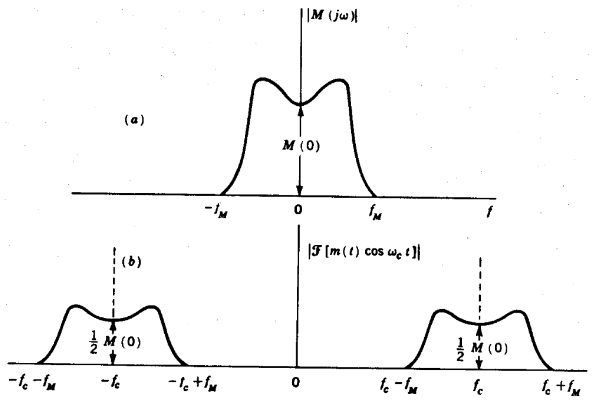


Figure 3.4: (a) Spectral density $|M(j\omega)|$ of a nonperiodic signal $m(t)$. (b) Spectral density of $m(t)\cos 2\pi f_c t$.

3.4 Recovery of Baseband Signal

Recovery of original signal $m(t)$ from a translated version may be done by reverse translation, which is accomplished by simply multiplying the translated signal with $\cos\omega_c t$

$$\begin{aligned}[m(t)\cos\omega_c t]\cos\omega_c t &= m(t)\cos^2\omega_c t \\ &= m(t)\left(\frac{1}{2}+\frac{1}{2}\cos 2\omega_c t\right) \\ &= \frac{m(t)}{2}+\frac{m(t)}{2}\cos 2\omega_c t\end{aligned} \tag{3.4.1}$$

The baseband signal $m(t)$ reappears. The higher order frequency (double-freq. signal) which has $f_c \gg f_m$ is easily removed by a LPF.

This method of signal recovery, for all its simplicity, is beset by an important inconvenience when applied in a physical communication system. Suppose that the auxiliary signal used for recovery differs in phase from the auxiliary signal used in the initial translation. If this phase angle is θ, the recovered baseband waveform will be proportional to $m(t)cos\theta$. Therefore, unless it is possible to maintain $\theta = 0$, the signal strength at recovery will suffer. If it should happen that $\theta = \pi/2$, the signal will be lost entirely. Or consider, for example, that θ drifts back and forth with time. Then in this case the signal strength will wax and wane, in addition, possibly, to disappearing entirely from time to time.

Alternatively, suppose that the recovery auxiliary signal is not precisely at frequency f_c but is instead at $f_c+\Delta f$. In this case, the recovered baseband signal will be proportional to $m(t)\cos 2\pi\Delta f$, resulting in a signal which will wax and wane or even be entirely unacceptable if Δf is comparable to f_m or larger than the frequencies present in the baseband signal.

It is noted, therefore, that signal recovery using a second multiplication requires that there should be available at the recovery point a signal which is precisely synchronous with the corresponding auxiliary signal at the point of the first multiplication. In such a synchronous or coherent system a fixed initial phase discrepancy is of no consequence since a simple phase shifter will correct the matter. What is essential is that, in any time interval, the number of cycles executed by the two auxiliary-signal sources be the same. Of course, in a physical system, where some signal distortion is tolerable, some lack of synchronism may be allowed.

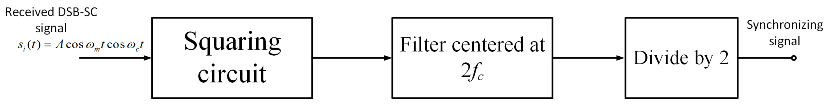


Figure 3.5: Simple squaring synchronizer

When the use of a common auxiliary signal is not feasible, it is necessary to resort to rather complicated means to provide a synchronous auxiliary signal at the location of the receiver. One commonly employed scheme is indicated in *Figure* 3.5. To illustrate the operation of the synchronizer, we can assume the baseband signal is a sinusoidal $\cos\omega_m t$. The received signal is $s_i(t) = A\cos\omega_m t\cos\omega_c t$, with A, a constant amplitude. The signal $s_i(t)$ does not have a spectral component at the angular frequency, ω_c, The output of the squaring circuit is

$$\begin{aligned} s_i^2(t) &= A^2\cos^2\omega_m t\cos^2\omega_c t \\ &= A^2\left(\frac{1}{2}+\frac{1}{2}\cos 2\omega_m t\right)\left(\frac{1}{2}+\frac{1}{2}\cos 2\omega_c t\right) \\ &= \frac{A^2}{4}\left[1+\frac{1}{2}\cos 2\left(\omega_c+\omega_m\right)t+\frac{1}{2}\cos 2\left(\omega_c-\omega_m\right)t+\cos 2\omega_m t+\cos 2\omega_c t\right] \end{aligned} \tag{3.4.2}$$

The filter selects the spectral component $\left(\frac{A^2}{2}\right)\cos 2\omega_c t$, which is then applied to a circuit which divides the frequency by a factor of 2.

Chapter 4

Amplitude (Linear) Modulation Fundamentals

***

There are too many people praying for mountains of difficulty to be removed, when what they really need is the courage to climb them!

– Unknown

A general linearly modulated carrier is represented by setting the instantaneous phase deviation ϕ in equal to zero in *Eqn.* 3.1.3. Thus, a linearly modulated carrier is represented by

$$x_c(t) = A_c m(t) \cos(2\pi f_c t) \tag{4.0.1}$$

in which the carrier amplitude $A(t)$ varies in one-to-one correspondence with the message signal.

4.1 Conventional Amplitude Modulation

A conventional AM signal consists of a large carrier component in addition to the double-sideband AM modulated signal.

The message $m(t)$, is transmitted through the communication channel by impressing it on a carrier or bandpass signal of the form

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

For AM, it is a requirement to set $\phi_c = 0$ giving the the modulated wave as

$$s(t) = A_c [1 + K_a m(t)] \cos(2\pi f_c t) \tag{4.1.1}$$

where $K_a = \frac{1}{A_c}$ is amplitude sensitivity of the modulator responsible for generating the modulated signal $s(t)$.

Quick Pick 4 - 1: Conventional AM

The transmitted signal of AM is mathematically expressed as

$$s(t) = A_c\left[1 + \mu m(t)\right]\cos\left(2\pi f_c t\right)$$

where the message waveform is contrained to satisfy $|\mu| \leq 1$ for $f_c \gg W$. Note that

$$A_c m(t)\cos\left(2\pi f_c t + \phi_c\right)$$

is a double-sideband AM signal.

As long as $|\mu| \leq 1$, the amplitude $A_c\left[1 + m(t)\right]$ is always positive. This is the reason why conventional AM is easy to demodulate. A 100 % modulation is when $\mu = 1$, the envelop varies between $A_{min} = 0$ and $A_{max} = 2A_c$

The information-bearing signal or message signal is denoted by $m(t)$; the terms "information-bearing signal" and "message signal" are used interchangeably throughout this book. Amplitude modulation (AM) is formally defined as a process in which the amplitude of the carrier wave is varied about a mean value, linearly with the message signal $m(t)$.

Typically, the carrier amplitude A_c and the message $m(t)$ are measured in volts, in that case, K_a is in V^{-1}. *Figure* 4.1 shows the AM wave and its components.
Let us define **modulation index** in terms of amplitude sensitivity as

$$\mu = K_a A_m \tag{4.1.2}$$

where A_m is the amplitude of the baseband (message) signal, $m(t)$.

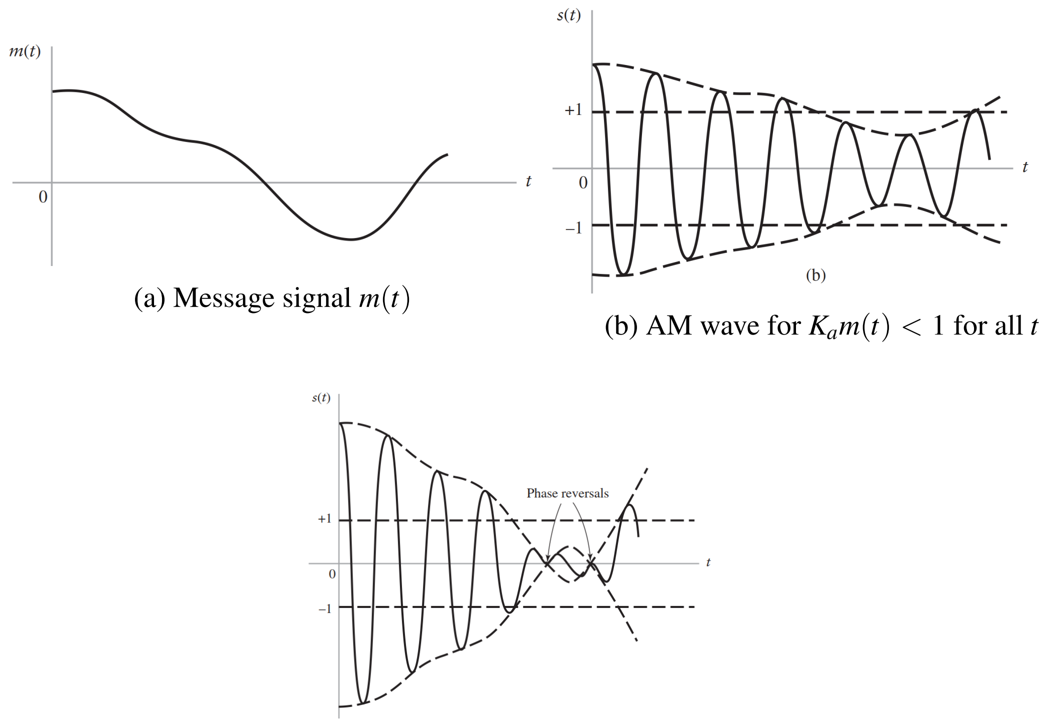
In amplitude modulation, information pertaining to the message signal resides solely on the **envelope**, which is defined by the amplitude of the modulated wave $s(t)$, that is $A_c|1 + K_a m(t)|$. From this, it can be observed that the envelope of $s(t)$ has the same wave form as the message signal provided that two conditions are met:

Quick Pick 4 - 2: Conventional AM

Continuing from **Quick Pick 4 - 1** on the value of μ

On the other hand, if $m(t) < -1$ for some t or $\mu > 1$, the AM signal is said to be **over-modulated** that results in phase reversals and demodulation of such signal is rendered more complex. In practice, $m(t)$ is scaled so that its magnitude is always less than unity or $\mu < 1$.

Another way of finding the modulation index is $\mu = m = \dfrac{A_m}{A_c}$. It is good to know that distortion caused by over-modulation also produces adjacent **channel interference**.



(a) Message signal $m(t)$

(b) AM wave for $K_a m(t) < 1$ for all t

(c) AM wave for $|K_a m(t)| > 1$ for some t

Figure 4.1: Illustration of amplitude modulation process

1. The modulation of $K_a m(t)$ is always less than unity, Figure 4.1b

$$|K_a m(t)| < 1, \quad \text{for all } t \text{ or } \mu < 1 \tag{4.1.3}$$

   When K_a is made larger to make $|K_a m(t)| < 1$ for any t, the carrier wave becomes **over modulated**, resulting in carrier *phase reversal* whenever $1 + K_a m(t)$ crosses zero. The modulated wave then exhibits envelope distortion, as in *Figure* 4.1c. It is, therefore, apparent that by avoiding over-modulation, a one-to-one relationship is maintained between the envelope of the AM wave and the modulating wave for all values of time. The absolute maximum value of μ multiplied by 100 is referred to as **percentage modulation**.

2. The carrier frequency is much greater than the highest frequency component W of the message signal, that is

$$f_c \gg W \tag{4.1.4}$$

   We call W, the message bandwidth. If the condition of *Eqn.* 4.1.4 is not satisfied, an envelope cannot be visualized (and therefore can't be detected) satisfactorily.

Provided that *Eqn.* 4.1.3 and 4.1.4 are satisfied, demodulation of the AM wave is achieved by using an **envelope detector**, which is defined as a device whose output traces the envelope of the AM wave acting as the input signal.

Conventional AM Wave – Alternative Approach

In amplitude modulation, the amplitude of a carrier signal is varied according to variations in the amplitude of modulating signal.

Let represent the modulating (message) signal as

$$m(t) = A_m \cos \omega_m t$$

and carrier signal as

$$c(t) = A_c \cos \omega_c t$$

Using the above mathematical expressions for modulating and carrier signals, the mathematical expression for amplitude of the complete modulated wave is given as

$$S_{Am} = A_c + m(t)$$

By substitution of $m(t)$ we get

$$S_{Am} = A_c + A_m \cos \omega_m t \tag{4.1.5}$$

Now, the instantaneous modulated wave $s(t)$ is given by

$$\begin{aligned} s(t) &= S_{AM} \cos \omega_c t \\ &= (A_c + A_m \cos \omega_m t) \cos \omega_c t \quad \rightarrow \text{ by substitution of } S_{AM} \\ &= A_c [1 + \mu \cos \omega_m t] \cos \omega_c t, \qquad \mu = \frac{A_m}{A_c} \end{aligned} \tag{4.1.6}$$

Expanding $s(t)$ without factoring A_c, we get

$$s(t) = A_c \cos \omega_c t + A_m \cos \omega_m t \cos \omega_c t \tag{4.1.7}$$

Using the appropriate trig. identities, we get

$$s(t) = \underbrace{A_c \cos (2\pi f_c t)}_{\text{Carrier}} + \underbrace{\frac{A_m}{2} \cos [2\pi (f_c - f_m) t]}_{\text{Lower sideband}} + \underbrace{\frac{A_m}{2} \cos [2\pi (f_c + f_m) t]}_{\text{Upper sideband}} \tag{4.1.8}$$

Now, remember that A_c and A_m are peak values of the carrier and modulating sinusoid waves, respectively. For power calculations, rms values must be used for the voltages.
The rms carrier and sideband voltages are then

$$s(t) = \frac{A_c}{\sqrt{2}} \cos (2\pi f_c t) + \frac{A_m}{2\sqrt{2}} \cos [2\pi (f_c - f_m) t] + \frac{A_m}{2\sqrt{2}} \cos [2\pi (f_c + f_m) t] \tag{4.1.9}$$

Taking R as the resistance of the antenna, the total transmit power is

$$\begin{aligned} P_T &= \frac{\left(A_c / \sqrt{2}\right)^2}{R} + \frac{\left(A_m / 2\sqrt{2}\right)^2}{R} + \frac{\left(A_m / 2\sqrt{2}\right)^2}{R} \\ &= \frac{A_c^2}{2R} + \frac{A_m^2}{8R} + \frac{A_m^2}{8R} \end{aligned} \tag{4.1.10}$$

Expressing A_m in terms of A_c by using the modulation index, *Eqn.* 4.1.10 is expressed as

$$\begin{aligned} P_T &= \frac{A_c^2}{2R} + \frac{(\mu A_c)^2}{8R} + \frac{(\mu A_c)^2}{8R} = \frac{A_c^2}{2R} + \frac{\mu^2 A_c^2}{8R} + \frac{\mu^2 A_c^2}{8R} \\ P_T &= \frac{A_c^2}{2R} \left(1 + \frac{\mu^2}{4} + \frac{\mu^2}{4}\right) \end{aligned} \tag{4.1.11}$$

The final equation is obtained by noting that $P_c = A_c^2 / 2R$, and factoring it out, yields

$$P_T = P_c \left(1 + \frac{\mu^2}{2}\right) \tag{4.1.12}$$

Example 4.13

An AM transmitter has a carrier power of 30 W. The percentage of modulation is 85 percent. Calculate (a) the total power and (b) the power in one sideband.

Solution

a)

$$P_T = P_c\left(1+\frac{\mu^2}{2}\right)$$
$$= 30\left[1+\frac{(0.85)^2}{2}\right] = 40.8 \text{ W}$$

b)

$$P_{SB}(\text{both}) = P_T - P_c = 40.8 - 30 = 10.8 \text{ W}$$

Power of one side only

$$P_{SB}(\text{one}) = \frac{P_{SB}}{2} = 5.4 \text{ W}$$

If we are after the total current, I_T, it can be derived from the total power as

$$I_T = I_c\sqrt{\left(1+\frac{\mu^2}{2}\right)} \tag{4.1.13}$$

The modulation index in terms of current is

$$\mu = \sqrt{2\left[\left(\frac{I_T}{I_c}\right)^2 - 1\right]} \tag{4.1.14}$$

4.1.1 Bandwidth of Conventional AM Signal

Assume that, $m(t)$ is a deterministic signal with Fourier transform, $M(f)$, the spectrum of the amplitude modulated signal $s(t)$ is given as

$$\begin{aligned} S(f) &= \mathscr{F}[K_a m(t)] \star \mathscr{F}[A_c\cos(2\pi f_c t)] + \mathscr{F}[A_c\cos(2\pi f_c t)] \\ &= \frac{A_c}{2}[\delta(f-f_c)+\delta(f+f_c)] + \frac{K_a A_c}{2}[M(f-f_c)+M(f+f_c)] \end{aligned} \tag{4.1.15}$$

It is clearly seen that, the spectrum (*Figure* 4.2) of conventional AM signal occupies a bandwidth twice the bandwidth of the message signal. The AM spectrum consists of carrier-frequency impulses and symmetrical sidebands centered at $\pm f_c$. For positive frequencies, the highest frequency component of the AM wave equals $f_c + W$ and the lowest frequency component equals $f_c - W$. The difference between these two frequencies defines the transmission bandwidth of the AM wave, which is exactly twice the message bandwidth W; that is

$$B_T = 2W \tag{4.1.16}$$

Note that, AM requires twice the bandwidth needed to transmit $x(t)$ at baseband without modulation. Transmission bandwidth is an important consideration for the comparison of modulation systems.

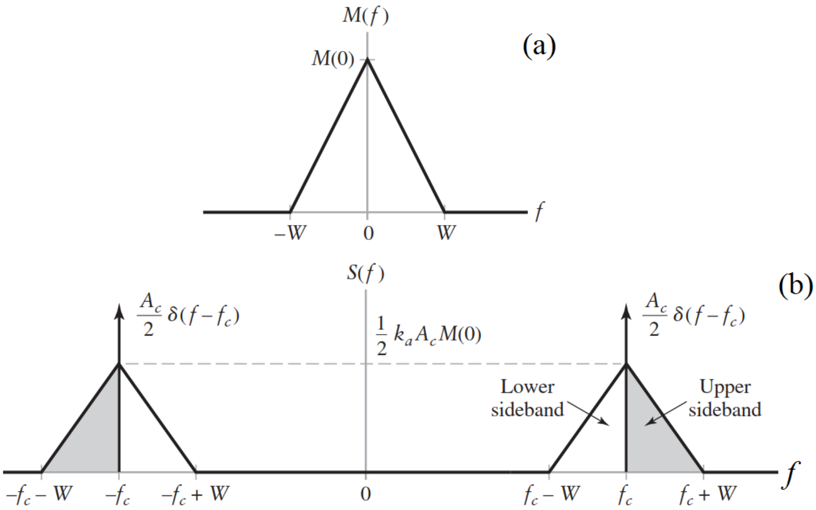


Figure 4.2: (a) Spectrum of message signal $m(t)$ (b) Spectrum of AM wave $s(t)$

4.1.1.1 Single-Tone Modulation

Consider a modulating wave $m(t)$ that consists of a single tone or frequency component given by

$$m(t) = A_m \cos(2\pi f_m t)$$

and carrier signal as

$$c(t) = A_c \cos(2\pi f_c t)$$

The corresponding AM wave is given by

$$s(t) = A_c\left[1 + \mu \cos(2\pi f_m t)\right]\cos(2\pi f_m t) \tag{4.1.17}$$

where $\mu = K_a A_m$ as before and μ is *modulation factor* or modulation index. μ is usually kept below unity to avoid distortion due to over-modulation.

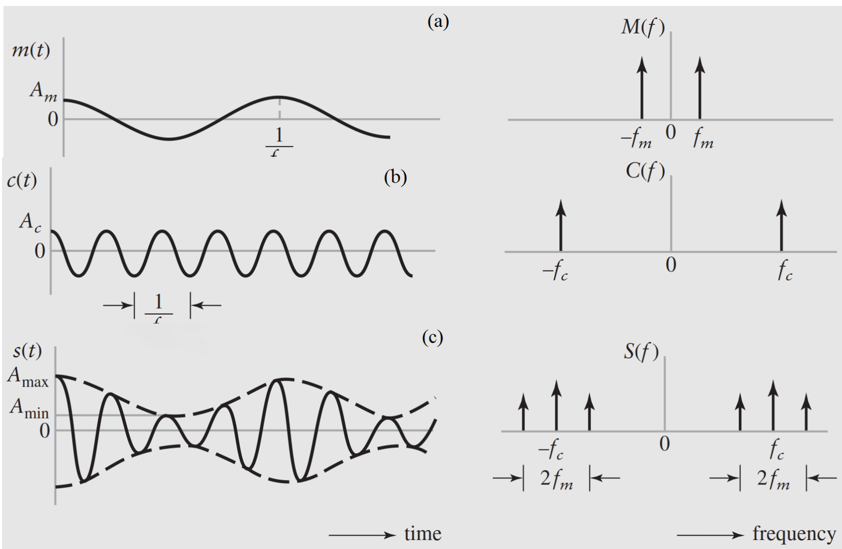


Figure 4.3: Illustration of the time-domain (on the left) and frequency-domain (on the right) characteristics of amplitude modulation produced by a single tone (a) Modulating wave (b) Carrier wave (c) AM wave.

Figure 4.3c shows sketch of $s(t)$ for $\mu < 1$.
Let A_{max} and A_{min} denote the maximum and minimum values of envelope of the modulated wave, respectively. Thus

$$\frac{A_{\max}}{A_{\min}} = \frac{A_c(1+\mu)}{A_c(1-\mu)} \tag{4.1.18}$$

Rearranging, the modulation factor is obtained

$$\mu = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}} \tag{4.1.19}$$

Expanding *Eqn.* 4.1.17 and using the appropriate Trig. identities, the equation can be written in the form

$$s(t) = A_c \cos(2\pi f_c t) + \frac{1}{2}\mu A_c [cos2\pi(f_c + f_m)t] + \frac{1}{2}\mu A_c [cos2\pi(f_c - f_m)t] \tag{4.1.20}$$

Fourier transform of $s(t)$ is

$$\begin{aligned} S(f) &= \frac{1}{2}A_c[\delta(f-f_c) + \delta(f+f_c)] \\ &+ \frac{1}{4}\mu A_c[\delta(f-f_c-f_m) + \delta(f-f_c-f_m)] \\ &+ \frac{1}{4}\mu A_c[\delta(f-f_c+f_m) + \delta(f+f_c-f_m)] \end{aligned} \tag{4.1.21}$$

Thus, the spectrum of an AM wave consists of delta functions at $\pm f_c$, $f_c \pm f_m$, and $-f_c \pm f_m$, as shown in *Figure* 4.3c.

4.1.2 Power Content of Conventional AM Signal

Conventional AM signal is similar to DSB, when $m(t)$ is substituted with $1 + K_a m(t)$. Later, we will see that in the DSB-SC case, the average (transmitted) power of the modulated signal is

$$P_T \triangleq \langle s^2(t) \rangle \tag{4.1.22}$$

Using *Eqn.* 4.1.17 and expanding after squaring

$$P_T = \frac{1}{2}A_c^2 \langle 1 + 2\mu m(t) + \mu^2 m^2(t) \rangle + \frac{1}{2}A_c^2 \left\langle [1 + \mu m(t)]^2 \cos 2\omega_c t \right\rangle \tag{4.1.23}$$

The second term average to zero under condition $f_c \gg W$. Thus, if $\langle m(t) \rangle = 0$ and $\langle m^2(t) \rangle = P_m$ then

$$P_T = \frac{1}{2}A_c^2 \left(1 + \mu^2 P_m\right) \tag{4.1.24}$$

It is assumed here that the average of $m(t)$ is zero. Conventional AM is not practical for transmitting signals with significant low-frequency content.
Eqn. 4.1.24 can be put in the form

$$P_T \simeq P_c + 2P_{sb} \tag{4.1.25}$$

where

$$P_c = \frac{1}{2}A_c^2, \qquad P_{sb} = \frac{1}{4}A_c^2 \mu^2 P_m = \frac{1}{2}\mu^2 P_m P_c \tag{4.1.26}$$

The term P_c represents the **unmodulated carrier power**, since $P_T = P_c$ when $\mu = 0$; the term P_{sb} represents the power per sideband since, when $\mu \neq 0$, P_T consists of the power in the carrier plus two symmetric sidebands. The modulated constraint $|\mu m(t)| \leq 1$ requires that $\mu^2 P_m \leq 1$, so $P_{sb} \leq \frac{1}{2}P_c$ and

$$P_c = P_T - 2P_{sb} \geq \frac{1}{2}P_T, \qquad P_{sb} \leq \frac{1}{4}P_T \tag{4.1.27}$$

Consequently, at least 50 percent of the total transmitted power resides in a carrier term that's independent of $m(t)$ and thus conveys no message information.

Alternative AM Power calculation

In radio transmission, the AM signal is amplified by a power amplifier and fed to the antenna. As you already know, AM signal is really a composite of several signal voltages, namely, the carrier and the two sidebands; each of these signals produces power in the antenna. The total transmitted power P_T is simply the sum of the carrier power P_c and the power in the two sidebands P_{USB} and P_{LSB}:

$$P_T = P_c + P_{LSB} + P_{USB} \tag{4.1.28}$$

Example 4.14

Suppose that the modulating signal, $m(t)$ is sinusoid of the form

$$m(t) = \cos 2\pi f_m t, \qquad f_m \ll f_c$$

Determine the DSB AM signal, its upper and lower sidebands and spectrum, taking the modulation index as μ.

Solution

From 4.1.17 where $\mu = K_a A_m$, the DSB AM is

$$\begin{aligned} s(t) &= A_c\left[1+\mu\cos 2\pi f_m t\right]\cos(2\pi f_c t + \phi_c) \\ &= A_c \cos\left(2\pi f_c t + \phi_c\right) + \frac{A_c\mu}{2}\cos\left[2\pi(f_c - f_m)t + \phi_c\right] + \frac{A_c\mu}{2}\cos\left[2\pi(f_c + f_m)t + \phi_c\right] \end{aligned}$$

The lower sideband component is

$$s_{LSB}(t) = \frac{A_c\mu}{2}\cos\left[2\pi(f_c - f_m)t + \phi_c\right]$$

and the upper sideband is

$$s_{LSB}(t) = \frac{A_c\mu}{2}\cos\left[2\pi(f_c + f_m)t + \phi_c\right]$$

Fourier transform of the modulated signal is

$$\begin{aligned} S(f) = {} & \frac{A_c}{2}\left[e^{j\phi_c}\delta\left(f - f_c\right) + e^{-j\phi_c}\delta\left(f + f_c\right)\right] \\ & + \frac{A_c\mu}{4}\left[e^{j\phi_c}\delta\left(f - f_c + f_m\right) + e^{-j\phi_c}\delta\left(f + f_c - f_m\right)\right] \\ & + \frac{A_c\mu}{4}\left[e^{j\phi_c}\delta\left(f - f_c - f_m\right) + e^{-j\phi_c}\delta\left(f + f_c + f_m\right)\right] \end{aligned}$$

The line spectrum is given below

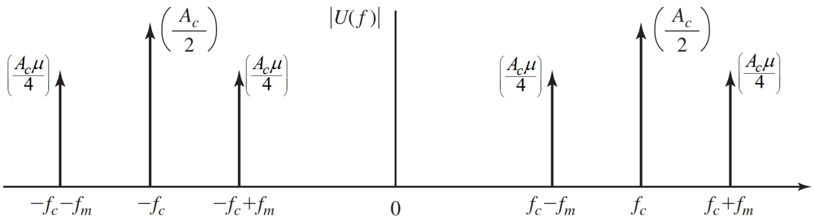


Figure 4.4: DSB AM spectrum

Example 4.15

An AM signal of the form

$$s(t) = A_c \cos(2\pi f_c t)\left[1 + \mu\cos\left(2\pi f_m + \pi/3\right)\right]$$

contains a total power of 1000 W with a modulation index of 0.8. Find the power contained in the carrier and the sidebands, also find the efficiency. Draw the amplitude and phase spectra of the AM signal.

Solution

The total power is given by

$$1000 = \langle x^2(t)\rangle = \frac{A_c^2}{2} + \frac{\mu^2 A_c^2}{2}\cdot\langle m^2(t)\rangle$$

It should be clear that in this problem $m(t) = cos(2\pi f_m t)$, so $\langle m^2(t)\rangle = 1/2$

Thus,

$$1000 = A_c^2\left[\frac{1}{2} + \frac{1}{4}\times 0.64\right] = \frac{33}{50}A_c^2$$

Hence,

$$A_c^2 = 1000 \times \frac{50}{33} = 1515.15$$

and

$$P_c = \frac{1}{2}A_c^2 = \frac{1515.15}{2} = 757.58 \text{ W}$$

The total sideband power is given by

$$P_{sb} = 1000 - P_c = 242.42 \text{ W}$$

Efficiency of the system is given by

$$\eta = \frac{242.42}{1000} = 0.242 \text{ or } 24.2\,\%$$

The magnitude and phase spectra can be plotted by first expanding $s(t)$

$$
\begin{aligned}
s(t) &= A_c \cos(2\pi f_c t) + \mu A_c \cos\left(2\pi f_m t + \pi/3\right)\cos(2\pi f_c t) \\
&= A_c \cos(2\pi f_c t) + \frac{\mu A_c}{2}\cos\left[2\pi(f_c + f_m)t + \pi/3\right] + \frac{\mu A_c}{2}\cos\left[2\pi(f_c - f_m)t - \pi/3\right]
\end{aligned}
$$

The spectrum is given below

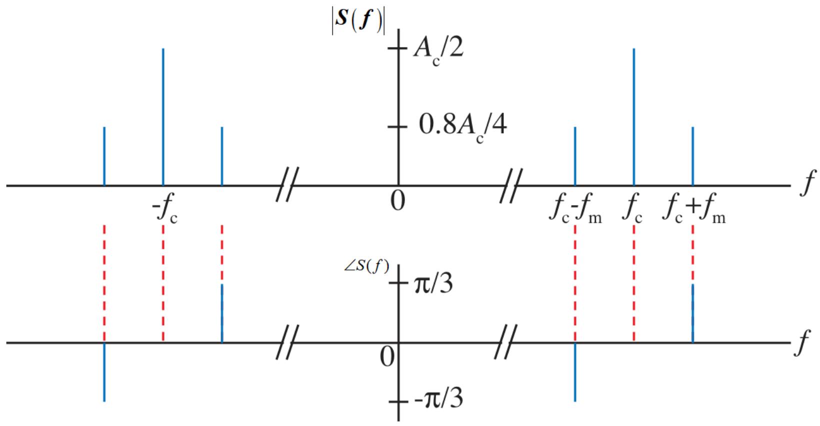


Figure 4.5: Amplitude and phase spectra for AM signal

4.1.3 Transmission Efficiency

The advantage of envelope detection in AM has its price. In AM, the carrier term does not carry any information, and hence, the carrier power is wasted. The sideband power is the useful power and the carrier power is the power wasted for convenience.

The transmission efficiency, η is the ratio of the total sideband power to the total power in the modulated wave (using *Section 4.1.2*, *Eqn.* 4.1.9 and there on)

$$
\begin{aligned}
\eta &= \frac{\text{Power in sidebands}}{\text{Total power } (P_T)} = \frac{P_{USB} + P_{LSB}}{P_c\left[1 + \frac{\mu^2}{2}\right]} \\
\eta &= \frac{\mu^2}{2 + \mu^2}
\end{aligned}
\tag{4.1.29}
$$

With the condition that $0 \le \mu \le 1$. It can be seen that, η increases monotonically with μ, and $\eta_{\max}$ occurs at $\mu = 1$, for which $\eta_{\max} = 33\%$

Note that, the efficiency can be stated as

$$
\eta = \frac{\mu^2 \left\langle m^2(t) \right\rangle}{1 + \mu^2 \left\langle m^2(t) \right\rangle} = \frac{\mu^2 P_m^2}{1 + \mu^2 P_m^2}
\tag{4.1.30}
$$

using *Eqn.* 4.1.24. Thus, for tone modulation, under best conditions $\mu = 1$, only one-third of the transmitted power is used for carrying message. For practical signals, the efficiency is even worse – in the orders of 25% or lower—compared to that of the DSB-SC case. The best condition implies $\mu = 1$. Smaller values of μ degrade efficiency further. For this reason volume compression and peak limiting are commonly used in AM to ensure that full modulation ($\mu = 1$) is maintained, most of the time.

Example 4.16

Determine η and the percentage of the total power carried by the sidebands of the AM wave for tone modulation when (a) $\mu = 0.5$ and (b) $\mu = 0.3$.

Solution
For $\mu = 0.5$,

$$\eta = \frac{\mu^2}{2+\mu^2} \times 100\% = \frac{(0.5)^2}{2+(0.5)^2} \times 100\% = 11.11\%$$

Hence, only about 11% of the total power is sidebands.

For $\mu = 0.3$

$$\eta = \frac{(0.3)^2}{2+(0.3)^2} \times 100\% = 4.3\%$$

Here, only 4.3% of the total power is the useful power (i.e. power in the sidebands).

Example 4.17

Suppose that on an AM signal, the $A_{max(p-p)}$ value read from the graticule on the oscilloscope screen is 5.9 divisions and $A_{min(p-p)}$ is 1.2 divisions.

i. What is the modulation index?

ii. Calculate A_c, A_m, and m if the vertical scale is 2 V per division. (Hint: Sketch the signal)

Solution – Getting modulation index (m) from AM wave

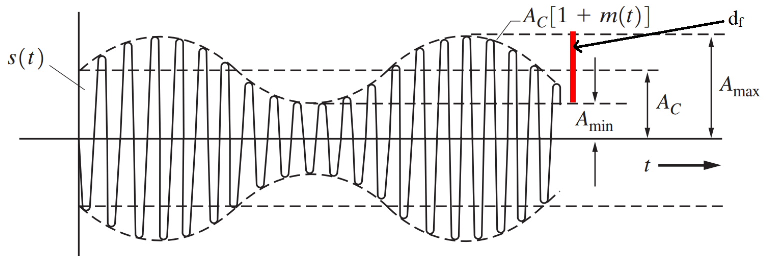


Figure 4.6: Complete AM modulated wave

i.

$$m = \frac{A_{\text{max}} - A_{\text{min}}}{A_{\text{max}} + A_{\text{min}}} = \frac{5.9 - 1.2}{5.9 + 1.2} = 0.662$$

ii.

$$A_c = \frac{A_{\text{max}} + A_{\text{min}}}{2} = \frac{5.9 + 1.2}{2} = 3.55 \quad @ \quad \frac{2\text{ V}}{\text{div}}$$

$$A_c = 3.55 \times 2 \text{ V} = 7.1 \text{ V}$$

$$A_m = \frac{A_{\text{max}} - A_{\text{min}}}{2} = \frac{5.9 - 1.2}{2} = 2.35 \ @ \ \frac{2 \text{ V}}{\text{div}}$$

$$A_m = 2.35 \times 2 \text{ V} = 4.7 \text{ V}$$

Now, m

$$m = \frac{A_m}{A_c} = \frac{4.7}{7.1} = 0.662$$

Example 4.18

Derive an expression for multitone amplitude modulation, total transmitted power and total modulation index.

Solution

The AM wave is given in *Eqn.* 4.1.1, that is

$$s(t) = A_c\left[1 + K_a m(t)\right]\cos\left(2\pi f_c t\right)$$

Let's take two distinct modulating signal

$$m_1(t) = A_{m_1}\cos 2\pi f_{m_1} t \ \text{ and } \ m_2(t) = A_{m_2}\cos 2\pi f_{m_2} t$$

Therefore,

$$\begin{aligned}
s(t) &= A_c\left[1 + K_a\left(m_1(t) + m_2(t)\right)\right]\cos 2\pi f_c t \\
&= A_c\left[1 + K_a\left(A_{m_1}\cos 2\pi f_{m_1} t + A_{m_2}\cos 2\pi f_{m_2} t\right)\right]\cos 2\pi f_c t \\
&= A_c\left[1 + \left(\underbrace{K_a A_{m_1}}_{\mu_1}\cos 2\pi f_{m_1} t + \underbrace{K_a A_{m_2}}_{\mu_2}\cos 2\pi f_{m_2} t\right)\right]\cos 2\pi f_c t \\
&= A_c\left[1 + \mu_1\cos 2\pi f_{m_1} t + \mu_2\cos 2\pi f_{m_2} t\right]\cos 2\pi f_c t
\end{aligned}$$

$$s(t) = A_c\cos 2\pi f_c t + \mu_1 A_c\cos 2\pi f_{m_1} t \bullet \cos 2\pi f_c t + \mu_2 A_c\cos 2\pi f_{m_2} t \bullet \cos 2\pi f_c t$$

Applying $\cos A\cos B = \frac{1}{2}\cos\left(A - B\right) + \frac{1}{2}\cos\left(A + B\right)$, $s(t)$ becomes

$$\begin{aligned}
s(t) = A_c\cos 2\pi f_c t &+ \frac{\mu_1 A_c}{2}\cos 2\pi\left(f_c - f_{m_1}\right)t + \frac{\mu_1 A_c}{2}\cos 2\pi\left(f_c + f_{m_1}\right)t \\
&+ \frac{\mu_2 A_c}{2}\cos 2\pi\left(f_c - f_{m_2}\right)t + \frac{\mu_2 A_c}{2}\cos 2\pi\left(f_c + f_{m_2}\right)t
\end{aligned}$$

It is clear here that, when we have two modulating signal frequencies, we get four additional frequencies, two upper sidebands, $f_c + f_{m_1}$, $f_c + f_{m_2}$ and two lower sidebands, $f_c - f_{m_1}$, $f_c - f_{m_2}$

The total transmitted power in the AM wave is calculated as below

$$\begin{aligned}
P_T &= P_c + P_{USB1} + P_{USB2} + P_{LSB1} + P_{LSB2} \\
&= \frac{\left(A_c/\sqrt{2}\right)^2}{R} + \frac{\mu_1^2 A_c^2}{8R} + \frac{\mu_1^2 A_c^2}{8R} + \frac{\mu_2^2 A_c^2}{8R} + \frac{\mu_2^2 A_c^2}{8R} \\
&= \frac{A_c^2}{2R} + \frac{\mu_1^2 A_c^2}{4R} + \frac{\mu_2^2 A_c^2}{4R} \\
&= \frac{A_c^2}{2R}\left[1 + \frac{\mu_1^2}{2} + \frac{\mu_2^2}{2}\right] \\
&= P_c\left[1 + \frac{\mu_T^2}{2}\right]
\end{aligned}$$

where $\mu_T = \sqrt{\mu_1^2 + \mu_2^2}$ In general, the total modulation index is given by

$$\mu_T = \sqrt{\mu_1^2 + \mu_2^2 + \mu_3^2 + \ldots + \mu_n^2}$$

Example 4.19

Given that the first input to AM Modulator is 500 kHz Carrier signal with amplitude of 20 V. The second input to the AM Modulator is a 10 kHz modulating signal with ±7.5 V. Determine the following:

i. USB and LSB

ii. Modulation index and percent modulation, μ

iii. Peak amplitude of modulation carrier and upper and lower side frequency

iv. Maximum & minimum amplitude of the envelope, V_{max} and V_{min}

v. Draw the output in frequency domain and time domain

Solution

i. USB and LSB
$f_{usb} = 500 + 10 = 510$ kHz and $f_{lsb} = 500 - 10 = 490$

ii. Modulation index and percent modulation, μ

$$\mu = \frac{E_m}{E_c} = \frac{7.5}{20} = 0.375$$

Percent modulation

$$\mu\% = 0.375 \times 100\% = 73.5\%$$

iii. Peak amplitude of modulation carrier and Upper and Lower side frequency.
Peak amplitude of the carrier is $E_c = 20$ V
Amplitude of the lower (E_{LSB}) and upper (E_{USB})

$$E_{LSB} = E_{USB} = \frac{E_m}{2} = \frac{7.5}{2} = 3.75 \text{ V}$$

Alternatively

$$E_{LSB} = E_{USB} = \frac{mE_c}{2} = \frac{(0.375)(20)}{2} = 3.75 \text{ V}$$

$$E_{LSB} = E_{USB} = \frac{mE_c}{2} = \frac{(0.375)(20)}{2} = 3.75 \text{ V}$$

iv. Maximum & minimum amplitude of the envelope, V_{max} and V_{min}

$$V_{\max} = E_c + E_m = 20 + 7.5 = 27.5\,\text{V}$$

$$V_{\min} = E_c - E_m = 20 - 7.5 = 12.5\,\text{V}$$

v. Draw the output in frequency domain and time domain

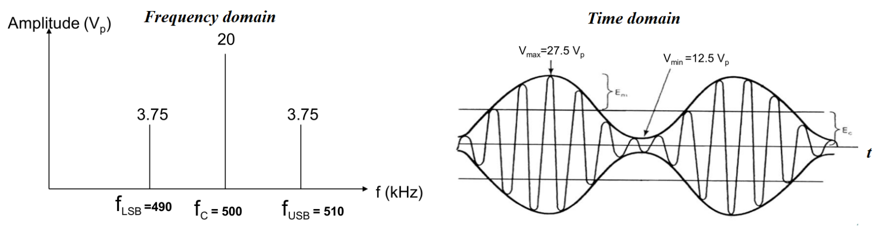


Figure 4.7: Freq. and Time domains representation

4.1.4 Generation of Conventional AM Wave

Since the process of modulation involves the generation of new frequency components, modulators are generally characterized as nonlinear or time-variant systems.

4.1.4.1 Power-Law or Square-Law Modulator

An amplitude-modulated wave can be obtained by combining the modulating signal and the carrier through a nonlinear device. The nonlinear device could be a diode or a transistor. Transistors have advantage over diode that it also provides amplification. One popular method for generating AM wave is to use a square-law modulator.

The Square-law modulator consist of **Summer** – It adds the carrier and the modulating signal, **Non-linear device** – a device with non-linear input-output relation and **Band-pass filter (BPF)** – this extract the deserved signal from the modulator's product. The input-to-output voltage $(v_o(t),\, v_{in}(t))$ relation of current passing through a diode connected to a load resistor is given by the following polynomial approximation:

$$v_o(t) = a_1 v_{in}(t) + a_2 v_{in}^2(t) + a_2 v_{in}^3(t) + \ldots. \tag{4.1.31}$$

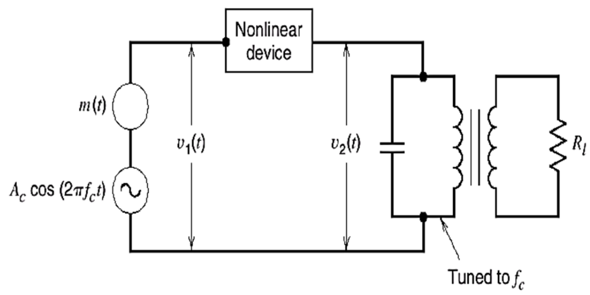


Figure 4.8: Square law modulator

where the a_is are constants. If $|v_{in}t|$ is very small, the higher power terms can be neglected, hence *Eqn.* 4.1.31 becomes

$$v_o(t) = a_1 v_{in}(t) + a_2 v_{in}^2(t) \tag{4.1.32}$$

Figure 4.9 shows the schematic of a square-law modulator for generating a conventional AM wave If, $v_{in}(t) = m(t) + A_c \cos \omega_c t$, then the output of the square-law device is

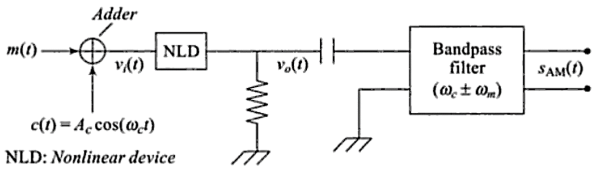


Figure 4.9: Square-law modulator for AM wave generation

$$\begin{aligned} v_o(t) &= a_1\left[m(t) + A_c \cos \omega_c t\right] + a_2\left[m(t) + A_c \cos \omega_c t\right]^2 \\ &= a_1 m(t) + a_2 m^2(t) + \frac{a_2 A_c^2}{2}\left[1 + \cos(2\omega_c t)\right] \\ &\underbrace{+ a_1 A_c \left[1 + \left(\frac{2a_2 A_c}{a_1}\right) m(t)\right] \cos(\omega_c t)}_{\text{Desired AM wave}} \end{aligned} \tag{4.1.33}$$

The desired AM wave is obtained by passing $v_o(t)$ through a bandpass filter centered at ω_c with a bandwidth equals to the AM bandwidth (i.e. $2\omega_m$). It is interesting to verify this by looking at the spectrum of $v_o(t)$ shown in *Figure* 4.10. In order to identify various terms in the spectrum, we have rearranged conveniently the expression for $v_o(t)$.

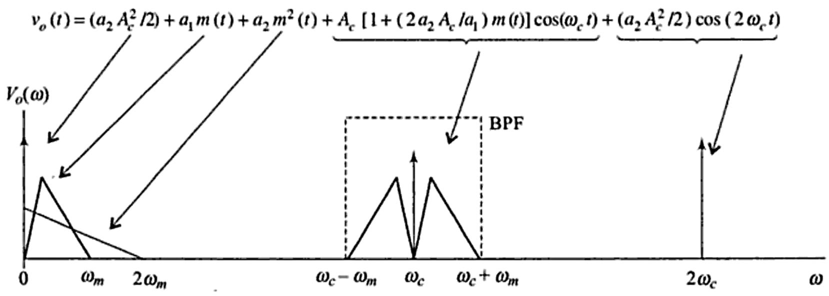


Figure 4.10: Spectrum of $v_o(t)$

It can be observed that the spectrum corresponding to $m^2(t)$ is different from that of $m(t)$. This follows from the frequency-domain convolution property of the Fourier transform. It is also clear from the figure that the desired AM term can be obtained by (ideal) bandpass filtering provided the lower edge of the filter $(\omega_c - \omega_m)$, is greater than the upper edge $(2\omega_m)$, of the nearest potential interfering component $(m^2(t))$. Hence, the required condition is

$$\omega_c - \omega_m \geq 2\omega_m \text{ or } \omega_c \geq 3\omega_m \tag{4.1.34}$$

For the square-law modulator to work, the carrier frequency should be at least three times the maximum frequency of the baseband signal.

4.1.4.2 The Switching Modulator

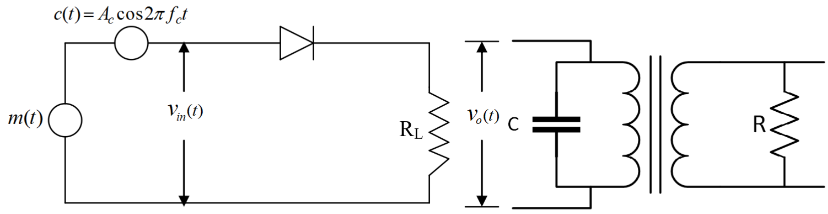


Figure 4.11: Switching circuit diagram

Assume that carrier wave $c(t)$ applied is large in amplitude, so that, it swings right across the characteristic curve of the diode. We also assume that the diode acts as an ideal switch; that is, it presents zero impedance when it is forward-biased and infinite impedance when it is reverse-biased. We may approximate the transfer characteristic of the diode-load resistor combination by a piecewise-linear characteristic.

The total input voltage, $v_{in}(t)$, applied to the diode terminal is the sum of both carrier and message signals

$$v_{in}(t) = m(t) + c(t) = m(t) + A_c \cos 2\pi f_c t \tag{4.1.35}$$

During the positive half cycle of the carrier signal i.e. if $c(t) > 0$, the diode is forward biased; it then acts as a closed switch. The output voltage $v_o(t)$ is the same as the input voltage $v_{in}(t)$. During the negative half cycle of the carrier signal i.e. if $c(t) < 0$, the diode is reverse biased that acts as an open switch. Here, the output voltage $v_o(t)$ is zero. Finally, the output voltage varies periodically between the values input voltage $v_{in}(t)$ and zero at a rate equal to the carrier frequency f_c

$$\begin{aligned} v_o(t) &= \begin{cases} v_{in}(t), & c(t) > 0 \\ 0, & c(t) < 0 \end{cases} \\ v_o(t) &= v_{in}(t) \cdot p(t) \\ &= [m(t) + A_c \cos 2\pi f_c t]\, p(t) \end{aligned} \tag{4.1.36}$$

where $p(t)$ is periodic pulse train with duty cycle one-half period, i.e. $T_o/2$, where $T_o = 1/f_c$ given by

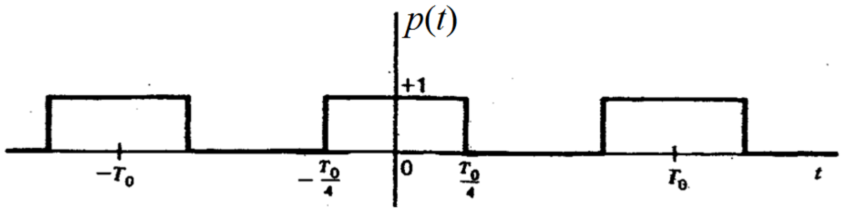


Figure 4.12: Periodic pulse train

Thus,

$$v_o(t) = [m(t) + A_c \cos 2\pi f_c t] \left[\frac{1}{2} + \frac{2}{\pi} \cos 2\pi f_c t - \frac{2}{3\pi} \cos 6\pi f_c t + - + - \right] \tag{4.1.37}$$

The required AM signal centered at f_c can be separated using band pass filter. The filter output is

$$s(t) = \frac{A_c}{2} \left[1 + \frac{4}{\pi} \frac{m(t)}{A_c} \right] \cos 2\pi f_c t \tag{4.1.38}$$

For a single-tone message, set $m(t) = A_m \cos 2\pi f_m t$ and the output becomes

$$s(t) = \frac{A_c}{2} \left[1 + \frac{4}{\pi} \frac{A_m}{A_c} \cos 2\pi f_m t \right] \cos 2\pi f_c t \tag{4.1.39}$$

Therefore, modulation index

$$m = \frac{4}{\pi} \frac{A_m}{A_c} \tag{4.1.40}$$

4.1.5 Demodulation of Conventional AM Signal

The process of recovering the message signal from the modulated signal is called demodulation or detection.

4.1.5.1 Envelope Detection

Normal (conventional) DSB AM signals are easily demodulated by means of an envelope detector. An envelope detector consists of a diode and an *RC* circuit that is basically a lowpass filter, *Figure* 4.13.

During positive half cycle of the input signal, the diode is forward-biased and the capacitor charges up to the peak value of the input signal. When the input voltage falls below the peak value, the diode becomes reverse biased, and the capacitor discharges slowly through the load resistor R_L. As a result, only positive half cycle of AM wave appears across R_L.
The discharge process continues until the next positive half cycle. When the input signal becomes greater than the voltage across the capacitor, the diode conducts again and the process repeats.

Selection of RC time constant

I. The capacitor C charges through 'D' and R_s when the diode is 'ON' and discharges through R_L when the diode is 'OFF'.

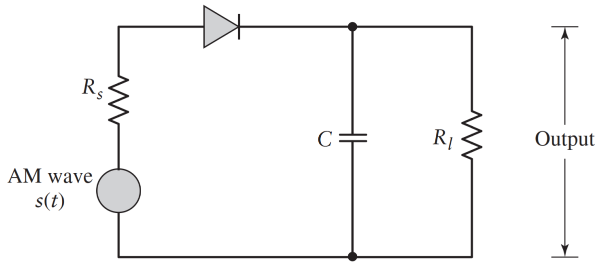


Figure 4.13: Envelope detector circuit diagram

II. The charging time constant R_sC must be short compared with the carrier period $\frac{1}{f_c}$. Therefore

$$R_sC \ll \frac{1}{f_c} \tag{4.1.41}$$

In this case the capacitor charges rapidly.

III. The discharging time constant R_LC should be long enough to ensure that the capacitor discharges slowly through the load resistor R_L between positive peak of the carrier wave. Thus

$$\frac{1}{f_c} \ll R_LC \ll \frac{1}{W} \tag{4.1.42}$$

where W is the maximum modulating frequency.

Figure 4.14 shows the effect of *RC* on the cut-off frequency of the envelope detector

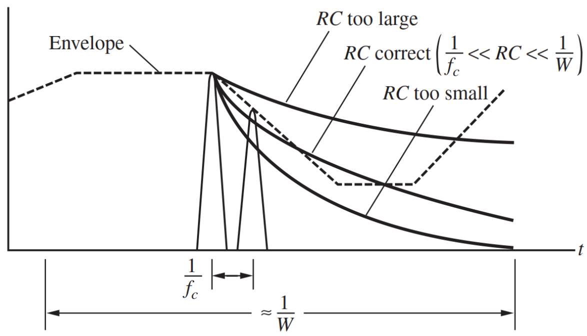


Figure 4.14: Effect of *RC* on the envelope detector

The result is such that the capacitor voltage or detector output is nearly the same as the envelope of the AM wave. The detector's output usually has a small ripple of the carrier frequency. This ripple is removed by using lowpass filter. One other AM detector that is equally good but will not be discussed is **Square-Law demodulator**.

Distortions in Envelope Demodulator Output

(a) *Normal output*; (b) *Diagonal Clipping*, when R_LC is too large or long; (c) *Ragged output*, when R_LC is too small or short; (d) *Negative clipping*, distortion due to over modulation.

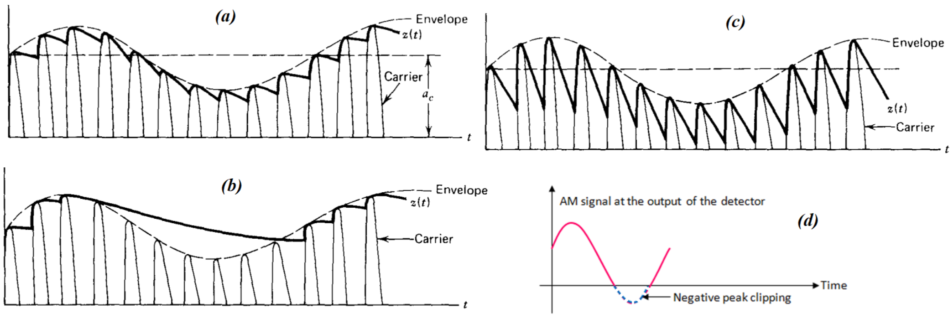


Figure 4.15: Distortion in Envelope detector output: (a) Normal output (b) Diagonal clipping (c) Ragged output (d) Negative clipping

4.1.6 Advantages, Disadvantages and Applications of AM

Advantages	Disadvantages
AM transmission are less complex	Power wasted in the transmitted signal. Most of the transmitted power is in the carrier wave which does not carry any information
Receivers as simple and detection is easy. AM requires larger bandwidth	AM is wasteful of channel bandwidth. The transmitted signal requires twice the bandwidth of the message. This is due to the transmission of both sidebands.
AM receivers are cost efficient	Easily affected by channel noise

4.2 Double Sideband-Suppressed Carrier Modulation (DSB-SC)

DSB-SC modulation results when $A(t)$ is proportional to the message signal $m(t)$. It is obtained by the product modulator that simply multiplies the message signal $m(t)$ by the carrier wave $x_c(t) = A_c \cos(2\pi f_c t)$. DSB-SC signal is given by

$$\begin{aligned} s(t) &= x_c(t)m(t) \\ &= A_c m(t)\cos(\omega_c t + \phi_c) \end{aligned} \tag{4.2.1}$$

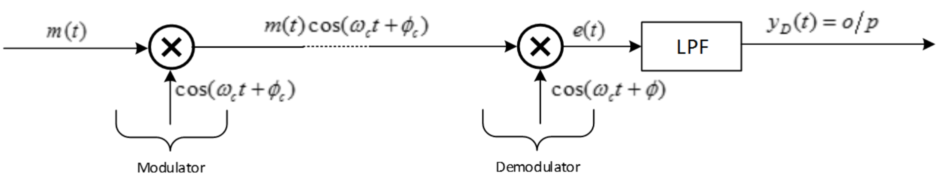


Figure 4.16: DSB-SC generation and demodulator

Unlike amplitude modulation, DSB-SC modulation (*Eqn.*4.2.1) is reduced to zero whenever the message signal is switched off. That is, $s(t)$ undergoes phase reversal whenever $m(t)$ crosses the x-axis.

Most noticeably, however, is the fact that the modulated signal undergoes a phase reversal whenever the message signal crosses zero, as indicated in *Figure* 4.17b for the message signal depicted in *Figure* 4.17a. The envelope of a DSB-SC modulated signal is therefore different from the message signal, which means that simple demodulation using an envelope detection is not a viable option for DSB-SC modulation.

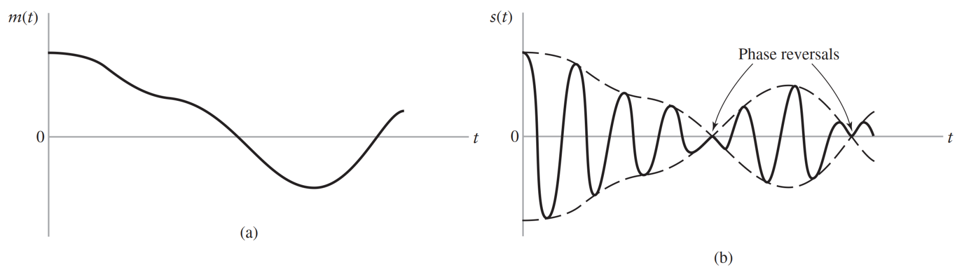


Figure 4.17: (a) Message signal $m(t)$. (b) DSB-SC modulated wave $s(t)$

4.2.1 Bandwidth Requirements

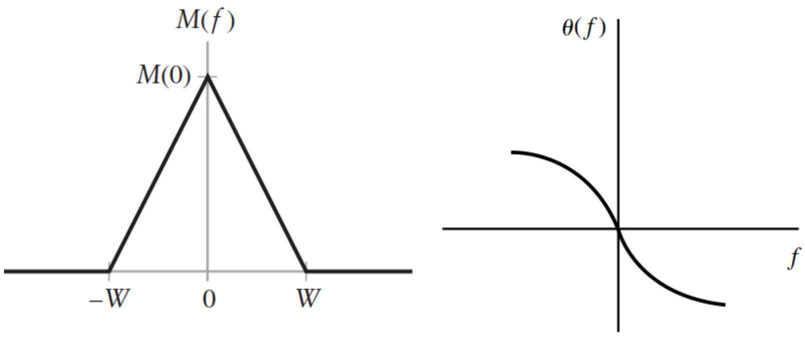
The spectrum of the modulated signal can be obtained by taking Fourier transform of $s(t)$

$$\begin{aligned} S(f) &= \mathscr{F}\left[m(t)\right] \star \mathscr{F}\left[A_c \cos\left(2\pi f_c t + \phi_c\right)\right] \\ &= M(f) \star \frac{A_c}{2}\left[e^{j\phi_c}\delta\left(f - f_c\right) + e^{-j\phi_c}\delta\left(f + f_c\right)\right] \\ &= \frac{A_c}{2}\left[M\left(f - f_c\right)e^{j\phi_c} + M\left(f + f_c\right)e^{-j\phi_c}\right] \end{aligned} \tag{4.2.2}$$

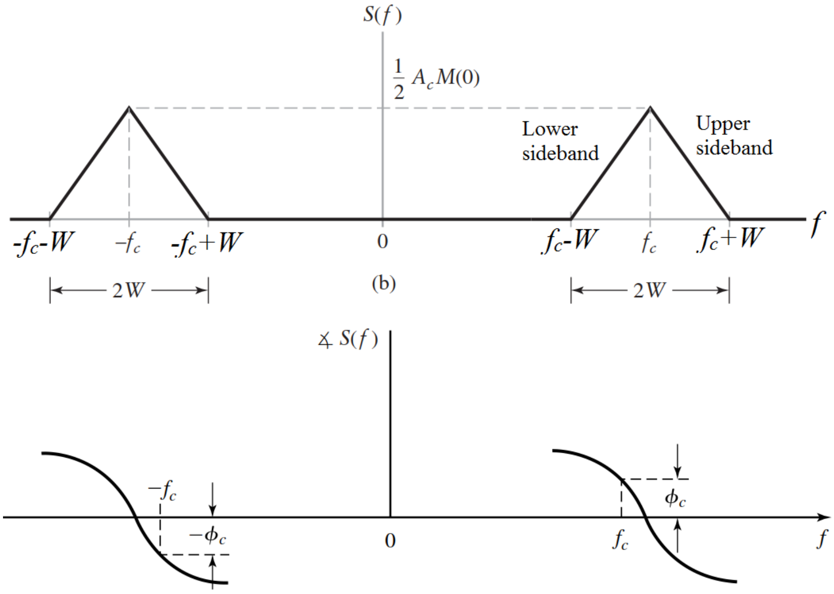
where $m(t) \rightleftharpoons M(f)$.
When the $m(t)$ is limited to the interval $-W \leq f \leq W$, (*Figure* 4.18a), the spectrum of $S(f)$ can be obtained as in *Figure* 4.18b. Of course, the transmission bandwidth required by DSB-SC modulation is **2W**, where W is the message frequency

It is observed that the magnitude of the spectrum of the message signal $m(t)$ has been translated or shifted in frequency by an amount f_c. The phase of the message signal has been translated in frequency and offset by the carrier phase ϕ_c. Furthermore, the bandwidth occupancy of amplitude-modulated signal is $2W$, whereas bandwidth of the message signal $m(t)$ is W. Therefore, channel bandwidth required to transmit modulated signal $s(t)$ is $B_c = 2W$.
Frequency content of modulated signal $s(t)$ in the frequency band $|f| > f_c$ is called **upper sideband** of $S(f)$, and the frequency content in the frequency band $|f| < f_c$ is called **lower sideband** of $S(f)$.

It is important to note that either one of the sidebands of $S(f)$ contains all frequencies that are in $M(f)$. That is, frequency content of $S(f)$ for $f > f_c$ corresponds to frequency content of $M(f)$ for $f > 0$, and frequency content of $S(f)$ for $f < -f_c$ corresponds to frequency content of $M(f)$ for $f < 0$. Hence, the upper sideband of $S(f)$ contains all frequencies in $M(f)$. A similar statement applies to the lower sideband of $S(f)$. Therefore, the lower sideband of $S(f)$ contains all frequency content of the message signal $M(f)$. Since $S(f)$ contains both the upper and the lower sidebands, it is called a double-sideband (DSB) AM signal.



(a) Spectrum of message signal $m(t)$



(b) Spectrum of DSB-SC modulated wave $s(t)$

Figure 4.18: Magnitude and phase spectra of the message, $m(t)$ and DSB-SC signal, $s(t)$

The other characteristic of the modulated signal $s(t)$ is that it does not contain a carrier component. That is, all the transmitted power is contained in the modulating (message) signal $m(t)$, since the carrier term is suppressed in the spectrum as there is no impulse at $\pm f_c$. This is evident from observing the spectrum of $S(f)$. We note that, as long as $m(t)$ does not have any DC component, there is no impulse in $S(f)$ at $f = f_c$, which would be the case if a carrier component was contained in the modulated signal $s(t)$. For this reason, $s(t)$ is called a suppressed-carrier (SC) signal. Therefore, $s(t)$ is a DSB-SC AM signal.

In short, insofar as bandwidth occupancy is concerned, DSB-SC offers no advantage over AM. Its only advantage lies in saving transmitted power, which is important enough when the available transmitted power is at a premium.

Example 4.20

Suppose that the modulating signal $m(t)$ is a sinusoid of the form

$$m(t) = A_m \cos 2\pi f_m t, \qquad f_m \ll f_c$$

Determine the DSB-SC AM signal and its upper and lower sidebands. Take the carrier signal as $c(t) = A_c \cos(2\pi f_c t + \phi_c)$.

Solution
The DSB-SC AM is expressed in the time domain as

$$\begin{aligned} s(t) &= m(t)c(t) \\ &= \frac{A_c A_m}{2} \cos[2\pi(f_c - f_m)t + \phi_c] + \frac{A_c A_m}{2} \cos[2\pi(f_c + f_m)t + \phi_c] \end{aligned}$$

Converting $s(t)$ into frequency domain using Fourier transform, the signal is in the form

$$\begin{aligned} S(f) = \frac{A_c A_m}{4} &\left[e^{j\phi_c} \delta(f - f_c + f_m) + e^{-j\phi_c} \delta(f + f_c - f_m)\right] \\ + \frac{A_c A_m}{4} &\left[e^{j\phi_c} \delta(f - f_c - f_m) + e^{-j\phi_c} \delta(f + f_c + f_m)\right] \end{aligned}$$

Using $S(f)$, the spectrum (*Figure* 4.19) of $s(t)$ can be drawn to outline the various frequency components.

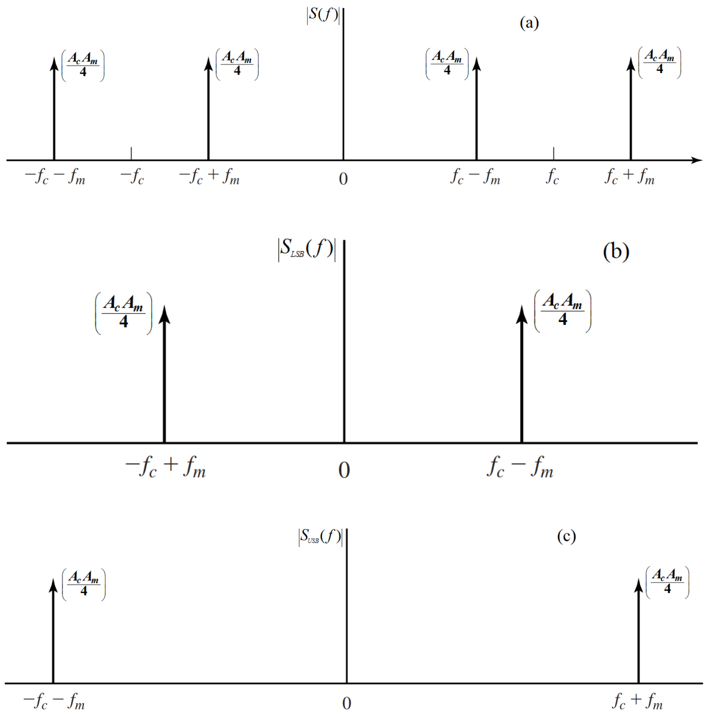


Figure 4.19: (a) The (magnitude) spectrum of a DSB-SC AM signal for a sinusoidal message signal (b) Lower sideband (c) Upper sideband

The lower sideband of $s(t)$ is

$$s_{LSB}(t) = \frac{A_c A_m}{2} \cos[2\pi(f_c - f_m)t + \phi_c]$$

and its spectrum is illustrated in *Figure* 4.19b. The upper sideband is give by

$$s_{USB}(t) = \frac{A_c A_m}{2} \cos\left[2\pi\left(f_c + f_m\right)t + \phi_c\right]$$

Its spectrum is illustrated in *Figure* 4.19c.

4.2.2 Power Content of DSB-SC Signals

Power content of DSB-SC signal would be calculated by assuming that the phase of the signal is set to zero (i.e. $\phi_c = 0$). This is because power in a signal is independent of the phase of the signal. Employing time-average autocorrelation function of the signal, $s(t)$, we get

$$\begin{aligned}
R_s(\tau) &= \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} s(t)s(t-\tau)dt \\
&= \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} A_c^2 m(t)m(t-\tau)\cos\left(2\pi f_c t\right)\cos\left(2\pi f_c(t-\tau)\right)dt \\
&= \frac{A_c^2}{2}\lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} m(t)m(t-\tau)\left[\cos\left(4\pi f_c t - 2\pi f_c \tau\right) + \cos\left(2\pi f_c \tau\right)\right]dt \\
&= \frac{A_c^2}{2} R_m(\tau)\cos\left(2\pi f_c \tau\right)
\end{aligned} \tag{4.2.3}$$

***NOTE*:**	$\lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} m(t)m(t-\tau)\cos\left(4\pi f_c t - 2\pi f_c \tau\right)dt = 0$

Now, taking Fourier transform of both sides of *Eqn.* 4.2.3, the power-spectral density of the modulated signal is given as

$$\begin{aligned}
S_s(f) &= \mathscr{F}\left[\frac{A_c^2}{2} R_m(\tau)\cos(2\pi f_c \tau)\right] \\
&= \frac{A_c^2}{4}\left[S_m\left(f - f_c\right) + S_m\left(f + f_c\right)\right]
\end{aligned} \tag{4.2.4}$$

Eqn. 4.2.4 shows that the power-spectral density of DSB-SC signal is the power-spectral density of the message shifted upward and downward by f_c and scaled by $A_c^2/4$.
By substituting $\tau = 0$ in *Eqn.* 4.2.3, the total power in the modulated signal can be found or by integrating *Eqn.* 4.2.4 over all frequencies.

Using the first approach from *Eqn.* 4.2.3, we have

$$\begin{aligned}
P_s &= \left.\frac{A_c^2}{2} R_m(\tau)\cos\left(2\pi f_c \tau\right)\right|_{\tau=0} \\
&= \frac{A_c^2}{2} R_m(0) \\
&= \frac{A_c^2}{2} P_m
\end{aligned} \tag{4.2.5}$$

where $P_m = R_m(0)$ is the power of or in the message signal.

Example 4.21

As a follow up to Example 4.20, determine the PSD of the modulated signal, the power in the modulated signal, and the power in each of the sidebands.

Solution
From Example 4.20, the message signal was $m(t) = A_m \cos 2\pi f_m t$, its PSD is given by

$$S_m(f) = \frac{A_m^2}{4}\delta(f - f_m) + \frac{A_m^2}{4}\delta(f + f_m)$$

Substituting in *Eqn.* 4.2.4, we obtain the PSD of the modulated signal as

$$S_s(f) = \frac{A_c^2 A_m^2}{16}\left[\delta(f - f_m - f_c) + \delta(f + f_m - f_c) + \delta(f - f_m + f_c) + \delta(f - f_m - f_c)\right]$$

The total power in the modulated signal is the integral of $S_s(f)$ and is given by

$$P_s = \int_{-\infty}^{\infty} S_s(f)df = \frac{A_c^2 A_m^2}{4}$$

With the use of symmetry of sidebands the powers the upper and lower sidebands, P_{LSB} and P_{USB} are equal and given by

$$P_{LSB} = P_{USB} = \frac{P_s}{2} = \frac{A_c^2 A_m^2}{8}$$

It can also be observed from the power-spectral density of the DSB-SC signal [as in *Eqn.* 4.2.4] that the bandwidth of the modulated signal is $2W$; twice the bandwidth of the message signal, and that there exists no impulse at the carrier frequency $\pm f_c$ in the power-spectral density. Therefore, the modulated signal is *suppressed carrier* (SC).

Alternatively
From *Eqn.* 4.2.1 and Example 4.20, the power can be obtained using the DSB-SC wave as

$$s(t) = \frac{A_c A_m}{2}\cos\left[2\pi(f_c - f_m)t + \phi_c\right] + \frac{A_c A_m}{2}\cos\left[2\pi(f_c + f_m)t + \phi_c\right]$$

The upper sideband frequency power content of the wave is

$$\begin{aligned} P_{USB} &= \left(\frac{A_m A_c/2}{\sqrt{2}}\right)^2 \\ &= \frac{A_m^2 A_c^2}{8} \end{aligned}$$

Similarly, the power content of the lower sideband is (due to symmetry)

$$P_{LSB} = \frac{A_m^2 A_c^2}{8}$$

Remember that, the power is taken as a power delivered to 1 Ω resistor.

The total power is given by

$$P_T = P_{USB} + P_{LSB}$$
$$= 2\left(\frac{A_m^2 A_c^2}{8}\right) = \frac{A_m^2 A_c^2}{4}$$

Note: For sinusoidal modulation, the average power in the lower or upper side-frequency with respect to the total power in the DSB-SC modulated wave is 50 %.

4.2.3 Generation (or Modulation) of DSB-SC Wave

A DSB-SC wave consists of product of modulating signal $m(t)$ and carrier signal, usually $c(t)$, to yield $s(t)$ (*Figure* 4.20).

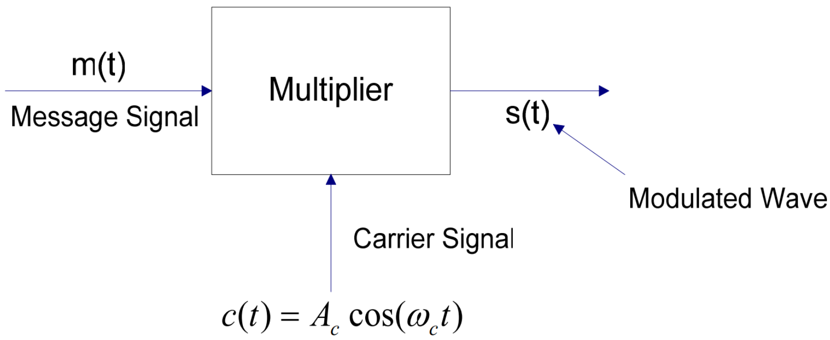


Figure 4.20: Generation of DSB-SC

The available modulators known as *product modulators* for this purpose is discussed below.

4.2.3.1 Multiplier Modulator

Modulation is achieved by multiplying $m(t)$ by $\cos \omega_c t$ using analog multiplier whose output is proportional to the product of the two input signals. This modulator has a lot of set backs and hence will not be discussed further.

4.2.3.2 Nonlinear Modulator

Modulation can also be achieved by using a nonlinear device such as a semiconductor diode or a transistor. *Figure* 4.21 shows one possible scheme that uses two identical nonlinear elements indicated as *NL*.

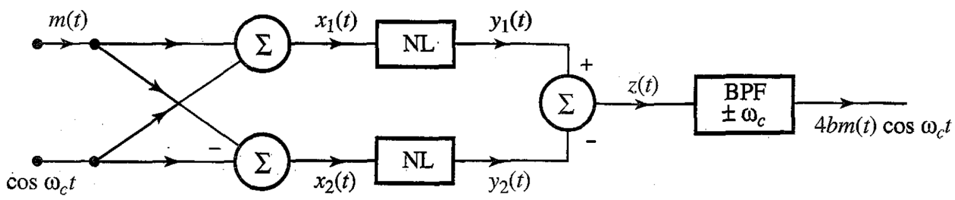


Figure 4.21: Nonlinear DSB-SC modulator

Suppose that the input-output characteristics of either of the NL is

$$y(t) = ax(t) + bx^2(t) \tag{4.2.6}$$

where $x(t)$ and $y(t)$ are the input and output respectively of the NL. The output $z(t)$ in *Figure* 4.21 is given by

$$z(t) = y_1(t) - y_2(t) = \left[ax_1(t) + bx_1^2(t)\right] - \left[ax_2(t) + bx_2^2(t)\right] \tag{4.2.7}$$

Let $x_1(t) = \cos \omega_c t + m(t)$ and $x_2(t) = \cos \omega_c t + m(t)$, then

$$\begin{aligned} y_1(t) - y_2(t) &= \left[a\left(\cos \omega_c t + m(t)\right) + b(\cos \omega_c t + m(t))^2\right] \\ &\quad - \left[a\left(\cos \omega_c t - m(t)\right) + b(\cos \omega_c t - m(t))^2\right] \\ &= 2am(t) + 4bm(t)\cos \omega_c t \end{aligned} \tag{4.2.8}$$

The spectrum of $m(t)$ is centered at the origin, whereas the spectrum of $m(t)\cos \omega_c t$ is centered at $\pm\omega_c$. Consequently, when $z(t)$ is passed through a bandpass filter tuned at ω_c, $am(t)$ is suppressed and the desired modulated signal $4bm(t)\cos \omega_c t$ passes through unharmed.

In this circuit there are two inputs; $m(t)$ and $\cos \omega_c t$. The summer output $z(t)$ does not contain the carrier signal $\cos \omega_c t$ input. Consequently, the carrier signal does not appear at the input of the final bandpass filter. This circuit acts as a balanced bridge for one of the inputs (the carrier signal). Circuits which have this characteristic are called balanced circuits.
The nonlinear modulator in *Figure* 4.21 is an example of a class of modulators known as **balanced modulators**. This circuit is balanced with respect to only one input (the carrier); the other input $m(t)$ still appears at the final bandpass filter, which must rejects $2am(t)$ in from the signal.

Balanced modulator

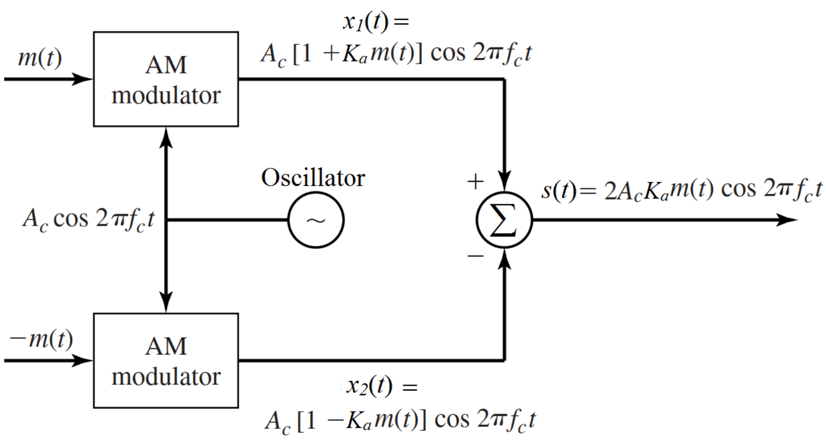


Figure 4.22: Balanced modulator

Balanced modulators consist of two amplitude modulators interconnected in a way to suppress the carrier. One input to the amplitude modulator is from an oscillator that generate the carrier wave. The second input to the top modulator is the message signal $m(t)$ while the bottom one takes $-m(t)$.

The outputs of the AM modulator are

$$x_1(t) = A_c\,[1 + K_a m(t)]\cos\omega_c t$$

and

$$x_2(t) = A_c\,[1 - K_a m(t)]\cos\omega_c t$$

The output of the summer is as before

$$s(t) = x_1(t) - x_2(t)$$

Putting parts together, we get

$$\begin{aligned} s(t) &= A_c\,[1 + K_a m(t)]\cos\omega_c t - A_c\,[1 - K_a m(t)]\cos\omega_c t \\ &= A_c\cos\omega_c t + A_c K_a m(t)\cos\omega_c t - A_c\cos\omega_c t + A_c K_a m(t)\cos\omega_c t \qquad (4.2.9) \\ s(t) &= 2A_c K_a m(t)\cos\omega_c t \end{aligned}$$

The balanced modulator output is equal to the product of the modulating signal $m(t)$ and carrier $c(t)$ except the scaling factor $2K_a$. K_a is the amplitude sensitivity of the modulator measured in per volts.
Now, taking Fourier transform of both side of *Eqn.* 4.2.9

$$S(f) = A_c K_a\,[M(f - f_c) + M(f + f_c)] \qquad (4.2.10)$$

Figure 4.23 gives the spectrum of the signal.

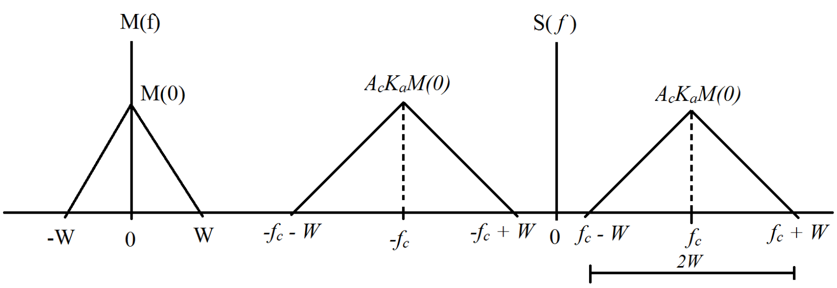


Figure 4.23: DSB-SC spectrum

4.2.3.3 Switching Modulator

A type of switching modulator that we will discuss is the **the Ring Modulator**.
Ring modulator consist of input transformer T_1, output transformer T_2 and four diodes connected in bridge.

The carrier amplitude A_c is made greater than the modulating signal amplitude, A_m, and the carrier frequency, f_c is also made greater then the modulating signal frequency, $f_m = W$; i.e. $f_c > W$. These conditions are set such that the diodes operation is controlled by $c(t)$ only. The modulating signal $m(t)$ is applied to the input transformer T_1 with the output appearing across the secondary transformer T_2.

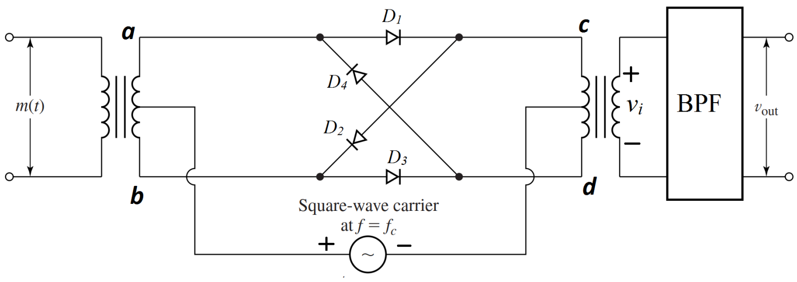


Figure 4.24: Ring modulator

Operation

During the positive half-cycles of the carrier, diodes D_1 and D_3 conduct (i.e. forward biased), and D_2 and D_4 are set opened. Hence, terminal a is connected to c, and terminal b is connected to d. Basically, the modulator multiplies the message signal $m(t)$ by +1 and therefore, $V_o(t) = m(t)$.

During the negative half-cycles of the carrier, diodes D_1 and D_3 are set opened, and D_2 and D_4 conduct, thus connecting terminal a to d and terminal b to c. This way, the output is proportional to $m(t)$ during the positive half-cycle and to $-m(t)$ (i.e. $V_o(t) = -m(t)$) during the negative half-cycle. The Ring modulator like the other product modulator output is $s(t) = c(t)m(t)$. Remember, $c(t)$ is given by the Fourier series

$$c(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos\left[2\pi f_c(2n-1)t\right] \tag{4.2.11}$$

$$s(t) = \frac{4}{\pi}m(t)\cos 2\pi f_c t - \frac{4}{3\pi}m(t)\cos 6\pi f_c t + \ldots\ldots \tag{4.2.12}$$

Taking Fourier transform of *Eqn.* 4.2.12, the spectrum of $s(t)$ can be drawn as

$$S(f) = \frac{2}{\pi}\left[M(f-f_c)+M(f+f_c)\right] - \frac{2}{3\pi}\left[M(f-3f_c)+M(f+3f_c)\right] \tag{4.2.13}$$

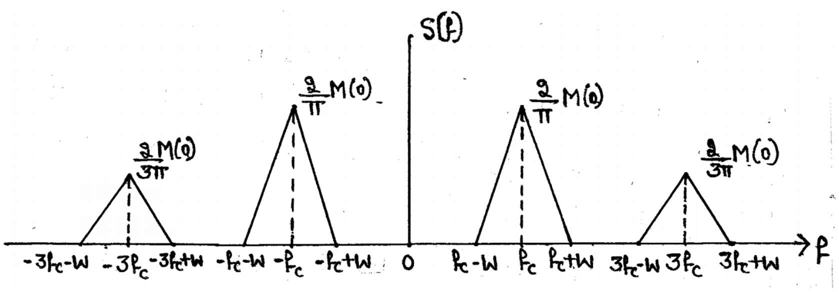


Figure 4.25: Amplitude spectrum of $s(t)$

DSB-SC wave is extracted from $s(t)$ by passing *Eqn.* 4.2.12 through an ideal BPF having carrier frequency f_c and bandwidth equal to $2W$ Hz.

The output of the BPF is

$$s(t) = \frac{4}{\pi}m(t)\cos 2\pi f_c t \tag{4.2.14}$$

4.2.4 Demodulation of DSB-SC Signals

At the receiver, the incoming signal is multiplied by a local carrier of frequency and phase in synchronism with the carrier used at the modulator. The product is then passed through a low-pass filter to filter out unwanted signals. The only difference between the modulator and the demodulator is the output filter.

In the modulator, the multiplier output is passed through a bandpass filter tuned at ω_c, whereas in the demodulator, the multiplier output is passed through a low-pass filter. For demodulation, the receiver must generate a carrier in phase and frequency synchronism with the incoming carrier. These types demodulators are called **synchronous** or **coherent demodulators**.

Assuming in the absence of noise, that is ideal channel, the received signal is equal to the modulated signal

$$\begin{aligned} r(t) &= s(t) \\ &= A_c m(t) \cos(2\pi f_c t + \phi_c) \end{aligned} \tag{4.2.15}$$

Suppose that the received signal is demodulated by first multiplying $r(t)$ by a locally generated sinusoid $\cos(2\pi f_c + \phi)$, where ϕ is the phase of the sinusoid. The product of the multiplication yields

$$\begin{aligned} r(t)\cos(2\pi f_c t + \phi) &= A_c m(t)\cos(2\pi f_c t + \phi_c)\cos(2\pi f_c t + \phi) \\ &= \frac{1}{2}A_c m(t)\cos(\phi_c - \phi) + \frac{1}{2}A_c m(t)\cos(4\pi f_c t + \phi + \phi_c) \end{aligned} \tag{4.2.16}$$

The product signal is passed through an ideal lowpass filter with a bandwidth of W. The LPF rejects double frequency components and passes only the lowpass components giving the output as

$$y_l(t) = \frac{1}{2}A_c m(t)\cos(\phi_c - \phi) \tag{4.2.17}$$

Note that $m(t)$ is multiplied by $\cos(\phi_c - \phi)$. Thus, the desired signal is scaled in amplitude by a factor that depends on the phase difference between the phase, ϕ_c, of the carrier in the received signal and the phase ϕ of the locally generated sinusoid. When $\phi_c \neq \phi$, the amplitude of the desired signal is reduced by the factor $\cos(\phi_c - \phi)$. If $\phi_c - \phi = 45^o$, the amplitude of the desired signal is reduced by $cos(45^o)$, the signal power is also reduced by a factor of two. If $\phi_c - \phi = 90^o$, the desired signal component vanishes.

The above discussion demonstrates the need for a phase-coherent or synchronous demodulator for recovering message signal $m(t)$ from received signal. That is, the phase ϕ of the locally generated sinusoid should ideally be equal to the phase, ϕ_c of the received carrier signal. A sinusoid that is phase-locked to the phase of the received carrier can be generated at the receiver in either ways. A voltage-controlled oscillator (VCO) that is controlled by a phase-locked loop (PLL) is commonly used.

PLL is used to generate a phase-coherent carrier signal that is mixed with the received signal in a balanced modulator. The output of the balanced modulator is then passed through a LPF of bandwidth W that passes the desired signal and rejects all signal and noise components above W Hz.

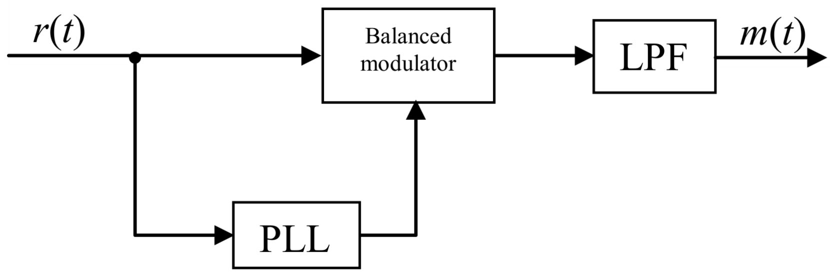


Figure 4.26: Demodulation of DSB-SC AM signal using PLL

Costas Loop Demodulator

A practical synchronous receiver system, suitable for demodulating DSB-SC waves, called the Costas-loop receiver is shown in *Figure* 4.27. It consists of two coherent detectors, both supplied with the incoming DSB-SC signal $A_c \cos \omega_c t$. The carrier in the upper channel and that in the lower channel are 90^o out-of-phase with each other. This carrier is generated by a voltage controlled oscillator (VCO). The detector in the upper path is referred to as the in-phase or I-channel and the detector in the lower path is referred to as the out-of-phase or Q-channel. The basic principle of operation of the costas receiver is that the two detectors together form a

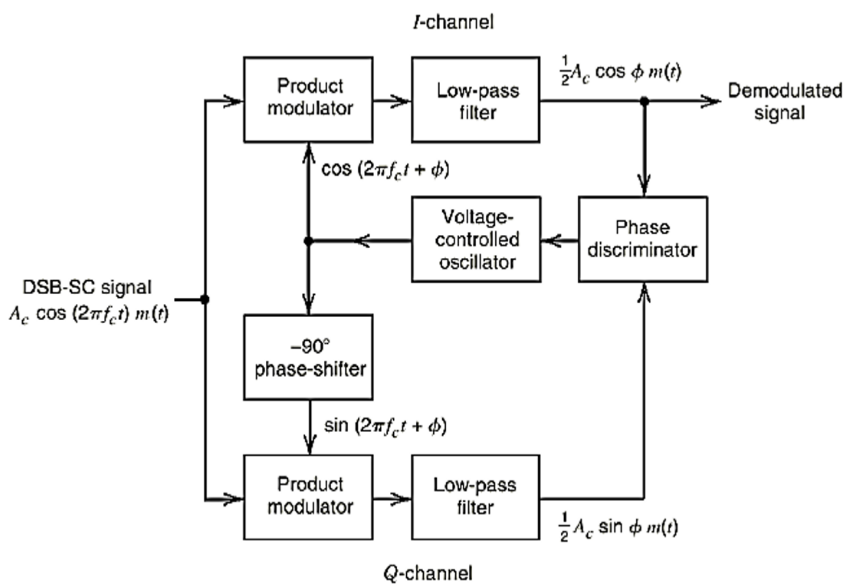


Figure 4.27: Costa Loop demodulator

negative feedback system acting in a way to maintain the local oscillator synchronous with the incoming carrier wave. Let us consider two situations:

i. Phase synchrony between the incoming carrier wave and the local oscillator carrier:
This is equivalent to $\phi = 0$. In this case, the I-channel output contains the demodulated signal $m(t)$. The Q-channel output is zero because of the quadrature null effect

ii. Loss of phase synchrony between the incoming carrier wave and the local oscillator carrier:
This is equivalent to $\phi \neq 0$. The I-channel output essentially remains unchanged. However, now, there will be some signal appearing at the Q-channel output. The amplitude of the signal is proportional to $\sin\phi$. If $\phi > 0$. the Q-channel output will have

the same polarity as the I-channel output. If $\phi < 0$. the polarity of the Q-channel output is the negative of that of the 1-channel output. The two channel outputs $A_c m(t)\cos\phi$ and $A_c m(t)\sin\phi$ are provided as inputs to the multiplier followed by a lowpass filter (*Figure* 4.27). The output of the multiplier is $\left(m^2(t)\sin 2\phi\right)/2$. The lowpass filter following the multiplier is of a very narrow bandwidth. Therefore, it essentially generates a DC output, a value proportional to $\sin 2\phi$. This is shown as $k\sin 2\phi$, at the input to the VCO in the figure. This DC signal serves as a correction signal for the VCO. The VCO generates a carrier with the required phase correction. The system then generates a carrier that is phase-locked to the incoming carrier wave. The multiplier along with the lowpass filter is essentially a phase-discriminator circuit.

If the phase-lock process is not rapid, then there will be distortions in the output. In case of circuits designed to receive modulated voice signals, the phase-lock process normally occurs so rapidly that the distortions in the output are not perceived. VCO is a system that generates a sinusoidal carrier with a frequency deviation/change proportional to its input voltage.

4.2.5 Advantages, Disadvantages, and Applications of DSB-SC

Advantages	**Disadvantages**
Low Power consumption or power saying	Design of receiver is complex
The modulation system is simple	Bandwidth required is same as that of AM
Efficiency is more than AM	The receiver must generate a replica of the carrier in order to demodulate a DSB-SC signal
Carrier wave is suppressed	Any phase and/or frequency error will result in a distorted estimate of the message signal
Linear modulation type is required	It is difficult to generate a perfect replica of the transmitted carrier
It can be used for point to point communication	A simple modification to the technique results in a less efficient transmission but simplifies the detection process greatly

- Analogue TV systems – to transmit colour information
- For transmitting stereo information in FM sound broadcast at VHF

Example 4.22

Find the value of A_o as well as the frequencies of the modulating signal and the carrier. What is the power in the modulated signal? Indicate assumptions, if any. [Hint: The power spectrum density of $A\cos(\omega_o t + \phi)$ is $\dfrac{\pi A^2}{2}[\delta(\omega-\omega_o)+\delta(\omega+\omega_o)]$

Solution

Given the DSB-SC modulated signal $x(t) = A_o \cos \omega_m t \cos \omega_c t$ Expanding, we have

$$x(t) = \frac{A_o}{2}\left[\cos\left(\omega_c - \omega_m\right)t + \cos\left(\omega_c + \omega_m\right)t\right]$$

The PSD of $x(t)$ is given by

$$S_{xx}(\omega) = \frac{\pi A_o^2}{8}\left[\delta(\omega - \omega_c + \omega_m) + \delta(\omega + \omega_c - \omega_m) + \delta(\omega - \omega_c - \omega_m) + \delta(\omega + \omega_c + \omega_m)\right]$$

Assuming $\omega_c > \omega_m$ Comparing with

$$S_x(\omega) = [\delta(\omega - 400\pi) + \delta(\omega - 360\pi) + \delta(\omega + 360\pi) + \delta(\omega + 400\pi)]$$

we obtain

$$\frac{\pi A_o^2}{8} = 50\pi, \qquad \omega_c + \omega_m = 400\pi, \qquad \omega_c - \omega_m = 360\pi$$

Solving

$$A_o = 20 \;\; \mathrm{V}, \omega_c = 380\pi \;\; \mathrm{rad/s}, \omega_m = 20\pi \;\; \mathrm{rad/s}$$

The power in the modulated signal is

4.3 Single-Sideband Amplitude Modulation

Since either the upper sideband (USB) or the lower sideband (LSB) contains the complete information of the message signal, we can conserve bandwidth by transmitting only one sideband. The modulation scheme is called *single-sideband* (SSB) modulation. It is widely used by the military and amateur radio in high-frequency (HF) communication systems. It is popular because the bandwidth required is the same as that of the modulating signal $m(t)$ (half of the bandwidth of AM or DSB-SC signal)

Consider a DSB-SC modulator using sinusoidal modulating wave

$$m(t) = A_m \cos(2\pi f_m t)$$

With the carrier $c(t) = A_c \cos(2\pi f_c t)$, the resulting DSB-SC modulated wave is defined by

$$\begin{aligned} S_{DSB}(t) &= c(t)m(t) \\ &= A_c A_m \cos(2\pi f_c t)\cos(2\pi f_m t) \\ &= \frac{1}{2}A_c A_m \cos\left[2\pi(f_c + f_m)t\right] + \frac{1}{2}A_c A_m \cos\left[2\pi(f_c - f_m)t\right] \end{aligned} \tag{4.3.1}$$

Eqn. 4.3.1 has two side-frequencies, one at $f_c + f_m$ and the other at $f_c - f_m$. Suppose we want to generate a sinusoidal SSB modulated wave by retaining the upper side-frequency at $f_c + f_m$. Suppressing the second term in *Eqn.* 4.3.1, the upper SSB modulated wave can be obtained as

$$S_{USSB}(t) = \frac{1}{2}A_c A_m \cos\left[2\pi(f_c + f_m)t\right] \tag{4.3.2}$$

Expanding *Eqn.* 4.3.2, yields

$$S_{USSB}(t) = \frac{1}{2}A_c A_m \cos(2\pi f_c t)\cos(2\pi f_m t) - \frac{1}{2}A_c A_m \sin(2\pi f_c t)\sin(2\pi f_m t) \tag{4.3.3}$$

If on the other hand, the lower side-frequency at $f_c - f_m$ is retained in the DSB-SC modulated wave, we get lower SSB modulated wave defined by

$$S_{LSSB}(t) = \frac{1}{2}A_cA_m\cos(2\pi f_ct)\cos(2\pi f_mt) + \frac{1}{2}A_cA_m\sin(2\pi f_ct)\sin(2\pi f_mt) \tag{4.3.4}$$

Examining *Eqn.* 4.3.4 and *Eqn.* 4.3.2, the difference is only the minus sign.
Eqn. 4.3.4 and *Eqn.* 4.3.2 can be combined together to get the general form of SSB modulated wave as

$$S_{SSB}(t) = \frac{1}{2}A_cA_m\cos(2\pi f_ct)\cos(2\pi f_mt) \mp \frac{1}{2}A_cA_m\sin(2\pi f_ct)\sin(2\pi f_mt) \tag{4.3.5}$$

where the plus sign applies to lower SSB and the minus sign to upper SSB.

In a more generalized form, SSB wave is defined as

$$S_{SSB}(t) = \frac{A_c}{2}m(t)\cos(2\pi f_ct) \mp \frac{A_c}{2}\hat{m}(t)\sin(2\pi f_ct) \tag{4.3.6}$$

where

$$m(t) = \sum_n a_n\cos(2\pi f_nt) \tag{4.3.7}$$

and

$$\hat{m}(t) = \sum_n a_n\sin(2\pi f_nt) \tag{4.3.8}$$

It can be noted that that the periodic signal *Eqn.* 4.3.8 can be derived from the periodic modulating signal *Eqn.* 4.3.7 simply by shifting the phase of the cosine terms in *Eqn.* 4.3.7 by -90^o. Specifically, given a Fourier transformable message signal $m(t)$ with its **Hilbert transform** denoted by the SSB modulated wave produced by $\hat{m}(t)$ is defined by

$$s(t) = \frac{A_c}{2}m(t)\cos(2\pi f_ct) \mp \frac{A_c}{2}\hat{m}(t)\sin(2\pi f_ct) \tag{4.3.9}$$

where $A_c\cos(2\pi f_ct)$ is the carrier, $A_c\sin(2\pi f_ct)$ is its -90^o phase-shifted version. The plus and minus signs apply to the lower SSB and upper SSB, respectively.

4.3.1 Generation of Single-Sideband AM

DSB-SC AM signal require a channel bandwidth of $B_c = 2W$ Hz for transmission, where W is the bandwidth of the baseband signal. SSB reduce the bandwidth required for transmission by transmitting either of the sideband.

Figure 4.28 shows generation of SSB from DSB-SC wave. The Hilbert transform filter is used to generate $\hat{m}(t)$ from $m(t)$.

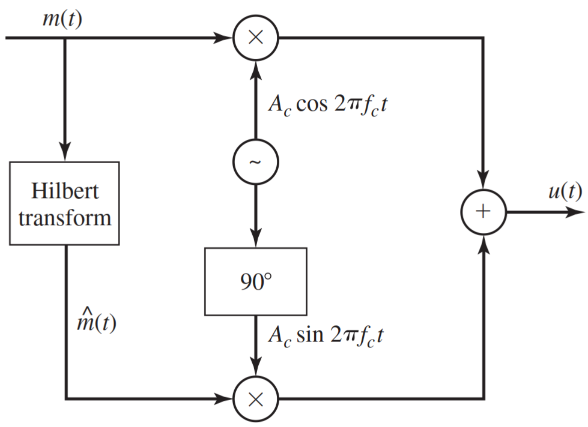


Figure 4.28: Generation of single-sideband AM signal

Another method that does not use Hilbert transform filter is shown in *Figure* 4.29.

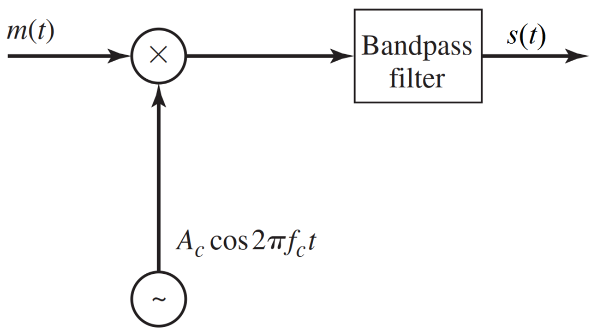


Figure 4.29: Generation of a single-sideband AM signal by filtering one of the sidebands of a DSB-SC AM signal

The method, illustrated in *Figure* 4.28, generates a DSB-SC AM signal and then employs a filter which selects either the upper sideband or the lower sideband of the double-sideband AM signal.

Example 4.23

Suppose that the modulating signal is a sinusoid of the form

$$m(t) = \cos(2\pi f_m t), \qquad f_m \ll f_c$$

Determine the two possible SSB AM signal

Solution
The Hilbert transform of $m(t)$ is

$$\hat{m}(t) = \sin(2\pi f_m t)$$

Hence,

$$s(t) = A_c \cos(2\pi f_m t)\cos(2\pi f_c t) \mp A_c \cos(2\pi f_m t)\sin(2\pi f_c t)$$

If the we take the upper $(-)$ sign, we obtain the upper sideband signal

$$s_{USSB}(t) = A_c \cos 2\pi(f_c + f_m)t$$

On the other hand, the lower $(+)$ gives

$$s_{LSSB}(t) = A_c \cos 2\pi(f_c - f_m)t$$

4.3.2 Demodulation of SSB AM Signals

To recover the message signal $m(t)$, in the received SSB AM signal, we require a phase coherent or synchronous demodulator, as was the case for DSB-SC AM signals. Using USSB signal as a case study, the received signal is

$$\begin{aligned} r(t)\cos(2\pi f_c t) &= s(t)\cos(2\pi f_c t + \phi) \\ &= \frac{1}{2}A_c m(t)\cos\phi + \frac{1}{2}A_c \hat{m}(t)\sin\phi + \text{ double frequency terms} \end{aligned}$$

By passing the equation (product signal) through an ideal lowpass filter, the double-frequency components are eliminated, leaving us with the final signal as

$$y(t) = \frac{1}{2}A_c m(t)\cos\phi + \frac{1}{2}A_c \hat{m}(t)\sin\phi$$

Further processing required to extract the intelligent signal.

4.3.3 Application of SSB AM

i. SSB is used in the systems which require minimum bandwidth such as telephone multiplex system and it is not used in broadcasting

ii. Point to point communications at frequency below 30 MHz -mobile communications, military, navigation radio etc where power saving is needed

4.4 Vestigial-Sideband AM

The stringent frequency-response requirements on the sideband filter in a SSB AM system can be relaxed by allowing a part, called a vestige, of the unwanted sideband to appear at the output of the modulator. Thus, we simplify the design of the sideband filter at the cost of a modest increase in the channel bandwidth required to transmit the signal. The resulting signal is called vestigial-sideband (VSB) AM.

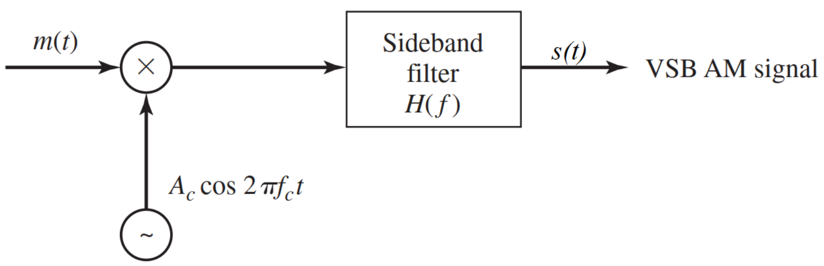


Figure 4.30: Generation of VSB AM signal.

To generate a VSB AM signal we begin by generating a DSB-SC AM signal and passing it through a sideband filter with frequency response $H(f)$ as shown in *Figure* 4.30. In the time domain the VSB signal may be expressed as

$$s(t) = [A_c m(t) \cos 2\pi f_c t] \star h(t) \tag{4.4.1}$$

where $h(t)$ is the impulse response of the VSB filter. In the frequency domain, the corresponding expression is

$$S(f) = \frac{A_c}{2} [M(f - f_c) + M(f + f_c)] H(f) \tag{4.4.2}$$

To determine the frequency-response characteristics of the filter, let us consider the demodulation of the VSB signal $s(t)$. We multiply $s(t)$ by the carrier component $\cos 2\pi f_c t$ and pass the result through an ideal lowpass filter, as shown in *Figure* 4.31.

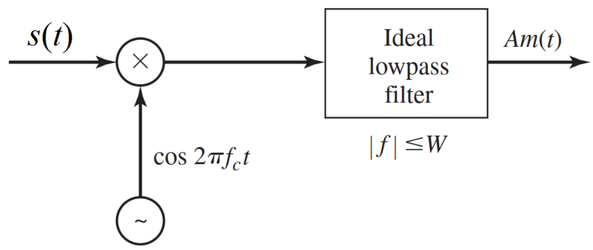


Figure 4.31: Demodulation of VSB signal.

Thus, the product signal is

$$v(t) = s(t) \cos 2\pi f_c t \tag{4.4.3}$$

or equivalently

$$V(f) = \frac{1}{2} [S(f - f_c) + S(f + f_c)] \tag{4.4.4}$$

If we substitute *Eqn.* 4.4.2 into *Eqn.* 4.4.4, we get

$$V(f) = \frac{A_c}{4} [M(f - 2f_c) + M(f)] H(f - f_c) + \frac{A_c}{4} [M(f) + M(f + 2f)] H(f + f_c) \tag{4.4.5}$$

The lowpass filter rejects the double-frequency terms and passes only the components in the frequency range $|f| \leq W$. Hence, the signal spectrum at the output of the ideal lowpass filter is

$$V_l(f) = \frac{A_c}{4} M(f) [H(f - f_c) + H(f + f_c)] \tag{4.4.6}$$

It is required that the message signal at the output of the lowpass filter be undistorted.
Hence, the VSB filter characteristic must satisfy the condition

$$H(f - f_c) + H(f + f_c) = \text{ constant}, \qquad |f| \leq W \tag{4.4.7}$$

This condition is satisfied by a filter that has the frequency-response characteristic shown in *Figure* 4.32. We note that $H(f)$ selects the upper sideband and a vestige of the lower sideband. It has odd symmetry about the carrier frequency f_c, in the frequency range $f_c - f_a < f < f_c + f_a$, where f_a is a conveniently selected frequency that is some small fraction of W; i.e. $f_a \ll W$. Thus, we obtain an undistorted version of the transmitted signal.

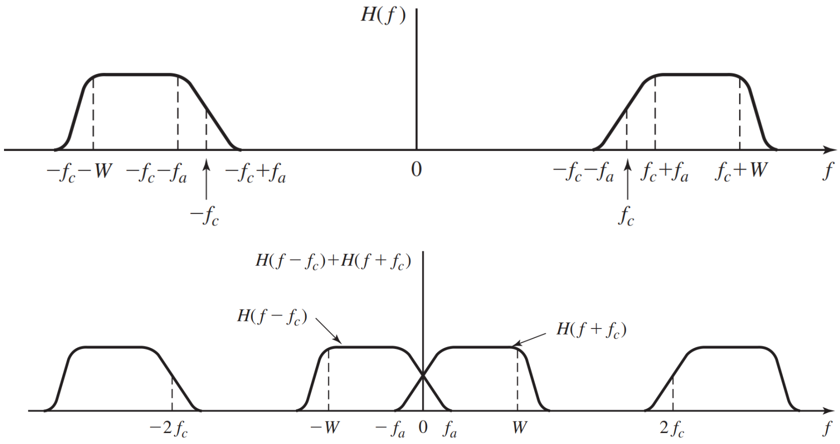


Figure 4.32: VSB filter characteristics

Figure 4.33 illustrates the frequency response of a VSB filter that selects the lower sideband and a vestige of the upper sideband. In practice, the VSB filter is designed to have some specified phase characteristic. To avoid distortion of the message signal, the VSB filter should be designed to have linear phase over its passband $f_c - f_a \leq |f| \leq f_c + W$.

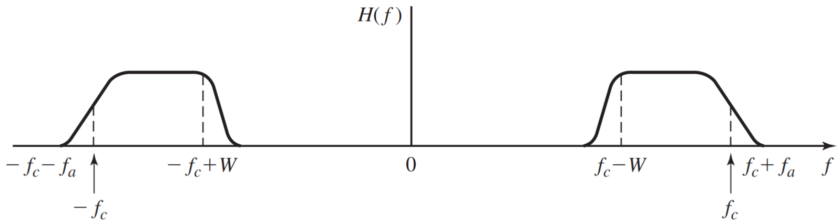


Figure 4.33: Frequency response of VSB filter for selecting the lower sideband of the message signals

Vestigial sideband (VSB) modulation distinguishes itself from SSB modulation in two practical respects:

1. Instead of completely removing a sideband, a trace or vestige of that sideband is transmitted; hence, the name “vestigial sideband”.

2. Instead of transmitting the other sideband in full, almost the whole of this second band is also transmitted.

Accordingly, the transmission bandwidth of a VSB modulated signal is defined by

$$B_T = f_v + W \tag{4.4.8}$$

where f_v is the vestige bandwidth and W is the message bandwidth. Typically, f_v is 25 percent of W, which means that the VSB bandwidth lies between the SSB bandwidth, W, and DSB-SC bandwidth, $2W$.

Example 4.24

Suppose that the message signal is given as

$$m(t) = 10 + 4\cos 2\pi t + 8\cos 4\pi t + 10\cos 20\pi t$$

Specify the frequency-response characteristic of a VSB filter that passes the upper sideband and the first frequency component of the lower sideband.

Solution
The spectrum of the DSB-SC AM signal $s(t) = m(t)\cos 3\pi f_c t$ is

$$\begin{aligned} S(f) = {} & 5\left[\delta(f - f_c) + \delta(f + f_c)\right] + 2\left[\delta(f - f_c - 1) + \delta(f + f_c + 1)\right] \\ & + 4\left[\delta(f - f_c - 2) + \delta(f + f_c + 2)\right] + 5\left[\delta(f - f_c - 10) + \delta(f + f_c + 10)\right] \end{aligned}$$

The VSB filter can be designed to have unity gain in the range $2 \le |f - f_c| \le 10$, a gain of $1/2$ at $f = f_c$ of $1/2 + \alpha$ at $f = f_c + 1$, and a gain of $1/2 - \alpha$, where α is some conveniently selected parameter that satisfies the condition $0 < \alpha < 1/2$.
Figure 4.34 illustrates the frequency-response characteristics of VSB filter.

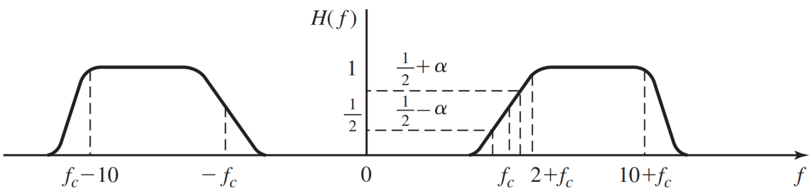


Figure 4.34: Frequency response characteristics of VSB filter

Example 4.25

Let the message signal sum of two sinusoids

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

Obtain the modulated VSB AM signal

Solution
The message signal is multiplied by a carrier signal $\cos\omega_c t$ to obtain DSB-SC signal as

$$\begin{aligned} s(t) &= (A_1\cos\omega_1 t + A_2\cos\omega_2 t)\cos\omega_c t \\ &= \frac{1}{2}A_1\cos(\omega_c - \omega_1)t + \frac{1}{2}A_2\cos(\omega_c - \omega_2)t + \frac{1}{2}A_1\cos(\omega_c + \omega_1)t + \frac{1}{2}A_2\cos(\omega_c - \omega_2)t \end{aligned}$$

Single sideband spectrum from the DSB-SC is

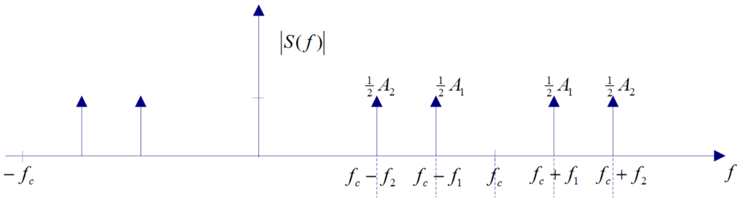


Figure 4.35: Single sideband of DSB-SC signal

A vestigial sideband filter is then used to generate the VSB modulated signal from DSB-SC modulated signal

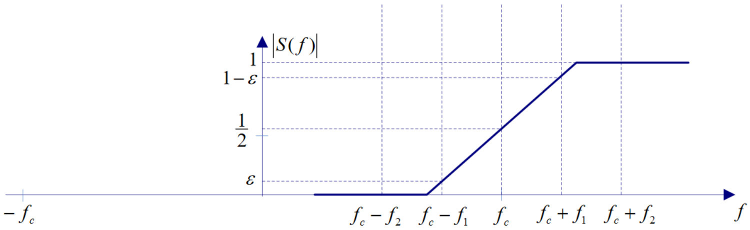


Figure 4.36: VSB modulated signal

The final VSB signal is

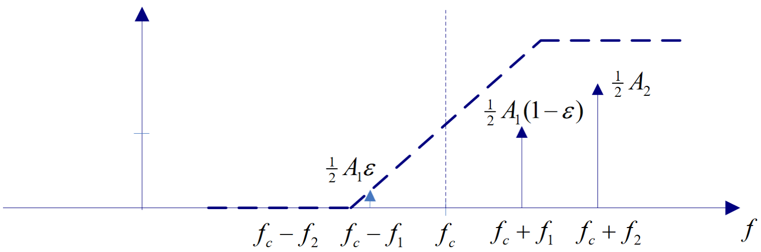


Figure 4.37: VSB modulated signal

The spectrum of VSB signal is

$$VSB_{\text{spectrum}} = |S(f)| \bullet |H(f)|$$

$$v_{VSB}(t) = \frac{1}{2}A_1 \cos(\omega_c - \omega_1)t + \frac{1}{2}A_1(1-\varepsilon)\cos(\omega_c + \omega_1)t + \frac{1}{2}A_1 \cos(\omega_c + \omega_1)t$$

Demodulation

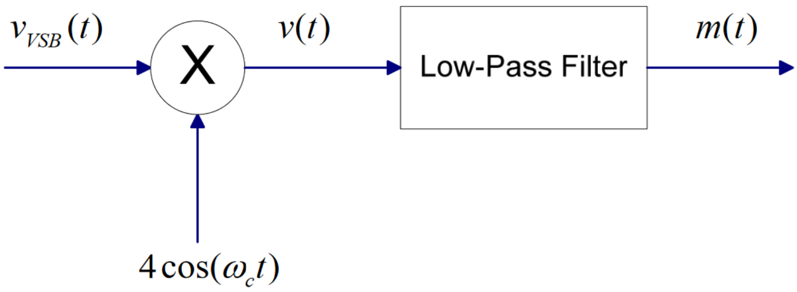


Figure 4.38: Block diagram of VSB signal demodulation

The input of the LPF is

$$\begin{aligned}
v(t) &= v_{VSB}(t) \cdot 4\cos\omega_c t \\
&= 2A_1\varepsilon\cos(\omega_c - \omega_1)t\cos\omega_c t \\
&\quad + 2A_1(1-\varepsilon)\cos(\omega_c + \omega_1)t\cos\omega_c t + 2A_2\varepsilon\cos(\omega_c + \omega_2)t\cos\omega_c t \\
&= [A_1\varepsilon\cos(2\omega_c - \omega_1)t + \cos\omega_c t] + A_1(1-\varepsilon)[\cos(2\omega_c + \omega_1)t + \cos\omega_c t] \\
&\quad + A_2\varepsilon[\cos(2\omega_c + \omega_2)t + \cos\omega_c t]
\end{aligned}$$

After low-pass filter, the required message is

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

4.4.1 Applications of VSB AM

VSB is mainly used in TV broadcasting for their video transmissions. TV signal consists of: Audio signal - is transmitted by FM and Video signal - is transmitted by VSB.

4.5 AM Receivers

4.5.1 Tuned Radio Frequency (TRF) Receiver

TRF receivers are simple to design and construct for low broadcast frequencies but present difficulties at higher frequencies. This is due to the risk of instability resulting from all the gain being achieved at the signal frequency. Tracking also becomes a problem as the frequency increases. In general, the defects of the tuned *RF* receiver are variation of sensitivity and selectivity with frequency over the tuning range.
At broadcast frequencies, the circuit simplicity and ease of alignment are the main advantages.

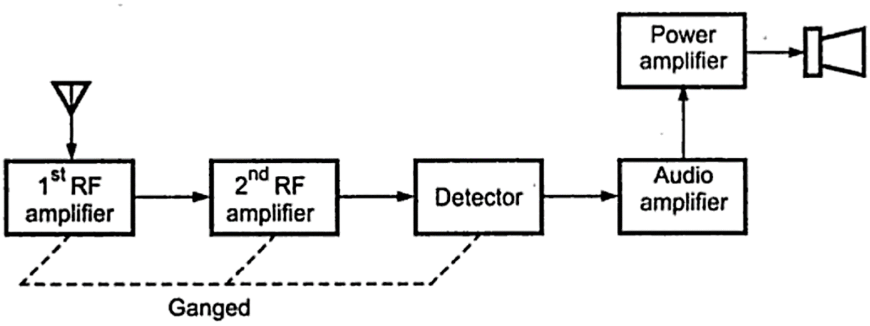


Figure 4.39: Block diagram of Tuned Radio Receiver

4.5.1.1 Problems in TRF

1. Tracking of the Tune circuit

2. Instability

3. Variable Bandwidth
   TRF receivers suffer from a variation in bandwidth over the tuning range. Consider a medium wave receiver required to tune over 535 kHz to 1640 kHz and it provides the necessary bandwidth of 10 kHz at 535 kHz. Let us calculate Q of this circuit.

$$Q = \frac{f}{BW} = \frac{535}{10} = 53.5$$

   Now consider the frequency at the other end of the broadcast band, i.e. 1640 kHz. At 1640 kHz, Q of the coil should be 164 (1640 kHz / 10 kHz). However, in practice due to various losses depending on frequency, it is better to comparatively maintained low Q. Let us assume that at 1640 kHz frequency, Q is increased to value 100 instead of 164. With this, Q of the tuned circuit bandwidth can be calculated as follows

$$BW = \frac{f}{Q} = \frac{1640}{100} = 16.4 \text{ kHz}$$

4.5.2 Superheterodyne AM Receiver

The radio receiver used in an AM system is called the **superheterodyne** AM receiver. This solves the basic problems associated with TRF receivers. In superheterodyne receivers, all incoming RF frequencies are converted to a fix lower frequency called intermediate frequency (IF). Then, the intermediate frequency is amplified and extracted to reproduce the original information. Since the characteristics of the IF amplifier are independent of the frequency to which the receiver is tuned, the selectivity and sensitivity of superheterodyne receivers are fairly uniform throughout its tuning range.

The RF section is basically a tunable filter and an amplifier that picks up the desired station by tuning the filter to the right frequency band. The next section, the frequency mixer (converter), translates the carrier from ω_c to a fixed IF frequency of 455 kHz by using a local oscillator whose frequency f_{LO} is exactly 455 kHz above the incoming carrier frequency f_c; that is, $f_{LO} = f_c + f_{IF}$ ($f_{IF} = 455$ kHz).

Note that this is up-conversion. The tuning of the local oscillator and the RF tunable filter is done by one knob. There is a tuning capacitors in both circuits that are ganged together and are designed to maintain the tuning frequency of the local oscillator at 455 kHz above the tuning frequency of the RF filter. This means every station that is tuned in is translated to a fixed carrier frequency of 455 kHz by the frequency converter. The reason for translating all stations to a fixed carrier frequency of 455 kHz is to obtain adequate selectivity. It is difficult to design sharp bandpass filters of bandwidth 10 kHz (the modulated audio spectrum) if the center frequency f_c is very high. This is particularly true if the filter used is tunable. Hence, the RF filter cannot provide adequate selectivity against adjacent channels. But when the signal is translated to an IF frequency by a converter, it is further amplified by an IF amplifier (usually a three-stage amplifier), which provide good selectivity.
This is because the IF frequency is reasonably low, and, second, its center frequency is fixed and factory-tuned. One important feature of the IF section is the ability to effectively suppress adjacent-channel interference, this only possible due to its high selectivity. For envelope detection, it can amplifies the signal.

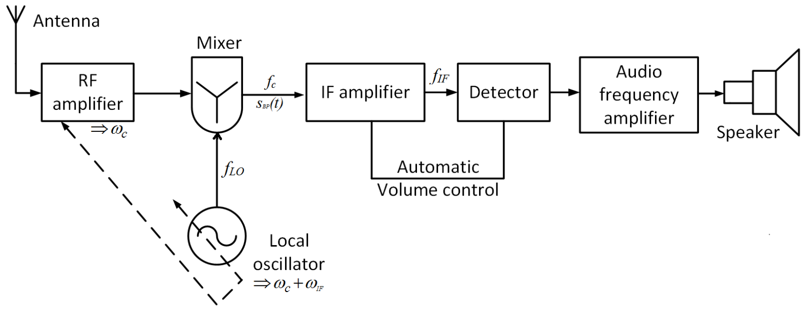


Figure 4.40: Superheterodyne AM Receiver

Practically, all of the selectivity is realized in the *IF* section; the *RF* section plays a negligible role. The main function of the *RF* section is image frequency suppression.
The mixer, or converter, output consists of components of the difference between the incoming (f_c) and the local-oscillator (f_{LO}) frequencies (that is, $f_{IF} = |f_{LO} - f_c|$). Now, if the incoming carrier frequency $f_c = 1000$ kHz, then $f_{LO} = f_c + f_{RF} = 1000 + 455 = 1455$ kHz. But another carrier, with $f'_c = 1455 + 455 = 1910$ kHz, will also be picked up because the difference $f'_c - f_{LO}$ is also 455 kHz. The station at 1910 kHz is said to be the im age of the station of 1000 kHz. Stations that are $2f_{IF} = 910$ kHz apart are called ***image*** stations and would both appear simultaneously at the *IF* output if *RF* filter is not provided at receiver input. The *RF* filter may provide poor selectivity against adjacent stations separated by 10 kHz, but it can provide reasonable selectivity against a station separated by 910 kHz. Thus, when we wish to tune in a station at 1000 kHz, the RF filter, tuned to 1000 kHz, provides adequate suppression of the image station at 1910 kHz.

The receiver (*Figure* 4.40) converts the incoming carrier frequency to the *IF* frequency by using a local oscillator of frequency f_{LO} higher than the incoming carrier frequency (up-conversion) and, hence, is called a superheterodyne receiver. The reason for up-conversion rather than down-conversion is that the former leads to a smaller tuning range (smaller ratio of the maximum to minimum tuning frequency) for the local oscillator than does the latter.

Superheterodyne – simplified
The superheterodyne implement three important frequencies in order to achieve a perfect receiver system. These frequencies are:

1. Radio Frequency (*RF*): – The center frequency is the signal Radio station broadcast on.
2. Intermediate Frequency: – A fixed frequency in the receiver system. The RF signal is downgraded to this frequency
3. Local Oscillator: – This is a tunable frequency in the system used to translate the *RF* signal to the *IF* frequency

In superheterodyne receiver, every AM radio signal is converted to a common *IF* frequency of $f_{IF} = 455$ kHz. This conversion permits the use of a single tuned *IF* amplifier for signals from any radio station in the frequency band. The *IF* bandwidth is made to have 10 kHz which matches the transmitted signal. The conversion is performed by combination of *RF* amplifier and the mixer. The frequency of the local oscillator is

$$f_{LO} = f_c \pm f_{IF} \tag{4.5.1}$$

Upper Sideband: $f_{LO} = f_C + f_{IF}$ Lower Sideband: $f_{LO} = f_c - f_{IF}$. By limiting the bandwidth of the RF amplifier to the range $B_c < B_{RF} < 2f_{IF}$, where B_c is the bandwidth of the AM radio signal (10 kHz), we can reject the radio signal transmitted at the so-called image frequency, $f_c' = f_{LO} + f_{IF}$. Mixing the LO output $2\pi f_{LO}t$ with received signal we get

$$\begin{aligned} r_1(t) &= A_c\,[1+m_1(t)]\cos 2\pi f_c t \\ r_2(t) &= A_c\,[1+m_1(t)]\cos 2\pi f'{}_c t \end{aligned} \tag{4.5.2}$$

where $f_c = f_{LO} - f_{IF}$ and $f_c' = f_{LO} + f_{IF}$, the mixer output also gives

$$\begin{aligned} y_1(t) &= A_c\,[1+m_1(t)]\cos 2\pi f_{IF}t + \text{double freq. term} \\ y_2(t) &= A_c\,[1+m_2(t)]\cos 2\pi f_{IF}t + \text{double freq. term} \end{aligned} \tag{4.5.3}$$

where $m_1(t)$ is the desired signal and $m_2(t)$ represent the signal transmitted by the radio station transmitting at the carrier frequency. Note that, the primary purpose of the RF filter is to reject any signal that might be present on the image frequency.

4.5.3 Characteristics of Radio Receiver

The performance of a Radio receiver can be measured in terms of the following receiver characteristics: Selectivity, Sensitivity, Fidelity, Image frequency and its rejection, and Double spotting

4.6 Review Questions

1. An amplitude wave has the form

$$s_{AM}(t) = 5\,(1 + 0.5\cos 1000\pi t + 0.5\cos 2000\pi t)\cos 10000\pi t$$

   a) Sketch the amplitude spectrum of $s_{AM}(t)$
   b) Find the total power, the sideband power and the power efficiency assuming unit-ohm resistive load
   c) Obtain the modulation index

2. *Figure* 4.41, we show the amplitude spectrum of real single-tone AM signal. Obtain the modulation index and efficiency of this signal.

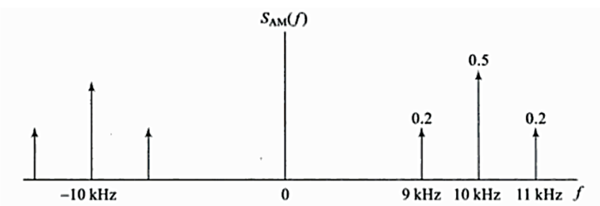


Figure 4.41: Amplitude spectrum of the sine wave modulation AM signal

3. Obtain the expression for the modulation efficiency of a general AM signal. Hence show that the maximum modulation efficiency of a square-wave modulating signal is 50

4. Show that an AM signal can be demodulated using coherent detection. Use *Figure* 4.42

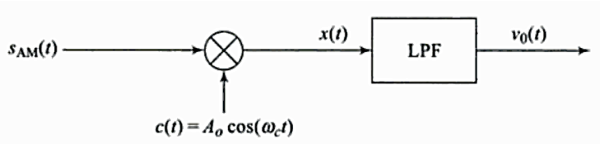


Figure 4.42: Coherent detector

5. An AM signal $s(t) = A_c[1+\mu\cos\omega_m t]\cos\omega_c t$ is passed through a filter which removes the lower sideband. If the resulting signal is applied to an envelope detector, determine
   (a) The expression for the detector output voltage
   (b) The ratio of the fundamental to the second harmonic in the output voltage

6. A conventional AM signal is of the form

$$\psi_{AM}(t) = [1+\alpha\cos\omega_m t+\alpha\cos 2\omega_m t]\cos\omega_c t; \qquad \alpha > 0$$

   Show that to avoid distortion, $\alpha \leq \dfrac{8}{9}$

7. Consider the periodic message signal shown in *Figure* 4.43. This signal is impressed (full AM) upon a carrier $c(t) = A\cos\left(2\pi\times 10^7 t\right)$ with 50 % modulation index.
   (i) Sketch the AM wave form and its spectrum
   (ii) If demodulation is performed using a conventional envelope detector with ideal diode with $\omega_o = \dfrac{1}{RC} = 10$ rad/sec, sketch the output of the envelope detector (time domain) for the following cases using MATLAB

$\alpha = 0.25$, $\beta = -0.5$	$\alpha = 0.25$, $\beta = 0$	$\alpha = 0.25$, $\beta = 1$
$\alpha = 0.5$, $\beta = -0.5$	$\alpha = 0.5$, $\beta = 0$	$\alpha = 0.5$, $\beta = 1$
$\alpha = 0.75$, $\beta = -0.5$	$\alpha = 0.75$, $\beta = 0$	$\alpha = 0.75$, $\beta = 1$

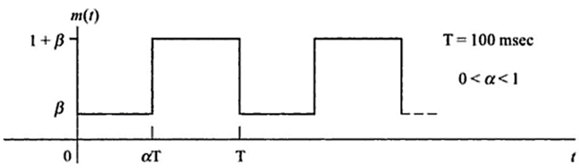


Figure 4.43: Periodic square wave

8. The message signal $m(t) = 2\cos 400t + 4\sin(500t + \pi/3)$ modulates the carrier signal $c(t) = A\cos(8000\pi t)$, using DSB amplitude modulation. Find the time domain and frequency domain representation of the modulated signal and plot the spectrum (Fourier transform) of the modulated signal. What is the power content of the modulated signal?

9. The tuned circuit of the oscillator in a simple AM transmitter employs a 40 μH coil and 12 nF capacitor. If the oscillator output is modulated by audio frequency of 5 kHz, what are tire lower and upper sideband frequencies and the bandwidth required to transmit this AM wave?

10. A sinusoidal carrier $e_o = 100\cos(2\pi 15^5 t)$ is amplitude modulated by a sinusoidal voltage $e_m = 50\cos(2\pi 10^3 t)$ up to a modulation depth of 50%. Calculate the frequency and amplitude of each sideband and rms voltage of the modulated carrier.

11. A carrier wave of a frequency of 20 kHz is amplitude-modulated by a modulating signal $f(t) = \cos 2\pi 10^3 t + \cos 4\pi 10^3 t$. Find the expression for the corresponding SSB-SC signal.

12. Show that if the output of the phase-shift modulator is an SSB signal

    (i) The difference of the signals at the summing junction produces the USB SSB signal and

    (ii) The sum produces the LSB SSB signal.

13. A SSB AM signal is generated by modulating an 800kHz carrier by the signal $m(t) = \cos 2000\pi t + 2\sin 2000\pi t$. The amplitude of the carrier is $A_c = 100$.

    (a) Determine the signal $\hat{m}(t)$

    (b) Determine the (time domain) expression for the lower sideband of the SSB AM signal.

    (c) Determine the magnitude spectrum of the lower sideband SSB signal.

14. The efficiency η of ordinary AM is defined as the percentage of the total power carried by the sidebands, that is $\eta = \frac{P_s}{P_t} \times 100\%$
    Where Ps is the power carried by the sidebands and P_t is the total power of the AM signal.

    a) Find η for $\mu = 0.5$.

    b) Show that for a single-tone AM, η_{max} is 33.33 % at $\mu = 1$.

15. The output signal from an AM modulator is $u(t) = 5\cos 1800\pi t + 20\cos 2000\pi t + 5\cos 2200\pi t$

    i) Determine the modulating signal $m(t)$ and the carrier $c(t)$?

    ii) Determine the modulation index?

    iii) Determine the ratio of the power in the sidebands to the power in the carrier?

16. A Certain AM transmitter is coupled to an antenna. The input power to the antenna is measured although monitoring of the input current, when there is no modulation, the current is 10.8 A. With modulation, the current rises to 12.5A.
    Determine the depth of modulation.

17. A sub-carrier of 70 kHz is amplitude- modulated by tones of 2.1 and 6.8 kHz. The resulting AM signal is then used to amplitude-modulate a carrier of 12.5 MHz. Calculate all sideband frequencies in the composite signal, and draw a frequency-domain display of the signal.
    Assume 100 percent modulation. What is the band-width occupied by the complete signal?

18. The baseband signal $m(t)$ in the frequency-translated signal $v(t) = m(t)\cos 2\pi f_c t$ is recovered by multiplying $v(t)$ by the waveform $\cos(2\pi f_c t + \theta)$.

(a) The product waveform is transmitted through a low-pass filter which rejects the double frequency signal. What is the output signal of the filter?

(b) What is the maximum allowable value for the phase θ if the recovered signal is to be 90 percent of the maximum possible value?

(c) If the baseband signal $m(t)$ is bandlimited to 10 kHz, what is the minimum value of f_c for which it is possible to recover $m(t)$ by filtering the product waveform $v(t)\cos(2\pi f_c t+\theta)$?

19. An AM signal has the equation

$$v(t) = (15 + 4\sin 44 \times 10^3 t) sin 46.5 \times 10^6 t \text{ V}$$

(a) Find the carrier frequency.

(b) Find the frequency of the modulating signal.

(c) Find the value of m.

(d) Find the peak voltage of the unmodulated carrier.

(e) Sketch the signal in the time domain showing voltage and time scales

20. The system shown in *Figure* 4.44 is used for scrambling audio signals. The output $y(t)$ is the scrambled version of the input $m(t)$.

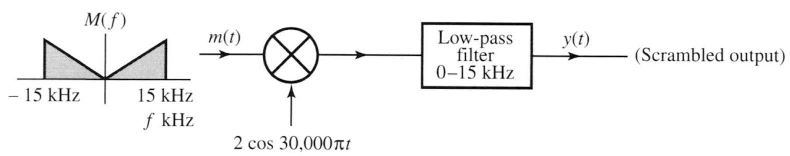
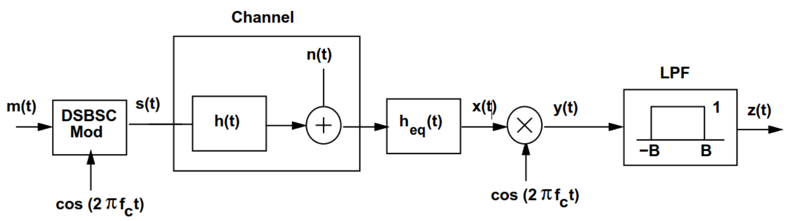


Figure 4.44: Audio scrambler

(a) Find the spectrum of the scrambled signal $y(t)$.

(b) Suggest a method for descrambling $y(t)$ to obtain $m(t)$

21. Consider the DSB-SC system shown in the Figure below, where $s(t) = m(t)\cos 2\pi f_c t$ with $f_c \gg B$.



Assume that the noise $n(t)$ has power spectral density $S_n(f) = 0.1$ mW/Hz. The PSD $S_m(f)$ of $m(t)$ in mW/Hz is

$$S_m(f) = \begin{cases} 10 - 10|f|/B, & |f| \leq B \\ 0, & |f| > B \end{cases}$$

The frequency response of the channel $H(f)$ is

$$H(f) = \begin{cases} 10, & |f - f_c| \le B/2 \\ 0.5, & B/2 < |f - f_c| < B \\ 0, & \text{otherwise} \end{cases}$$

(a. What is the power of $m(t)$.

(b. Sketch the PSD of the modulated signal $s(t)$.

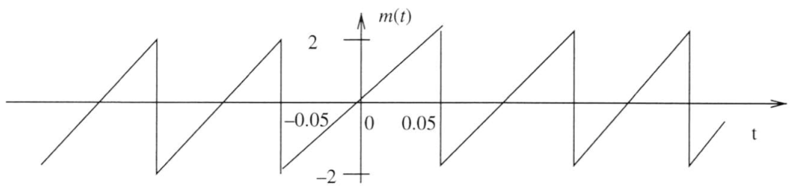
(c. Find the equalizer $H_{eq}(f)$ such that in the absence of noise (i.e., for $n(t) = 0$), $z(t) = m(t)$.

(d. Find the PSD and power of $z(t)$ due to noise only (i.e., for $m(t) = 0$), and due to signal only (i.e., for $n(t) = 0$), using $H_{eq}(f)$ is the equalizer you found in part (c).

(e. Find the SNR of the receiver output with the equalizer of part (c).

22. In an amplitude modulation system, the message signal is given by *Figure* 22 and the carrier frequency is 1 KHz. The modulator output is

$$s_{\rm AM}(t) = 2(b + 0.5m(t))\cos\omega_c t\,.$$



(a. Determine the average message power.

(b. If $b = 1$, determine the modulation index and the modulation power efficiency.

(c. If $b = 0.5$, repeat parts (a) and (b)

4.7 Textbooks and References

[1] Principles of Communication Systems by Taub and Schilling,2nd Edition. McGraw Hill.

[2] Communication Systems by Siman Haykin, 4th Edition, John Wiley and Sons Inc.

[3] Modern digital and analog communication system, by B. P. Lathi, 3rd Edition, Oxford University Press.

[4] Digital and analog communication systems, by L. W. Couch, 6th Edition, Pearson Education, Pvt. Ltd

[5] Introduction to Analog and Digital Communications, Second Edition, by Simon Haykin and Michael Moher

[6] Analog Communications, First Edition 2009, A. P. Godse and U. A. Bakshi

[7] Analog Communication(Jntu) By Thomas-Chandrasekhar

[8] Communication Systems Engineering, Second Edition by John G. Proakis and Masoud Salehi

[9] Introduction to Analog and Digital Communications, Second Edition by Simon Haykin and Michael Moher

[10] Signals and Systems by Alan V. Oppenheim

Chapter 5

Angle (Exponential) Modulation Techniques

***

> Worrying does not empty tomorrow of its troubles, it empties today of its strength.
>
> – Corrie ten Boom

This is another method of modulating a sinusoidal carrier wave – using the phase angle. Modulation in which either the phase or frequency of the carrier wave is varied according to the message signal. In frequency-modulation (FM) systems, the frequency of the carrier f_c is changed by the message signal and in phase-modulation (PM) systems the phase of the carrier is changed according to the variations in the message signal. Frequency and phase modulation are obviously quite nonlinear, and very often they are jointly referred to as **angle-modulation** methods. Frequency and phase-modulation systems generally expand the bandwidth such that the *effective bandwidth* of the modulated signal is usually many times the bandwidth of the message signal.

The major benefit of these systems is their high degree of noise immunity. In fact these systems trade-off bandwidth for high noise immunity. That is the reason that FM systems are widely used in high-fidelity music broadcasting and point-to-point communication systems where the transmitter power is quite limited. The amplitude and power of an angle-modulation signal do not change with modulation. Thus, an FM signal has no envelope. This is actually an advantage; an FM receiver does not have to respond to amplitude variations, and this lets it ignore noise to some extent. Similarly, FM equipment can use nonlinear amplifiers throughout, since amplitude linearity is not important.

5.1 Basic Definitions of Angle Modulation

Let $\theta_i(t)$ denote the angle or **instantaneous phase** of a modulated sinusoidal carrier at time t; it is assumed to be a function of the information-bearing signal or message signal. We express the resulting angle-modulated wave as

$$s(t) = A_c \cos\left[\theta_i(t)\right] \tag{5.1.1}$$

If $\theta_i(t)$ increases monotonically with time, then the average frequency over small interval from t to $t+\Delta t$ is given by

$$f_{\Delta t}(t) = \frac{\theta_t\left(t+\Delta t\right) - \theta_i(t)}{2\pi\Delta t} \tag{5.1.2}$$

Allowing Δt to approach zero leads to the definition of *instantaneous frequency* of the angle modulated signal $s(t)$

$$\begin{aligned} f_i(t) &= \lim_{\Delta t \to 0} f_{\Delta t}(t) \\ &= \lim_{\Delta t \to 0} \left[\frac{\theta_t\left(t+\Delta t\right) - \theta_i(t)}{2\pi\Delta t}\right] = \frac{1}{2\pi}\frac{d\theta_i(t)}{dt} \end{aligned} \tag{5.1.3}$$

From *Eqn.* 5.1.1, $s(t)$ can be interpreted as rotating phasor of length A_c and angle $\theta i(t)$. The angular velocity of such a phasor is $d\theta_i(t)\big/dt$ measured in radians per second. And the corresponding phasor rotates with a constant angular velocity equal to radians per second. The constant ϕ defines the angle of the unmodulated carrier at time, $t=0$. The instantaneous of $s(t)$ is given by

$$\theta_i(t) = 2\pi f_c t + \phi(t) \qquad \text{for } m(t) = 0 \tag{5.1.4}$$

Hence, *Eqn.* 5.1.3 becomes

$$f_i(t) = f_c + \frac{1}{2\pi}\frac{d}{dt}\phi(t) \tag{5.1.5}$$

The two common ways in which $\theta_i(t)$ can be vary in angle modulation method are **phase modulation (PM)** and **frequency modulation (FM)**. If $m(t)$ is the message signal, then PM and FM are defined as describe in subsequent sections.

5.2 Phase Modulation Systems (PM)

PM is such that the angle modulation in which the instantaneous angle $\theta_i(t)$ is varied linearly with the message signal $m(t)$ given by

$$\theta_i(t) = 2\pi f_c t + k_p m(t) \tag{5.2.1}$$

where $\phi(t) = k_p m(t)$. The term $2\pi f_c t$ represent the angle of the unmodulated carrier. k_p represent the *phase sensitivity factor* of the modulator or more popularly **phase deviation constant**, expressed in radians per volt on the assumption that $m(t)$ is a voltage waveform.
The phase modulated wave $s(t)$ is

$$s_{PM}(t) = A_c \cos\left[2\pi f_c t + k_p m(t)\right] \tag{5.2.2}$$

5.3 Frequency Modulation (FM)

Similarly, FM implies that the frequency deviation of the carrier is proportional to the modulating signal. In other words, it is the form of angle modulation in which the instantaneous frequency $f_i(t)$ is varied linearly with the message signal $m(t)$ given by

$$f_i(t) = f_c + k_f m(t) \tag{5.3.1}$$

where

$$k_f m(t) = f_i(t) - f_c = \frac{1}{2\pi}\frac{d}{dt}\phi(t) \tag{5.3.2}$$

The constant term f_c represents the frequency of the unmodulated carrier. k_f represent the *frequency sensitivity factor* of the modulator or more popularly **frequency deviation constant**, expressed in hertz per volt on the assumption that $m(t)$ is a voltage waveform.

Integrating *Eqn.* 5.3.1 with respect to time and multiplying the results by 2π yields

$$\begin{aligned}\theta_i(t) &= 2\pi\int_{0|-\infty}^{t} f_i(\tau)d\tau \\ &= 2\pi f_c t + k_f\int_{0\ |\ -\infty}^{t} m(\tau)d\tau\end{aligned} \tag{5.3.3}$$

where the second term accounts for the increase or decrease in the instantaneous phase $\theta_i(t)$ due to the message signal $m(t)$. The lower limit of the integration indicates that the integration can either start from zero or minus infinity.
The frequency modulated wave is therefore given by

$$s_{FM}(t) = A_c\cos\left[2\pi f_c t + k_f\int_{0\ |\ -\infty}^{t} m(\tau)d\tau\right] \tag{5.3.4}$$

Example 5.26

An 18 MHz carrier is modulated by a 400 Hz audio sine wave if the carrier voltage is 5.0 V and the maximum deviation is 12 kHz. Write down the equation for the modulated PM and FM wave.

Solution
Carrier frequency

$$\omega_c = 2\pi f_c = 2\pi\times 18\times 10^6 = 1131.1\times 10^6 \text{ rad/sec}$$

Modulating frequency

$$\omega_m = 2\pi f_m = 2\pi\times 400 = 2513.3\text{rad/sec}$$

The modulation index

$$\begin{aligned}\beta = \beta_f = \beta_p &= \frac{\Delta f}{f_m} \\ &= \frac{12\times 10^3}{400} = 30\end{aligned}$$

Hence, the required equations for PM and FM are

$$\begin{aligned} s_{PM}(t) &= V_{c\max} \sin\left[\omega_c t + \beta_p \sin \omega_m t\right] \\ &= 5 \sin\left[1131.1 \times 10^6 t + 30 \sin 2513.3t\right] \end{aligned}$$

Now, equation for FM

$$\begin{aligned} s_{FM}(t) &= V_{c\max} \sin\left[\omega_c t + \beta_f \sin \omega_m t\right] \\ &= 5 \sin\left[1131.1 \times 10^6 t + 30 \sin 2513.3t\right] \end{aligned}$$

When the $f_m = 1.6\,\text{kHz}$, $\beta_f = 7.5$ but β_p remain unchanged. The new equation for FM based on the new f_m is

$$s_{FM}(t) = 5 \sin\left[1131.1 \times 10^6 t + 7.5 \sin 2513.3t\right]$$

5.4 Relationship Between FM and PM Waves

The demodulation of an FM signal involves finding the instantaneous frequency of the modulated signal and then subtracting the carrier frequency from it. In the demodulation of PM, the demodulation process is done by finding the phase of the signal and then recovering $m(t)$. The maximum phase deviation in a PM system is given by

$$\Delta\phi_{\max} = k_p \max\left[|m(t)|\right] \tag{5.4.1}$$

and the maximum frequency-deviation in an FM system is given by

$$\Delta f_{\max} = k_f \max\left[|m(t)|\right] = k_f \frac{A_m}{2\pi} \tag{5.4.2}$$

or

$$\Delta f_{\max} = \max\left\{\frac{1}{2\pi}\left[\frac{d\phi(t)}{dt}\right]\right\} \tag{5.4.3}$$

The analysis on *Section* 5.2 and 5.3 leads to interesting result observation that is, if we phase modulate the carrier with the integral of a message, it is equivalent to frequency modulation of the carrier with the original message. PM and FM can be related using the integral of phase angle ϕ

$$\phi(t) = \begin{cases} k_p m(t), & \text{PM} \\ k_f \int_{-\infty}^{t} m(\tau)d\tau, & \text{FM} \end{cases}, \quad \frac{d}{dt}\phi(t) = \begin{cases} k_p \frac{d}{dt} m(t), & \text{PM} \\ 2\pi k_f m(t), & \text{FM} \end{cases} \tag{5.4.4}$$

If $m(t) = A_m cos \omega_m t$, then

$$\phi(t) = \begin{cases} k_p A_m \cos \omega_m t, & \text{for PM} \\ \frac{k_f A_m}{\omega_m} \sin \omega_m t, & \text{for FM with } \phi(-\infty) = 0 \end{cases} \tag{5.4.5}$$

Eqn. 5.4.4 shows that if we frequency modulate the carrier with the derivative of a message, the result is equivalent to phase modulation of the carrier with the message itself. The relation between PM and FM is shown in *Figure* 5.1.

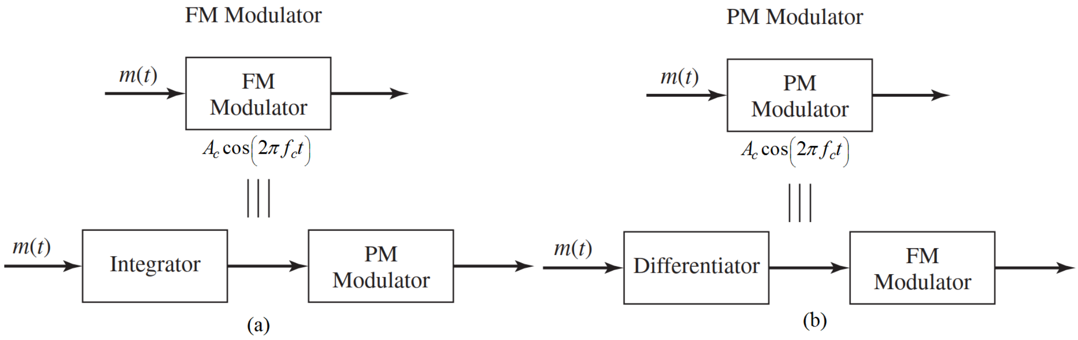


Figure 5.1: Illustration of the relationship between frequency modulation and phase modulation. (a) Scheme for generating an FM wave by using a phase modulator (b) Scheme for generating a PM wave by using a frequency modulator

It follows therefore that phase modulation and frequency modulation are uniquely related to each other. This relationship, in turn, means that we may deduce the properties of phase modulation from those of frequency modulation and vice versa. For this reason, in this chapter we will be focusing much of the discussion on frequency modulation. *Figure* 5.2 shows comparison of AM, FM and PM waves using square wave as message signal.

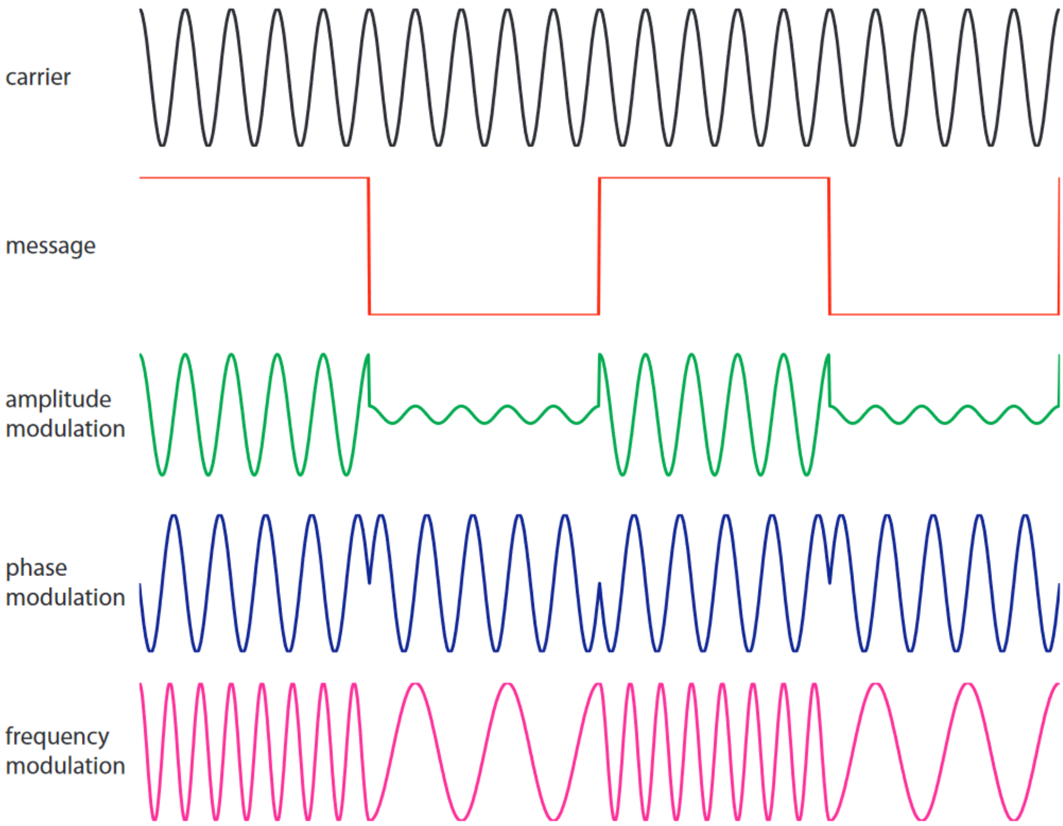


Figure 5.2: Illustration of AM, PM, and FM waves produced by a single tone

5.4.1 Modulation Indexes of PM and FM

Example 5.27 best describes the derivation of both PM and FM modulation indexes.

Example 5.27

Consider the message signal $A_m \cos(2\pi f_m t)$ is used to either frequency modulate or phase modulate the $A_c \cos(2\pi f_c t)$. Find the modulated signal in each case.

Solution
In PM, we have

$$\phi(t) = k_p m(t) = k_p A_m \cos(2\pi f_m t)$$

and in FM,

$$\begin{aligned}\phi(t) &= k_f \int_{-\infty}^{t} m(\tau) d\tau \\ &= k_f \int_{-\infty}^{t} A_m \cos(2\pi f_m \tau)\, d\tau \\ &= \frac{k_f A_m}{\omega_m} \sin(2\pi f_m t)\end{aligned}$$

The modulated signal for both is

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t + k_p A_m \cos(2\pi f_m t)), & \text{PM} \\ A_c \cos\left(2\pi f_c t + \frac{k_f A_m}{\omega_m} \sin(2\pi f_m t)\right), & \text{FM} \end{cases} \tag{5.4.6}$$

Hence by observation and variable definition

$$\beta_p = k_p A_m \tag{5.4.7}$$

and

$$\beta_f = \frac{k_f A_m}{\omega_m} = \frac{\Delta f}{\omega_m} \tag{5.4.8}$$

where β_p and β_f are modulation indexes of PM and FM systems respectively.
The modulated wave using the modulation indexes is

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t + \beta_p \cos(2\pi f_m t)), & \text{PM} \\ A_c \cos(2\pi f_c t + \beta_f \sin(2\pi f_m t)), & \text{FM} \end{cases} \tag{5.4.9}$$

We can extend the definition of the modulation indexes for a general nonsinusoidal signal $m(t)$ as

$$\beta_p = k_p \max[|m(t)|] \qquad \Rightarrow \text{PM} \tag{5.4.10}$$

$$\beta_f = \frac{k_f \max[|m(t)|]}{W} \qquad \Rightarrow \text{FM} \tag{5.4.11}$$

Note that W is the frequency to the modulating signal $m(t)$ analogous to f_m.
In terms of maximum phase and frequency deviation $\Delta\phi_{\max}$ and $\Delta f_{\max}$, we have

$$\beta_p = \Delta\phi_{\max} \tag{5.4.12}$$

$$\beta_f = \frac{\Delta f_{\max}}{W} = \frac{\Delta \omega_{\max}}{W} \tag{5.4.13}$$

Note the following features of angle modulated signal:

I. For FM signal, the maximum frequency deviation takes place when modulating signal is at positive and negative peaks.

II. For PM signal the maximum frequency deviation takes place near zero crossings of the modulating signal.

III. Both FM and PM waveforms are identical except the phase shift.

IV. It is difficult to identify PM or FM from angle modulated waveform

V. Depending on the value of β (the modulation index), for FM, we either have **Narrowband FM**, where $\beta_f \ll 1$ or **Wideband FM**, where $\beta_f \approx 1$.

Example 5.28

Let $s(t)$ be a general angle signal given by

$$s(t) = A_c \cos[\theta_i(t)] = A_c \cos[\omega_c t + \phi(t)]$$

It is given that when $m(t) = cos\omega_m t$, $s(t)$ has an instantaneous frequency given by $f_i(t) = f_c + 2\pi k(f_m)^2 cos\omega_m t$, where k is a suitable constant.
(a) Find the expression for $\theta_i(t)$. (b) If $m(t)$ is different from $cos\omega_m t$, what is the expression of $\theta_i(t)$ and $s(t)$?

Solution
(a) Using *Eqns.* 5.1.3 and 5.1.5,

$$f_i(t) = \frac{1}{2\pi}\frac{d\theta_i(t)}{dt} = f_c + 2\pi k(f_m)^2 \cos\omega_m t$$
$$\frac{d\theta_i(t)}{dt} = 2\pi f_c + \omega_m^2 k \cos(\omega_m t)$$

Integrating w.r.t. t, $\theta_i(t)$ can be obtained as

$$\theta_i(t) = \int d\theta_i(t)dt = 2\pi f_c t + \omega_m k \sin(\omega_m t)$$
$$= 2\pi f_c t - k\frac{d}{dt}[\cos(\omega_m t)]$$

(b) Generally,

$$\theta_i(t) = 2\pi f_c t - k\frac{d}{dt}[m(t)]$$

and

$$s(t) = A_c \cos\left[2\pi f_c t - k\frac{dm(t)}{dt}\right]$$

5.5 Properties of Angle Modulated Waves

Angle-modulated waves are characterized by some important properties, it is these properties that put angle-modulated waves in a family of their own, and distinguish them from the family of amplitude-modulated waves, as illustrated by the waveforms shown in *Figure* 5.2 sinusoidal modulation.

5.5.1 Constancy of Power

From both *Eqn*. 5.3.4 and 5.2.2, it is readily seen that the amplitude of PM and FM waves is maintained at a constant value equal to the carrier amplitude for all time t, irrespective of the sensitivity factors k_p and k_f or modulation indexes, β_p and β_f.
Consequently, the average transmitted power of angle-modulated waves is a constant, as shown by

$$P_{av} = \frac{1}{2}A_c^2 \tag{5.5.1}$$

where it is assumed that the load resistor is 1 ohm as usual in communication systems.

5.5.2 Nonlinearity of the Angle Modulation

PM and FM waves violate the principle of superposition. Suppose, for example, that the message signal $m(t)$ is made up of two different components and as shown by

$$m(t) = m_1(t) + m_2(t) \tag{5.5.2}$$

Let $s_{PM}(t)$, $s_{PM_1}(t)$ and$s_{PM_2}(t)$ denote the PM waves produced by $m(t)$, $m_1(t)$ and $m_2(t)$ accordance with *Eqn.* 5.2.2, respectively. Then *Eqn.* 5.2.2 may be express as these PM waves

$$s_{PM}(t) = A_c \cos\left[2\pi f_c t + k_p\left(m_1(t) + m_2(t)\right)\right] \tag{5.5.3}$$

where

$$s_{PM_1}(t) = A_c \cos\left[2\pi f_c t + k_p m_1(t)\right] \tag{5.5.4}$$

and

$$s_{PM_2}(t) = A_c \cos\left[2\pi f_c t + k_p m_2(t)\right] \tag{5.5.5}$$

From these expressions, it is readily noticed that

$$s_{PM}(t) = s_{PM_1}(t) + s_{PM_2}(t) \tag{5.5.6}$$

Other properties include:

I. Visualization difficulty of message waveform

II. Trade-off of increased transmission bandwidth for improved noise performance

5.6 Spectra of Angle Modulated Signals

The angle modulated signal, $s(t) = A_c \cos\left[\theta_i(t)\right]$ can be put in the complex exponential form $g(t) = A_c e^{j\theta_i(t)}$, which is a nonlinear function of $m(t)$ Taking the Fourier transform of $s(t)$, we have

$$S(f) = \frac{1}{2}\left[G(f - f_c) + G^*(-f - f_c)\right] \tag{5.6.1}$$

where

$$G(f) = F\left[g(t)\right] = F\left[A_c e^{j\theta_i(t)}\right] \tag{5.6.2}$$

5.7 Wideband Frequency Modulation (WBFM)

WBFM is obtain when the modulation index of angle modulationd signal is approximately equal to unity, that is

$$\beta \approx 1 \tag{5.7.1}$$

A WBFM signal has theoretically infinite bandwidth. Spectrum calculation of WBFM signal is a tedious process. For, practical applications however the Bandwidth of a WBFM signal is calculated as using Carson's rule (*Eqn.* 5.7.11). The following subsection discusses WBFM into details.

5.7.1 Spectrum of Angle Modulation by Sinusoidal Signal

Assume that the modulation on the PM signal is

$$m_p(t) = A_m \sin \omega_m t$$

Then

$$\theta(t) = \beta \sin \omega_m t$$

where $\beta_p = k_p A_m = \beta$ is the phase modulation index.
The same phase function $\theta(t)$ can be used in FM modulated signal with the message

$$m_f(t) = A_m \cos \omega_m t$$

and $\beta = \beta_f = k_f A_m / \omega_m$ and maximum deviation as $\Delta f_{\max} = k_f \frac{A_m}{2\pi}$. Putting together, the modulated signal can be written as

$$s(t) = A_c \cos(2\pi f_c t + \beta \sin 2\pi f_m t) \rightarrow \text{PM} \tag{5.7.2}$$

Or

$$s(t) = \text{Re}\left(A_c e^{j2\pi f_c t} e^{j\beta \sin 2\pi f_m t}\right) \tag{5.7.3}$$

The complex envelope is

$$g(t) = A_c e^{j\theta_i(t)} = A_c e^{j\beta \sin \omega_m t}$$

with period $T_m = 1/f_m$. The Fourier series representation that is valid over all time ($\infty < t < \infty$) is

$$g(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_m t} \tag{5.7.4}$$

where

$$c_n = \frac{A_c}{T_m} \int_{-T_m/2}^{T_m/2} \left(e^{j\beta \sin \omega_m t}\right) e^{-jn\omega_m t} dt$$

which reduces to

$$c_n = A_c \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(\beta \sin \theta - n\theta)} d\theta\right] = A_c J_n(\beta) \tag{5.7.5}$$

Eqn. 5.7.5 is known as the *Bessel function of the first kind of the nth order,*$J_n(\beta)$. The Fourier series for the complex exponential part is

$$e^{j\beta \sin 2\pi f_m t} = \sum_{n=-\infty}^{\infty} J_n(\beta) e^{j2\pi n f_m t} \tag{5.7.6}$$

The modulated signal becomes after substitution of *Eqn.* 5.7.6 into *Eqn.* 5.7.3

$$\begin{aligned} s(t) &= \mathrm{Re}\left[A_c \sum_{n=-\infty}^{\infty} J_n(\beta)\, e^{j2\pi n f_m t} e^{j2\pi f_c t}\right] \\ &= \sum_{n=-\infty}^{\infty} A_c J_n(\beta)\cos\left(2\pi(f_c + n f_m)t\right) \end{aligned} \tag{5.7.7}$$

Fourier transform of *Eqn.* 5.7.7 is

$$S(f) = \frac{A_c}{2}\sum_{n=-\infty}^{\infty} J_n(\beta)\left[\delta\left(f - f_c - n f_m\right) + \delta\left(f + f_c + n f_m\right)\right] \tag{5.7.8}$$

Eqn. 5.7.8 indicates the following:

(i) FM signal has infinite number of sidebands at frequency, $f_c \pm n f_m$

(ii) Relative amplitude of all the spectral lines depends on the value of $J_n(\beta)$

(iii) The number of significant sidebands depends on the modulation index (β). With $(\beta \ll 1)$, only $J_o(\beta)$ and $J_1(\beta)$ are significant. But for $\beta \gg 1$, many sidebands exists.

(iv) The average power of an FM wave is $P = 0.5{A^2}_c$ based on Bessel function property.

A series expansion of the Bessel function is given by

$$J_n(\beta) = \sum_{k=0}^{\infty} \frac{(-)^k \left(\frac{\beta}{2}\right)^{n+2k}}{k!\,(k+n)!}$$

For a small β, we can the approximation below

$$J_n(\beta) \approx \frac{\beta^n}{2^n n!} \tag{5.7.9}$$

Thus for a small modulation index β, only the first sideband corresponding to $n = 1$ is of importance. Examination of the *Eqn.* 5.7.5 shows that (by making a change in variable)

$$J_{-n}(\beta) = (-1)^n J_n(\beta) \tag{5.7.10}$$

A plot of the Bessel functions for various orders n as a function of β is shown in *Figure* 5.3.

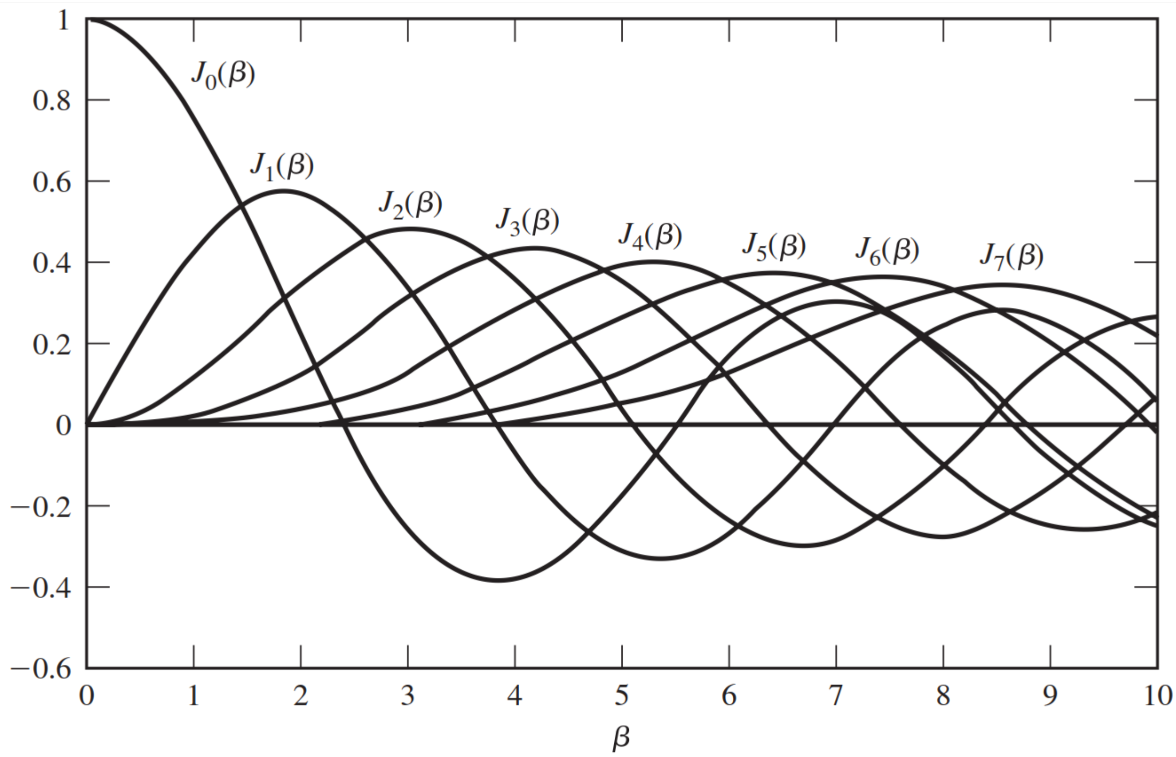


Figure 5.3: Plots of the Bessel function of the first kind, $J_n(\beta)$, for varying order n

The spectrum diagram of $S(f)$ is shown in *Figure* 5.4 for $\beta =$ 0.2, 1.0, 2.0, 5.0, and 8.0.

5.7.2 Carson's Rule and Transmission Bandwidth of FM Signal

Figure 5.4 also shows that the bandwidth of the angle-modulated signal depends on β and f_m. In fact, it can be shown that 98 % of the total power is contained in the bandwidth

$$B_T = 2(\beta + 1)W \tag{5.7.11}$$

where β is either the phase modulation index or the frequency modulation index and W is the bandwidth of the modulating signal (which is f_m for sinusoidal modulation). *Eqn.* 5.7.11 is popularly known as **Carson's rule**.
Other forms of Carson's rule is given as

$$B_T \approx 2\left(\Delta f + f_m\right) = 2\Delta f\left(1 + \frac{1}{\beta}\right) \tag{5.7.12}$$

Note, Δf is frequency deviation. Other forms bandwidth of angle modulated signal is

$$\begin{aligned} B_T &= \begin{cases} 2\left(k_p A_m + 1\right) f_m, & \text{PM} \\ 2\left(\frac{k_f A_m}{f_m} + 1\right) f_m, & \text{FM} \end{cases} \\ &= \begin{cases} 2\left(k_p A_m + 1\right) f_m, & \text{PM} \\ 2\left(k_f A_m + f_m\right), & \text{FM} \end{cases} \end{aligned} \tag{5.7.13}$$

Specifically, the bandwidth required to transmit an FM wave generated by an arbitrary modulating wave is based on a worst-case i.e. tone-modulation analysis. We first determine the so-called deviation ratio D, defined as the ratio of the frequency deviation Δf, which corresponds to the maximum possible amplitude of the modulation wave to the highest modulation frequency W. These conditions represent the extreme cases possible. We may formally write

$$D = \frac{\Delta f}{W} \tag{5.7.14}$$

The deviation ratio D plays the same role for nonsinusoidal modulation that the modulation index plays for the case of sinusoidal modulation.

5.7.3 Properties of Bessel Function

1. For different integer values of n, *Eqn.* 5.7.10 is positive for even values of n and negative for odd values of n.

2. For small values of modulation index β

$$\left.\begin{aligned} J_o(\beta) &\approx 1, \\ J_1(\beta) &\approx \tfrac{\beta}{2}, \\ J_n(\beta) &\approx 0, \qquad n > 2 \end{aligned}\right\} \tag{5.7.15}$$

3. The equality

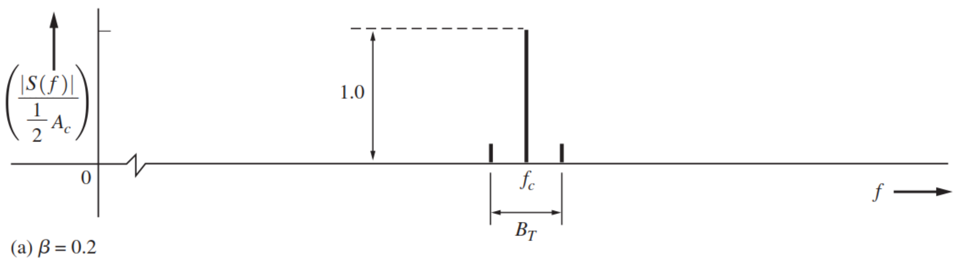
$$\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1 \tag{5.7.16}$$

true for arbitrary β.

4. The average power for FM or PM was given in *Eqn.* 5.5.1 as $P_{av} = \frac{1}{2}A_c^2$. Now, When the carrier is modulated to generate the FM wave, the power in the side-frequencies may appear only at the expense of the power originally in the carrier, thereby making the amplitude of the carrier component dependent on β. Note that the average power of an FM wave may also be determined by

$$P = \frac{1}{2}A_c^2 \sum_{n-\infty}^{\infty} J_n^2(\beta) \tag{5.7.17}$$

Substituting *Eqn.* 5.7.16 into *Eqn.* 5.7.17 yields *Eqn.* 5.5.1



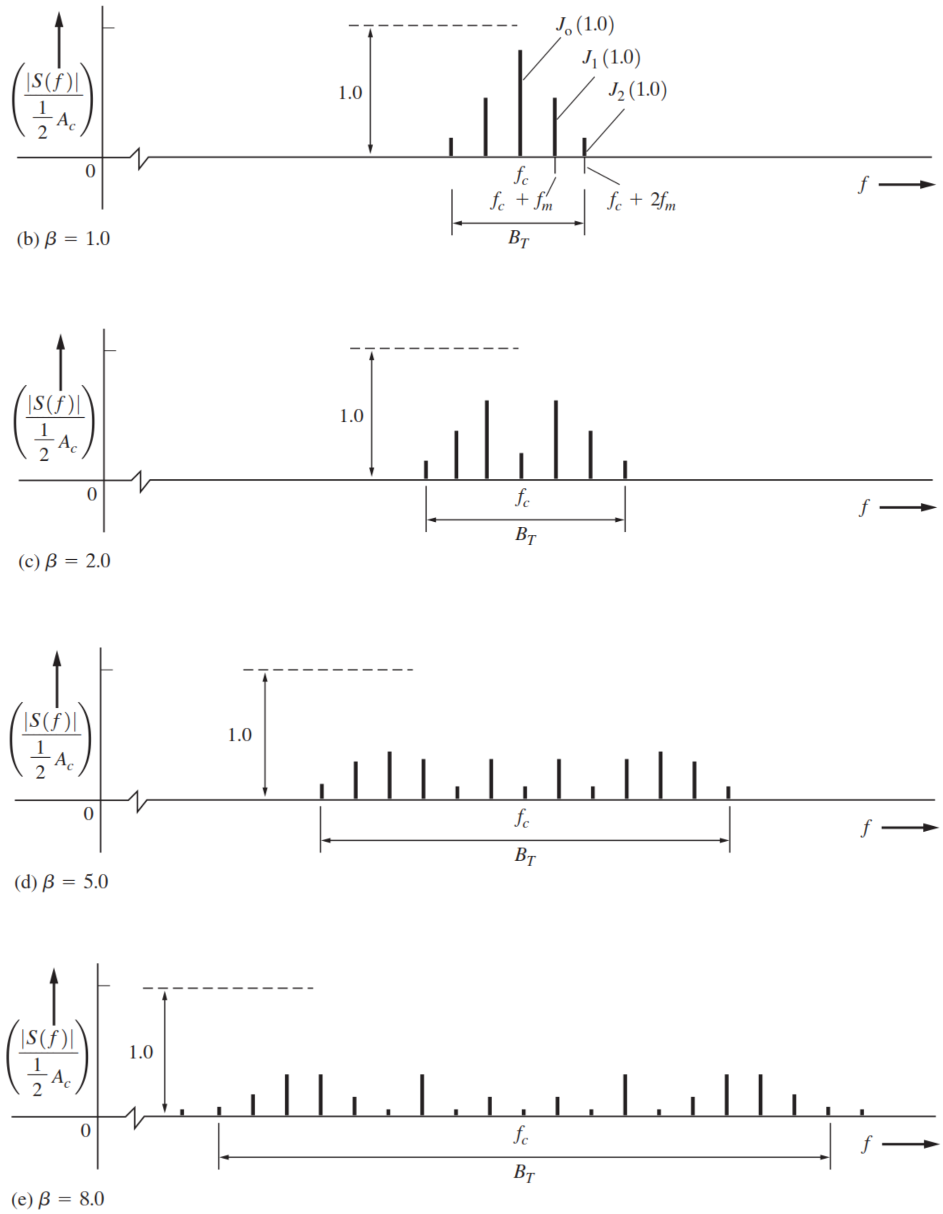


Figure 5.4: Magnitude spectra for FM or PM with sinusoidal modulation for various modulation indexes

Example 5.29

An angle modulated signal is given by

$$s(t) = \cos\left[2\pi\left(2\times 10^6 t\right) + 30\sin(150t) + 40\cos(150t)\right]$$

Find the maximum phase and frequency derivations.

Solution
The terms

$$[30\sin(150t) + 40\cos(150t)]$$

can be expressed in the form

$$[\cos\alpha\sin(150t)+\sin\alpha\cos(150t)]$$

As

$$\sqrt{30^2+40^2}=50$$

We have

$$\begin{aligned}30\sin(150t)+40\cos(150t) &= 50\left[\frac{3}{5}\sin(150t)+\frac{4}{5}\cos(150t)\right]\\ &= 50\sin(150t+\phi)\ \text{ where } \phi=\tan^{-1}\frac{4}{3}\end{aligned}$$

Evidently, the maximum phase deviation is (100π)

Let

$$\psi(t)=100\pi\sin(150t+\phi)$$

, then,

$$\begin{aligned}\frac{1}{2\pi}\frac{d\psi}{dt} &= 50\cos(150t+\phi)\cdot 150\\ &= 7500\cos(150t+\phi)\end{aligned}$$

Hence, maximum frequency deviation = 7500 Hz.

Example 5.30

Find the bandwidth of a single tone modulated FM signal described by

$$s(t)=10\cos\left[2\pi 10^8 t+6\sin(2\pi 10^3 t)\right]$$

Solution

Comparing the given $s(t)$ with the angle modulated signal, the modulation index with $\beta=6$ and message frequency, $f_m=1000$ Hz.
By Carson's rule The transmission bandwidth is

$$\begin{aligned}B_T &= 2(\beta+1)f_m\\ &= 2(7)1000 = 14\text{ kHz}\end{aligned}$$

Example 5.31

A carrier wave of frequency 91 MHz is frequency modulated by a sine wave of amplitude 10 Volts and 15 kHz. The frequency sensitivity of the modulator is 3 kHz/V.

(a) Determine the approximate bandwidth of FM wave using Carson's Rule.

(b) Repeat part (a), assuming that the amplitude of the modulating wave is doubled.

(c) Repeat part (a), assuming that the frequency of the modulating wave is doubled.

Solution

(a) Modulation index

$$\beta = \frac{\Delta\omega}{\omega_m} = \frac{\Delta(2\pi f)}{2\pi f_m} = \frac{\Delta f}{f_m} = \frac{k_f A_m}{f_m} = \frac{3\times 10}{15} = 2$$

By Carson's rule, the bandwidth is

$$B_T = 2\,(\beta+1)\,f_m = 90\text{ kHz}$$

(b) When the amplitude, A_m is doubled.
New modulation index is

$$\beta = \frac{3\times 20}{15} = 4$$

Bandwidth, $B_T = 150$ kHz

(c) When the frequency of the message signal, f_m is doubled. New modulation index is $\beta = 1$. Bandwidth, $B_T = 120$ kHz

Example 5.32

A 20 MHz carrier is frequency modulated by a sinusoidal signal such that the peak frequency deviation is 100 kHz. Determine the modulation index and the approximate bandwidth of the FM signal if the frequency of the modulating signal is: (a) 1 kHz; (b) 50 kHz; (c) 500 kHz.

Solution
We are given $f_\Delta = 100$ kHz and $f_c = 20\text{ MHz} \gg f_m$. For sinusoidal modulation, $f_\Delta = \beta f_m$, that is, $\beta = D = f_\Delta / f_m$.

a. With $f_m = 1000$ Hz, $\beta = f_\Delta/1000 = 10^5/1000 = 100$. This is WBFM and $B_T \approx 2(\beta+1)f_m = 202$ kHz.

b. With $f_m = 50$ kHz, $\beta = 2$ and $B_T \approx 300$ kHz

c. With $f_m = 500$ kHz, $\beta = 0.2$. This is NBFM and $B_T \approx 2f_m = 1$ MHz

TABLE 5–2 FOUR-PLACE VALUES OF THE BESSEL FUNCTIONS $J_n(\beta)$

n \ β:	0.5	1	2	3	4	5	6	7	8	9	10
0	0.9385	0.7652	0.2239	−0.2601	−0.3971	−0.1776	0.1506	0.3001	0.1717	−0.09033	−0.2459
1	0.2423	0.4401	0.5767	0.3391	−0.06604	−0.3276	−0.2767	−0.004683	0.2346	0.2453	0.04347
2	0.03060	0.1149	0.3528	0.4861	0.3641	0.04657	−0.2429	−0.3014	−0.1130	0.1448	0.2546
3	0.002564	0.01956	0.1289	0.3091	0.4302	0.3648	0.1148	−0.1676	−0.2911	−0.1809	0.05838
4		0.002477	0.03400	0.1320	0.2811	0.3912	0.3576	0.1578	−0.1054	−0.2655	−0.2196
5			0.007040	0.04303	0.1321	0.2611	0.3621	0.3479	0.1858	−0.05504	−0.2341
6			0.001202	0.01139	0.04909	0.1310	0.2458	0.3392	0.3376	0.2043	−0.01446
7				0.002547	0.01518	0.05338	0.1296	0.2336	0.3206	0.3275	0.2167
8					0.004029	0.01841	0.05653	0.1280	0.2235	0.3051	0.3179
9						0.005520	0.02117	0.05892	0.1263	0.2149	0.2919
10						0.001468	0.006964	0.02354	0.06077	0.1247	0.2075
11							0.002048	0.008335	0.02560	0.06222	0.1231
12								0.002656	0.009624	0.02739	0.06337
13									0.003275	0.01083	0.02897
14									0.001019	0.003895	0.01196
15										0.001286	0.004508
16											0.001567

Figure 5.5: Four-place values of Bessel functions $J_n(\beta)$

Example 5.33

Let the carrier be given by $c(t) = 10\cos(2\pi f_c t)$ and let the message signal be $\cos(20\pi t)$. Further assume that the message is used to frequency modulate the carrier with $k_f = 50$. Find the expression for the modulated signal and determine how many harmonics should be selected to contain 99 % of the modulated signal power.

Solution
The power content of the carrier signal is given by

$$P_c = \frac{A_c^2}{2} = \frac{100}{2} = 50$$

The modulated signal is represented by

$$\begin{aligned} s(t) &= 10\cos\left(2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} \cos(20\pi\tau)d\tau\right) \\ &= 10\cos\left(2\pi f_c t + \frac{50}{10}\sin(20\pi t)\right) \end{aligned}$$

$$= 10\cos\left(2\pi f_c t + 5\sin(20\pi t)\right)$$

The modulation index is given by

$$\beta = k_f \frac{\max\left[|m(t)|\right]}{f_m} = 5$$

Therefore, the FM modulated signal is

$$\begin{aligned} s(t) &= \sum_{n=-\infty}^{\infty} A_c J_n(\beta)\cos\left(2\pi(f_c + nf_m)t\right) \\ &= \sum_{n=-\infty}^{\infty} 10 J_n(5)\cos\left(2\pi(f_c + 10 f_m)t\right) \end{aligned}$$

To make sure that at least 99 % of the total power is within the effective bandwidth, we have to choose k large enough such that

$$\sum_{n=-k}^{k} \frac{100 J_n^2(5)}{2} \geq 0.99 \times 50$$

The solution to this nonlinear equation is by trial and error using the table of Bessel functions. Using using the symmetry proper of Bessel function in given in *Eqn.* 5.7.10, we get

$$50\left[J_0^2(5) + 2\sum_{n=1}^{k} J_n^2(5)\right] \geq 49.5$$

Starting with small values of k and increasing it, we see that the smallest value of k for which the left-hand side exceeds the right-hand side is $k = 6$. This means that, if the modulated signal is passed through an ideal bandpass filter centered at fc with a bandwidth of at least 120 Hz, only 1 % of the signal power will be eliminated. This gives us a practical way to define the effective bandwidth of the angle-modulated signal as being 120 Hz. *Figure* 5.6 shows the frequencies present in the effective bandwidth of the modulated signal.

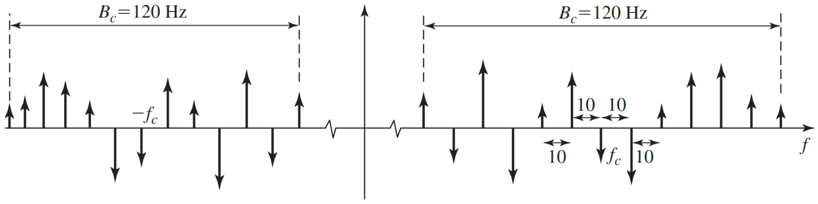


Figure 5.6: Harmonics present inside the effective bandwidth

Example 5.34

An FM signal has a deviation of 3 kHz and a modulating frequency of 1 kHz. Its total power is 5 W, developed across a 50 Ω resistive load. The carrier frequency is 160 MHz.

(a) Calculate the RMS signal voltage.

(b) Calculate the RMS voltage at the carrier frequency and each of the first three sets of sidebands.

(c) Calculate the frequency of each sideband for the first three sideband pairs.

(d) Calculate the power at the carrier frequency, and in each sideband, for the first three pairs. Determine what percentage of the total signal power is unaccounted for by the components described above.

(e) Sketch the signal in the frequency domain, as it would appear on a spectrum analyzer. The vertical scale should be power in dBm, and the horizontal scale should be frequency.

(f) Find the bandwidth of the FM signal.

Solution

(a) The signal power does not change with modulation, and neither does the voltage, which can easily be found from the power equation.

$$\begin{aligned} P_T &= \frac{V_T^2}{R_L} \\ V_T &= \sqrt{P_T R_L} \\ &= \sqrt{5 \times 50} = 15.8 \text{ V (rms)} \end{aligned}$$

(b) The modulation index must be found in order to use Bessel functions to find the carrier and sideband voltages

$$\beta_f = \frac{\Delta f}{f_m} = \frac{3}{1} = 3$$

From the Bessel function table, the coefficients for the carrier and the first three sideband pairs are:

$$J_o = -0.26 \qquad J_1 = 0.34 \qquad J_2 = 0.49 \qquad J3 = 0.31$$

These are normalized voltages, so they will have to be multiplied by the total RMS signal voltage to get the RMS sideband and carrier-frequency voltages. For the carrier,

$$V_c = J_o V_T$$

J_o has a negative sign. This simply indicates a phase relationship between the components of the signal. It would be required if we wanted to add together all the components to get the resultant signal. For our present purpose, however, it can be ignored, and we can use

$$V_o = |J_o|\, V_T = 0.26 \times 15.8 = 4.11 \text{ V}$$

Similarly

$$V_1 = 5.37 \text{ V}, \qquad V_2 = 7.74 \text{ V}, \qquad V_3 = 4.9 \text{ V}$$

(c) The sidebands are separated from the carrier frequency by multiples of the modulating frequency. Here, $f_c = 160$ MHz and $f_m = 1$ kHz, so there are sidebands at each of the following frequencies.

$$f_{USB1} = 160.001 \text{ MHz}, \qquad f_{USB2} = 160.002 \text{ MHz}, \qquad f_{USB3} = 160.003 \text{ MHz}$$

$$f_{LSB1} = 159.999 \text{ MHz}, \qquad f_{LSB2} = 159.998 \text{ MHz}, \qquad f_{LSB3} = 159.997 \text{ MHz}$$

(d) Since each of the components of the signal is a sinusoid, the usual equation can be used to calculate power. All the components appear across the same 50 Ω load.

$$P_c = \frac{V_c^2}{R_L} = \frac{4.11^2}{50} = 0.338 \text{ W}$$

Similarly,

$$P_1 = 0.576 \text{ W}, \qquad P_2 = 1.2 \text{ W}, \qquad P_3 = 0.48 \text{ W}$$

(e) To find the total power in the carrier and the first three sets of sidebands, it is only necessary to add the powers calculated above, counting each of the sideband powers twice, because each of the calculated powers represents one of a pair of sidebands.

$$\begin{aligned} P_T &= P_c + 2\,(P_1 + P_2 + P_3) \\ &= 0.338 + 2\,(0.576 + 1.2 + 0.48) = 4.85 \text{ W} \end{aligned}$$

This is not quite the total signal power, which was given as 5 W. The remainder is in the additional sidebands. To find how much is unaccounted for by the carrier and the first three sets of sidebands, we can subtract. Let the difference P_x.

$$P_x = 5 - 4.85 = 0.15 \text{ W}$$

As a percentage of the total power this is

$$P_x(\%) = \frac{0.15}{5} \times 100 = 3 \text{ \%}$$

All the information we need for the sketch is on hand, except that the power values have to be converted to dBm using the equation

$$P(\text{dBm}) = 10 \log \frac{P}{1 \text{ mW}}$$

This gives

$$P_c(\text{dBm}) = 25.3 \text{ dBm}, \quad P_1(\text{dBm}) = 27.6 \text{ dBm}, \quad P_2(\text{dBm}) = 30.8 \text{ dBm},$$
$$P_3(\text{dBm}) = 26.8 \text{ dBm}$$

The sketch is shown in *Figure* 5.7

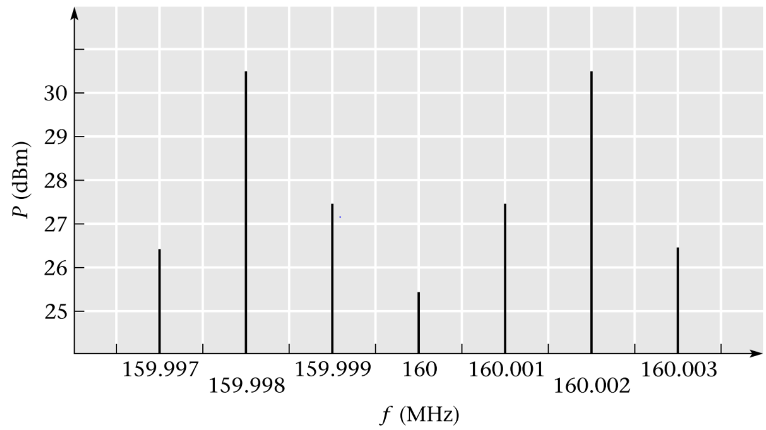


Figure 5.7: Sketch for Question 5.34(e)

(f) Using Carson's rule the bandwidth of the FM signal is calculated as

$$\begin{aligned} B_T &\approx 2\,(\Delta f + f_m) \\ &= 2(2+1) = 8 \text{ kHz} \end{aligned}$$

We found that 97 % of the power was contained in a bandwidth of 6 kHz. An 8-kHz bandwidth would contain more of the signal power.

5.8 Narrowband Angle Modulation (NBFM)

An angle-modulated carrier can be represented in exponential form as

$$\begin{aligned} s(t) &= \mathrm{Re}\left(A_c e^{j2\pi f_c t} e^{j\phi(t)}\right) \\ &= A_c \cos\left[2\pi f_c t + \beta \sin(2\pi f_m t)\right] \end{aligned} \tag{5.8.1}$$

where $\phi(t) = \beta \sin(2\pi f_m t)$

Now, if in an angle modulated system, the deviation constants k_p and k_f and the message signal $m(t)$ are such that for all t, $\phi(t) \ll 1$, then, $s(t)$ can be approximated simply by expanding $s(t)$ using $\cos(A+B) = \cos A \cos B - \sin A \sin B$. This gives

$$\begin{aligned} s(t) &= A_c \cos 2\pi f_c t \cos \phi(t) - A_c \sin 2\pi f_c t \sin \phi(t) \\ &\approx A_c \cos 2\pi f_c t - A_c \phi(t) \sin 2\pi f_c t \end{aligned} \tag{5.8.2}$$

Eqn. 5.8.2 shows that in this case the modulated signal is very similar to a conventional AM signal. The only difference is that, the message signal $m(t)$ is modulated on a sine carrier rather than a cosine carrier. The bandwidth of this signal is similar to the bandwidth of a conventional AM signal, which is twice the bandwidth of the message signal.
The output of modulator contains a carrier component and a term that is function of $m(t)$ multiplies by a 90^o phase-shifted carrier. This multiplication generates a pair of sidebands. Thus, if $\phi(t)$ has a bandwidth W, the bandwidth of a narrowband angle modulator output is $2W$. It is important to note, however, that the carrier and the resultant of the sidebands for narrowband angle modulation with sinusoidal modulation are in phase quadrature, whereas for AM they are not. *Figure* 5.8 shows generation of narrowband FM wave

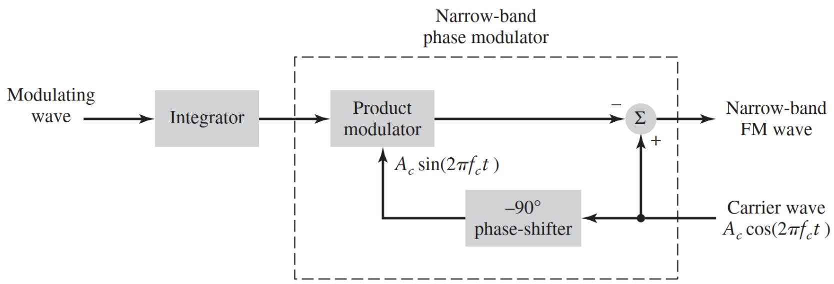


Figure 5.8: Generation (indirect method) of narrow-band FM wave

5.8.1 Spectrum of Narrowband Angle Modulation

Substituting ϕ and expanding *Eqn.* 5.8.2 yields

$$s(t) \approx A_c \left[\cos(2\pi f_c t) \cos(\beta \sin 2\pi f_m t) - \sin(2\pi f_c t) \sin(\beta \sin 2\pi f_m t)\right] \tag{5.8.3}$$

NOTE:	$\cos(A+B) = \cos A \cos B - \sin A \sin B$ where $A = 2\pi f_c t$ and $B = \beta \sin 2\pi f_m t$

In NBFM, β is small, hence it is possible to approximate these

$$\begin{aligned} \cos(\beta \sin 2\pi f_m t) &\approx 1 \\ \sin(\beta \sin 2\pi f_c t) &\approx \beta \sin 2\pi f_m t \end{aligned}$$

Eqn. 5.8.3 then becomes

$$s(t) \approx A_c \cos(2\pi f_c t) - A_c \beta \sin(2\pi f_c t) \sin 2\pi f_m t \tag{5.8.4}$$

***NOTE*:**	$\sin A \sin B = \frac{1}{2}\cos(A-B) - \frac{1}{2}\cos(A+B)$

Eqn. 5.8.3 finally becomes

$$s(t) \approx A_c \cos(2\pi f_c t) + \frac{1}{2}\beta A_c \{\cos[2\pi(f_c + f_m)t] - \cos[2\pi(f_c - f_m)t]\} \tag{5.8.5}$$

Recall that the AM wave is given by

$$s_{AM}(t) = A_c \cos(2\pi f_c t) + \frac{1}{2}\mu A_c \{\cos[2\pi(f_c + f_m)t] + \cos[2\pi(f_c - f_m)t]\} \tag{5.8.6}$$

The similarity of NBFM and AM, allows it to possesses some AM properties. The only difference between *Eqn.* 5.8.6 and *Eqn.* 5.8.5 is the sign reversal of the lower sideband. Thus NBFM requires the same bandwidth as AM i.e. $2f_m$.

To get the spectrum of *Eqn.* 5.8.5, we need to take the Fourier transform first

$$\begin{aligned} S(f) = &\frac{A_c}{2}[\delta(f - f_c) + \delta(f + f_c)] - \frac{\beta A_c}{4}[\delta(f - f_c + f_m) + \delta(f + f_c - f_m)] \\ &+ \frac{\beta A_c}{4}[\delta(f - f_c - f_m) + \delta(f + f_c + f_m)] \end{aligned} \tag{5.8.7}$$

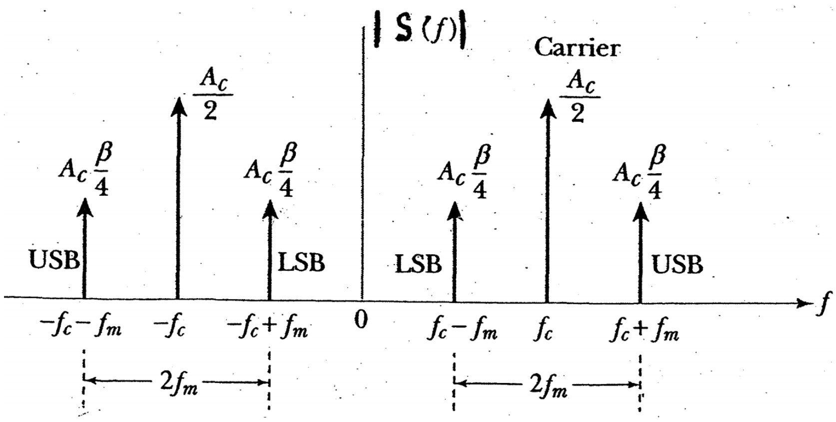


Figure 5.9: Spectral content of a NBFM wave for single-tone modulation!

Example 5.35

An angle modulated signal with carrier frequency $\omega_c = 2\pi \times 10^5$ is described by the equation

$$u(t) = 10\cos(\omega_c t + 5\sin 3000t + 10\sin 2000\pi t)$$

(a) Find the power of the modulated signal

(b) Find the maximum frequency deviation Δf_{max}

(c) Find the deviation ratio β

(d) Find the phase deviation $\Delta\phi$

(e) Estimate the bandwidth of $u(t)$

Solution
The signal bandwidth is the highest frequency in $m(t)$. In this case $W = 2000\pi/2\pi = 1000$ Hz.

1. The carrier amplitude is 10, the power is

$$P = \frac{10^2}{2} = 50$$

2. To find frequency deviation Δf, we find the instantaneous frequency ω_i given by

$$\omega_i = \frac{d}{dt}\theta(t) = \omega_c + 15000\cos 3000t + 20000\cos 2000\pi t$$

The carrier deviation is $15000\cos 3000t + 20000\pi\cos 2000\pi t$. At some point, the two sinusoids will add up, hence the maximum deviation, $\Delta\omega_{max}$ is expressed as $15000 + 20000\pi$ (i.e. taking the amplitude of the sinusoids). Hence,

$$\Delta f_{max} = \frac{\Delta\omega_{max}}{2\pi} = 12387.32 \text{ Hz}$$

3. Deviation ratio

$$\beta = \frac{\Delta f_{max}}{W} = \frac{12387.32}{1000} = 12.387$$

4. The angle $\theta(t) = \omega_c t + (5\sin 3000t + 10\sin 2000\pi t)$.
The phase deviation is maximum value of the angle inside the parentheses, and is given by $\Delta\phi = 15$ rad.

5. Bandwidth of $u(t)$

$$B_T = 2(\Delta f + W) = 26774.65 \text{ Hz}$$

Example 5.36

A 10 MHz carrier signal is frequency modulated by a sinusoidal signal of unity amplitude and with a FM factor $k_f = 10$ Hz/V. Find the approximate bandwidth of the frequency modulated signal if the modulating frequency (single tone) is 10 kHz.

Solution
Start with the expression:

$$s(t) = A_c\cos\left[(2\pi f_c t) + \frac{k_f A_m}{f_m}\cos(2\pi f_m t)\right]$$

Modulation index

$$\beta = \frac{k_f A_m}{f_m} = \frac{(10\,\text{Hz/V})\,(1\,\text{V})}{10^4\,\text{Hz}} = 10^{-3}$$

So this is clearly a narrowband FM (NBFM) case. Since it is NBFM we can use the equation that the bandwidth is approximately twice the modulating time frequency f_m.
NBFM bandwidth $B_T = 2f_m = 20$ kHz

5.9 Generation of Angle Modulated Signal

5.9.1 Indirect Method - Armstrong Modulator

Generation of Narrowband FM

This modulator involves splitting the carrier wave $A_c\cos(2\pi f_m t)$ into two paths. One path is direct; the other path contains a degree phase-shifting network and a product modulator, the combination of which generates a DSB-SC modulated wave. The difference between these two signals produces a narrow-band FM wave, but with some amplitude distortion.
Figure 5.10 shows indirect method of generating NBFM. In indirect method, the message

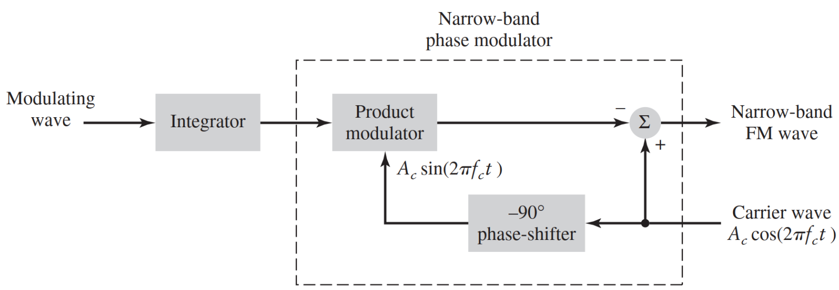


Figure 5.10: Block diagram of an indirect method for generating a narrowband FM wave

signal $m(t)$ is first passed through an integrator before before applying it to the phase modulation as shown *Figure* 5.10. The carrier signal is generated by using crystal oscillator because it provides very high frequency stability.

The operation of indirect method is divided into two parts as follows

I. Generate a NBFM wave using a phase modulator

II. Using the frequency multiplier of the mixer to obtain the required values of frequency deviation and modulation index – WBFM.

In the order to minimize the distortion in the phase modulator, the maximum phase deviation or modulation index β is kept small thereby resulting in a NBFM signal.
Let

$$s'(t) = A_c\cos\left[2\pi f_c t + k_f \int_{0|-\infty}^{t} m(\tau)d\tau\right] \tag{5.9.1}$$

where f_c is the frequency of the crystal oscillator of k_f is the frequency sensitivity constant in Hz/V. For a single-tone modulation signal defined by $m(t) = A_m\cos 2\pi f_m t$, then *Eqn.* 5.9.1 becomes

$$s'(t) = A_c\cos\left[2\pi f_c t + \beta\sin 2\pi f_m t\right] \tag{5.9.2}$$

The instantaneous frequency of *Eqn.* 5.9.2 is $f_i(t) = f_c + k_f m(t)$.

Generation of Wideband FM

A WBFM signal can be obtained from a NBFM signal using a device called a frequency multiplier producing the defined WBFM wave as shown *Figure* 5.11

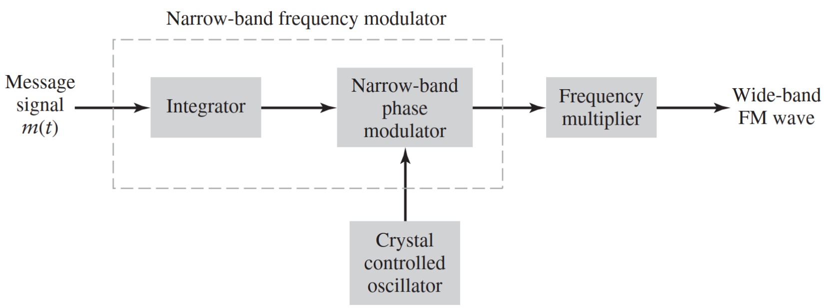


Figure 5.11: Block diagram of the indirect method of generating a wide-band FM wave

A frequency multiplier consist of a memoryless nonlinear device followed by a BPF. It is a nonlinear device designed to multiply the frequencies of an input signal by a given factor. The input-to-output relation of such a nonlinear device is expressed in the general form as

$$v(t) = a_1 s'(t) + a_2 s'^2(t) + \ldots + a_n s'^n(t)$$

where $a_1,\ a_2, \ldots,\ a_n$ are co-efficients and n is the highest order of the nonlinearity. Substituting *Eqn.* 5.9.2 and simplifying, we can find the FM wave having carrier frequencies $f_c,\ 2f_c, \ldots,\ nf_c$ with deviation as $\Delta f_c,\ 2\Delta f_c, \ldots,\ n\Delta f_c$.

The BPF has two functions to perform:

I. To pass the FM wave centered at carrier frequency nf_1 and having deviation of $n\Delta f_1$

II. To suppress all other FM spectra and the DC term.

The output of the frequency multiplier produces the desired WBFM wave having the following time-domain expression.

$$s(t) = A_c \cos\left(2\pi f_c t + nk_f \int_{0|-\infty}^{t} m(\tau) d\tau\right) \tag{5.9.3}$$

where the instantaneous frequency is given by $f_i(t) = nf_c + nk_f m(t)$

Illustration of Indirect method of WBFM signal generation

Consider a transmitter with the following specifications (old transmitters); carrier freq. $f_c = 200$ kHz, freq. deviation, $\Delta f = 25$ Hz and message freq. $f_m = 15$ kHz. The deviation ratio of NBFM before frequency multiplication is 0.00167 (typical deviation ratio of NBFM). The final peak frequency deviation for commercial FM is 75 kHz, the frequency multiplication factor required is $(75)(103)/25 = 3000$. After frequency multiplication, the new carrier frequency becomes $(200)(103)(3000) = 600$ MHz, which is outside the FM band of 88 to 108 MHz. A frequency down converter with a local oscillator frequency in the range of 600 ± 100 MHz is used to translate the spectrum from 600 MHz down to the FM band.

Another approach for generating an angle-modulated signal is to first generate a narrowband angle-modulated signal, and then change it to a wideband signal. Due to the similarity of conventional AM signals, generation of narrowband angle-modulated signals is straight forward.

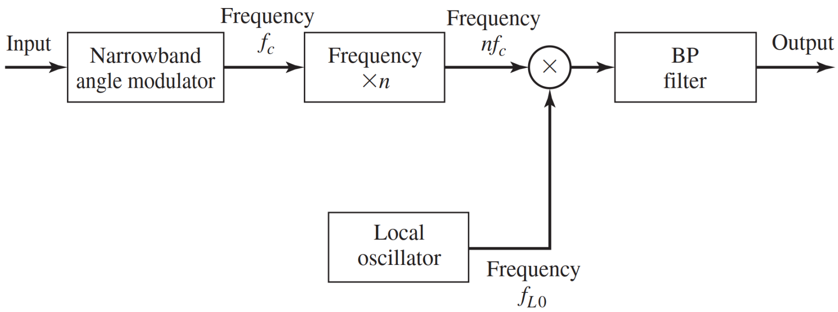


Figure 5.12: Indirect generation of angle-modulated signals

Figure 5.12 shows the block diagram of a system that generates wideband angle-modulated signals from narrowband angle-modulated signals. The first stage of such a system is, of course, a narrowband angle-modulator. The narrowband angle-modulated signal enters a frequency multiplier that multiplies the instantaneous frequency of the input by some constant n. This is usually done by applying the input signal to a nonlinear element and then passing its output through a bandpass filter tuned to the desired central frequency. If the narrowband- modulated signal is represented by

$$s_{nb}(t) = A_c \cos(2\pi f_c t + \phi) \tag{5.9.4}$$

The output of the frequency multiplier (output of the bandpass filter) is given by

$$y(t) = A_c \cos(2\pi n f_c t + n\phi(t)) \tag{5.9.5}$$

In general, this is, of course, a wideband angle-modulated signal. However, there is no guarantee that the carrier frequency of this signal, nf_c, will be the desired carrier frequency. The last stage of the modulator performs an up or down conversion to shift the modulated signal to the desired center frequency. This stage consists of a mixer and a bandpass filter. If the frequency of the local oscillator of the mixer is f_{LO} and we are using a down converter, the final wideband angle-modulated signal is given by

$$s_{wb}(t) = A_c \cos(2\pi(nf_c - f_{LO})t + n\phi(t)) \tag{5.9.6}$$

Since we can freely choose n and f_{LO}, we can generate any modulation index at any desired carrier frequency by this method.

Example 5.37

An FM signal

$$s(t) = 5\cos\left(2\pi 10^6 t + \sin 20000\pi t\right)$$

is input to a square-law nonlinearity (with the characteristic: $y = 2x^2$, where x is the input and y is the output), filtered by a BPF. The centre frequency of the BPF is 2.03 MHz and the bandwidth is 10 kHz. Determine the output $z(t)$ and sketch its magnitude spectrum.

Solution

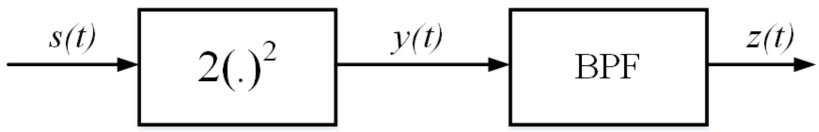


Figure 5.13: Illustration

$$y(t) = 2s^2(t) = 2A^2\cos^2(\omega_c t + \phi(t))$$
$$= A^2 + A^2\cos(2\omega_c t + 2\phi(t))$$

Rejecting DC component by BPF, we have a form of $y(t)$ as

$$y_1(t) = A^2\cos(2\omega_c t + 2\phi(t))$$

From the question,

$$y_1(t) = 25\cos\left(2\pi \times 2 \times 10^6 t + 2\sin 20000\pi(t)\right)$$

This is an FM signal with $f_c = 2$ MHz, $\beta = 2$ and $f_m = 10$ kHz.

$$y_1(t) = 25\sum_{n=-\infty}^{\infty} J_n(2)\cos\left[2\pi\left(2\times 10^6 + n\times 10\times 10^3\right)t\right]$$

The only tone that will pass through the BPF with center freq. 2.03 MHz and BW $= 10$ kHz will be the tone at 2.03 MHz, where $n = 3$.
The output of BPF indicated on the *Figure* 5.13.

$$z(t) = 25\cdot J_3(2)\cos\left[2\pi\left(2\times 10^6 + 30\times 10^3\right)t\right]$$

Note that $J_3(2) = 0.129$. Finally,

$$z(t) = 3.225\cos\left[2\pi\left(2\times 10^6 + 30\times 10^3\right)t\right] = 3.225\cos\left[2\pi(2.03\times 10^6)t\right]$$

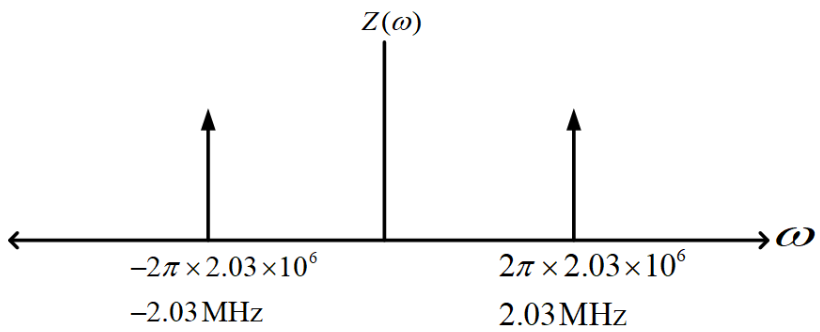


Figure 5.14: Magnitude spectrum

5.9.2 Direct Method

5.9.2.1 Varactor Diode Modulator

In a voltage-controlled oscillator (VCO), the frequency is controlled by an external voltage. The oscillation frequency varies linearly with the control voltage. We can generate an FM wave by using the modulating signal $m(t)$ as a control signal. This gives the instantaneous frequency as

$$\omega_i(t) = \omega_c + k_f m(t)$$

One can construct a VCO using an operational amplifier and an hysteric comparator (such as a Schmitt trigger circuit). Varying one of the reactive parameters (C or L) of the resonant circuit of an oscillator can be to achieved the same results. A reverse-biased semiconductor diode acts as a capacitor whose capacitance varies with the bias voltage.

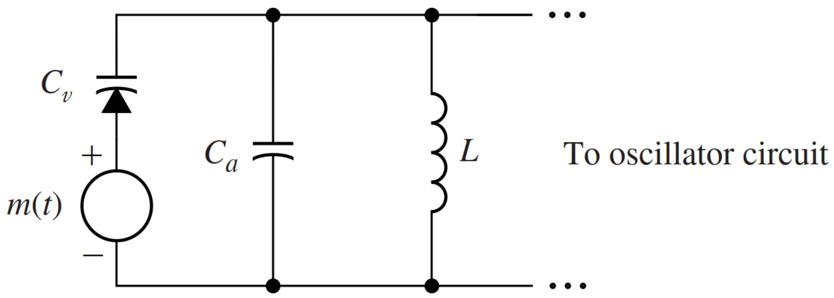


Figure 5.15: Varactor diode implementation of an anle modulator

The capacitance of these diodes – *varactor* (*Figure* 5.15), can be approximated as a linear function of the bias voltage $m(t)$ over a limited range. In Hartley or Colpitt oscillators, for instance, the frequency of oscillation is given by

$$\omega_i = \frac{1}{\sqrt{LC}} \tag{5.9.7}$$

If the capacitance C is varied by the modulating signal $m(t)$, that is if

$$C = C_o - km(t) \tag{5.9.8}$$

Following

$$\begin{aligned} \omega_i &= \frac{1}{\sqrt{LC_o\left[1 - \frac{km(t)}{C_o}\right]}} \\ &= \frac{1}{\sqrt{LC_o}\left[1 - \frac{km(t)}{C_o}\right]^{1/2}} \\ &\approx \frac{1}{\sqrt{LC_o}}\left[1 + \frac{km(t)}{2C_o}\right], \qquad \frac{km(t)}{C_o} \ll 1 \end{aligned} \tag{5.9.9}$$

Applying binomial approximation $(1+x)^n \approx 1 + nx$ for $|x| \ll 1$. Thus,

$$\begin{aligned} \omega_i &= \omega_c\left[1 + \frac{km(t)}{2C_o}\right], \qquad \omega_c = \frac{1}{\sqrt{LC_o}} \\ &= \omega_c + k_f m(t), \qquad k_f = \frac{k\omega_c}{2C_o} \end{aligned} \tag{5.9.10}$$

Because $C = C_o - km(t)$, the maximum capacitance deviation is

$$\Delta C = kA_m = \frac{2k_f C_o A_m}{\omega_c} \tag{5.9.11}$$

Hence,

$$\frac{\Delta C}{C_o} = \frac{2k_f A_m}{\omega_c} = \frac{2\Delta f}{f_c} \tag{5.9.12}$$

In practice, $\Delta f / f_c$ is usually small, and, hence, ΔC is a small fraction of C_o, which helps limit the harmonic distortion that arises because of the approximation used in this derivation.
Direct FM generation generally produces sufficient frequency deviation and requires little frequency multiplication. But this method has poor frequency stability. In practice, feedback is used to stabilize the frequency.

5.9.2.2 Basic Reactance Modulator

The operation of this modulator is based on certain simple operating conditions. That is, the impedance z seen at the input terminals A-A of *Figure* 5.16 is reactive – made inductive or capacitive by a single component change. The circuit shown is the basic circuit of a **FET reactance modulator**, which behaves as a three-terminal reactance that may be connected across the tank circuit of the oscillator to be frequency-modulated. It can be made inductive or capacitive by a simple component change. The value of this reactance is proportional to the transconductance of the device, which can be made to depend on the gate bias and its variations.

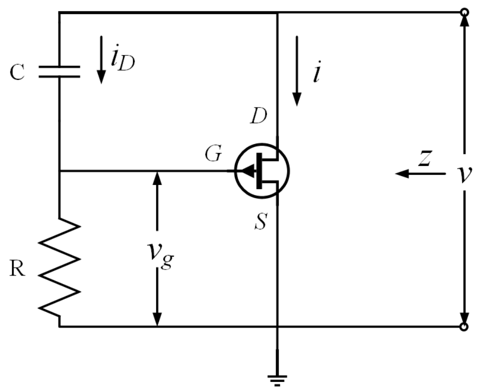


Figure 5.16: Basic reactance modulator

Theory
In order to determine z, a voltage v is applied to the terminals A-A between which the impedance is to be measured, and the resulting current i is calculated. The applied voltage is then divided by this current, giving the impedance seen when looking into the terminals. In order for this impedance to be a pure capacitive reactance (for example), two requirements must be fulfilled:

i. Bias network current i_b must be negligible compared to the drain current, i.e. $(i_b \ll i)$.

ii. Drain-to-gate impedance, X_c must be greater than the gate-to-source impedance, R, preferably by more that 5 : 1.

Now,

$$v_g = r_b R = \frac{v}{R - jX_c} R \tag{5.9.13}$$

The FET drain current is

$$i = g_m v_g = \frac{g_m v}{R - jX_c} R \tag{5.9.14}$$

Therefore, the impedance seen at the terminals A-A is

$$z = \frac{v}{i} = v + \frac{g_m R v}{R - jX_c} = \frac{R - jX_c}{g_m R} = \frac{1}{g_m}\left(1 - \frac{jX_c}{R}\right) \tag{5.9.15}$$

If $X_c \gg R$ in *Eqn.* 5.9.15 will reduce to

$$z = -j\frac{X_c}{g_m R} \quad \Rightarrow \text{(capacitive reactance)} \tag{5.9.16}$$

Eqn. 5.9.16 can be written as

$$X_{eq} = \frac{X_c}{g_m R} = \frac{1}{2\pi f g_m RC} = \frac{1}{2\pi f C_{eq}} \tag{5.9.17}$$

From *Eqn.* 5.9.17, it is seen that under such conditions, the input impedance of the device at A-A is a pure reactance given by

$$X_{eq} = g_m RC \tag{5.9.18}$$

The following can be observed from *Eqn.* 5.9.18:

1. The equivalent capacitance C_{eq} depends on the device transconductance and therefore, can be varied by the bias voltage.

2. This capacitance can be adjusted to any value by varying the components R and C.

3. The expression $g_m RC$ has the correct dimensions of capacitance; R, measured in ohms, and g_m, measured in Siemens (S), cancel each other's dimensions leaving C are required.

4. Gate-to-drain impedance X_c must be larger than gate-to-source impedance R as stated earlier.

5. z would have been a resistive component if X_c/R had not been much greater than unity.

6. The resistive component for this particular FET reactance modulator will be $1/1$ This component contains g_m, it will vary with the applied modulating voltage. This variable resistance (like the variable reactance) will appear directly across the tank circuit of the master oscillator, varying its Q and therefore its output voltage. A certain amount of amplitude modulation will be created. This applies to all the forms of reactance modulator. If the situation is unavoidable, the oscillator being modulated must be followed by an amplitude limiter.

The gate-to-drain impedance is, in practice, made five to ten times the gate-to-source impedance. Let $X_c = nR$ (at the carrier frequency) in the capacitive RC reactance FET so far discussed. Then

$$\begin{aligned} X_c &= \tfrac{1}{\omega C} = nR \\ C &= \tfrac{1}{\omega nR} = \tfrac{1}{2\pi f nR} \end{aligned} \tag{5.9.19}$$

Putting *Eqn.* 5.9.19 into *Eqn.* 5.9.18 gives

$$C_{eq} = g_m RC = \frac{g_m R}{2\pi f n R} = \frac{g_m}{2\pi f n} \tag{5.9.20}$$

In practical situations the frequency of operation and the ratio of X_c to R are the usual starting data from which other calculations are made. It is left to the student to find other types of Reactance Modulator.

Example 5.38

Determine the value of the capacitive reactance obtainable from a reactance FET whose g_m, is 12 milliSiemens (12 mS). Assume that the gate-tu-source resistance is one-ninth of the reactance of the gate-to-drain capacitor and that the frequency is 5 MHz.

Solution

$$C_{eq} = \frac{g_m}{2\pi f n}$$

$$\therefore \quad 2\pi f C_{eq} = \frac{g_m}{n} = \frac{1}{X_{C_{eq}}}$$

$$X_{C_{eq}} = \frac{n}{g_m} = \frac{9}{12 \times 10^{-3}} = 750 \ \Omega$$

Example 5.39

It is required to provide a maximum deviation of 75 kHz for the 88 MHz carrier frequency of a VHP FM transmitter. A FET is used as a capacitive reactance modulator, and the linear portion of its g_m-v_{gs}, curve lies from 320 μS (at which $v_{gs} = -2$ V) to 830 μS (at which $v_{gs} = -0.5$ V). Assuming that R_{gs} is one-tenth of $X_{C_{gd}}$, calculate:

(a) The rms value oft he required modulating voltage

(b) The value oft he fixed capacitance and inductance of the oscillator tuned circuit across which the reactance modulator is connected

Solution

(a) V_m peak-to-peak $= 2 - 0.5 = 1.5$ V

$$V_{m_{rms}} = 1.5 \Big/ 2\sqrt{2} = 0.53 \text{ V}$$

(b)

$$C_n = \frac{g_{m_{\min}}}{2\pi f n} = \frac{3.2\times 10^{-4}}{2\pi\times 8.8\times 10^{7}\times 10} = \frac{3.2\times 10^{-12}}{2\pi\times 8.8} = 5.8\times 10^{-14}$$
$$= 0.058 \text{ pF}$$
$$C_x = \frac{C_n g_{m_{\max}}}{g_{m_{\min}}} = 0.058\times\frac{820}{320} = 0.15 \text{ pF}$$

Now,

$$\frac{f_x}{f_n} = \frac{1}{2\pi\sqrt{L(C+C_n)}} + \frac{1}{2\pi\sqrt{L(C+C_x)}}$$
$$= \sqrt{\frac{C+C_x}{C+C_n}}$$
$$\left(\frac{f_x}{f_n}\right)^2 = \frac{C+C_x}{C+C_n}$$
$$\frac{f_x^2}{f_n^2} - 1 = \frac{C+C_x}{C+C_n} - 1$$
$$\frac{f_x^2 - f_n^2}{f_n^2} = \frac{C+C_x-C-C_n}{C+C_n}$$
$$\frac{(f_x+f_n)(f_x-f_n)}{f_n^2} = \frac{4f\delta}{f_n^2} \approx \frac{C_x-C_n}{C+C_n}$$
$$C+C_n = \frac{(C_x-C_n)f^2}{4f\delta}$$
$$C = \frac{(C_x-C_n)f^2}{4f\delta} - C_n$$
$$= \frac{(0.150-0.058)\times 88}{4\times 0.075} - 0.058$$
$$\approx \frac{0.092\times 88}{0.3} = 27 \text{ pF}$$
$$\therefore \qquad f = \frac{1}{2\pi\sqrt{L(C+C_{av})}} \approx \frac{1}{2\pi\sqrt{LC}}$$

So,

$$L = \frac{1}{4\pi^2 f^2 C} = \frac{1}{4\pi^2\times 8.8^2\times 10^{14}\times 2.7\times 10^{-11}}$$
$$= \frac{10^{-3}}{39.5\times 77.4\times 2.7} = \frac{10^{-5}}{82.5} = 1.21\times 10^{-7}$$
$$= 0.121\,\mu\text{H}$$

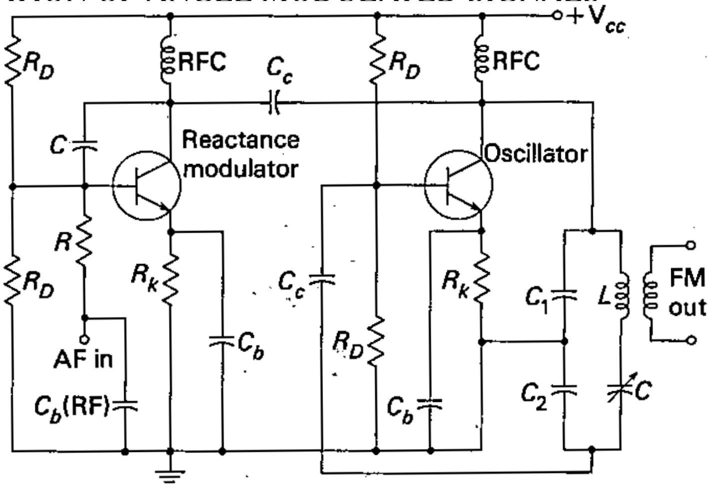


Figure 5.17: Typical example of Reactance Modulator circuit – Transistor RM

5.10 Demodulation of Angle Modulated Signals

An FM detector must satisfies the following requirement in order to demodulate FM wave properly.

1. It must convert frequency variation into amplitude variation
2. The conversion must be linear and efficient
3. The demodulator circuit should be insensitive to amplitude changes. It should respond only to the frequency changes.
4. It should not be too critical in its adjustment and operation

A variety of techniques and circuits have been developed for demodulating FM signals. We shall consider a few of these techniques falling under the following categories:

I) FM-to-AM conversion

II) Phase shift discrimination
This method of FM demodulation involves converting frequency variations into phase variations and detecting the phase changes. In other words, this method makes use of linear phase networks instead of the linear amplitude characteristic of the circuits used in the previous method. Under this category, we have the Foster-Seely discriminator (and its variant the ratio detector) and the quadrature detector.

III) Zero crossing detection

IV) Phase Locked Loop (PLL)
PLL is a versatile building block of the present day communication systems. Besides FM demodulation, it has a large number of other applications such as carrier tracking, timing recovery, frequency synthesis etc. The basic aim of a PLL is to lock (or synchronize) the instantaneous angle of a VCO output to the instantaneous angle of a signal that is given as input to the PLL. In the case of demodulation of FM, the input signal to PLL is the received FM signal.

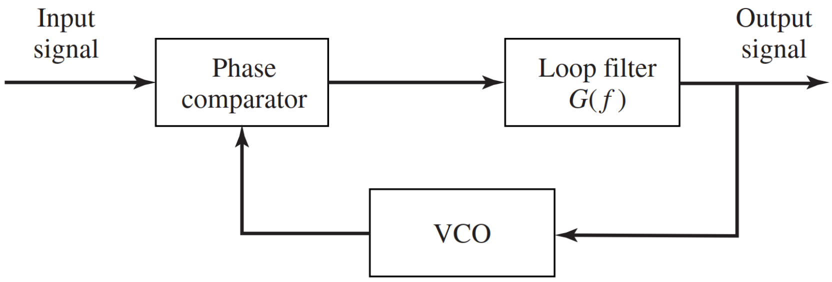


Figure 5.18: Block diagram of PLL-FM demodulator

The phase-locked loop is a feedback system whose operation is closely linked to frequency modulation. It is commonly used for carrier synchronization, and indirect frequency demodulation. The latter application is the subject of interest here. Basically, the phase-locked loop consists of three major components:

(a) **Voltage-controlled oscillator** (VCO), which performs frequency modulation on its own control signal.

(b) **Multiplier**, which multiplies an incoming FM wave by the output of the voltage-con- trolled oscillator.

(c) **Loop filter** of a low-pass kind, the function of which is to remove the high-frequency components contained in the multiplier's output signal and thereby shape the over- all frequency response of the system.

5.10.1 Comparison of FM and AM

The frequency modulation (FM) has the following advantages over the amplitude modulation (AM) :

1. FM receivers may be fitted with amplitude limiters to remove the amplitude variations caused by noise. This makes FM reception a good deal more immune to noise than AM reception.

2. It is possible to reduce noise still further by increasing the frequency-deviation. This is a feature which AM does not have because it is not possible to exceed 100 percent modulation without causing severe distortion.

3. Standard frequency allocations provide a guard band between commercial FM stations. Due to this, there is a less adjacent-channel interference than in AM.

4. FM broadcasts operate in the upper VHF and UHF frequency ranges at which there happens to be less noise than in the MF and HF ranges occupied by AM broadcasts.

5. The amplitude of FM wave is constant. It is therefore, independent of the modulation depth whereas in AM, modulation depth governs the transmitted power. This permits the use of low- level modulation in FM transmitter and use of efficient class C amplifiers in all stages following the modulator. Further since all amplifiers handle constant power, the average power handled equals the peak power. In AM transmitter, the maximum power is four times, the average power. Finally in FM, all the transmitted power is useful whereas in AM, most of the power is carrier power which does not contain any information.

However, the FM have some drawbacks over AM which are given below :

i. FM transmitting and receiving equipments particularly used for modulation and demodulation are more complex and more costly.

ii. A much wider channel typically 200 kHz is required in FM as against only 10 kHz in AM broadcast. This forms serious limitation of FM.

iii. In FM, the area of reception is small as it is limited to only one line of sight.

5.10.2 Comparison of FM and PM

FM and PM are quite similar except by the manner in which the modulation index is defined.

1. For FM, the phase angle is proportional to modulating wave whereas for PM, the time derivative of phase angle is proportional to modulating wave.

2. In FM, the frequency deviation is proportional to the amplitude of the modulating signal whereas in PM, the phase deviation is proportional to the amplitude of the modulating signal and therefore independent of its frequency.

3. In case of FM, the modulation index is inversely proportional to the modulating frequency whereas in PM, the modulation index is proportional to the modulating voltage.

It means that under identical conditions FM and PM are indistinguishable for a single modulating frequency. When the modulating frequency is changed, the PM modulation index will remain constant but the FM modulation index will decrease as modulating frequency is increased and vice-versa.

5.11 Review Questions

1. Sketch FM and PM waves for the modulating signal $m(t)$ shown in *Figure* 5.19. The constants k_f and k_p are $2\pi \times 10^5$ and 10π respectively, and the carrier f_c is 100 MHz

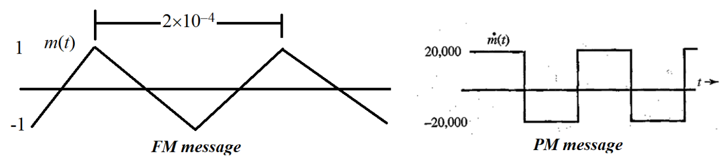


Figure 5.19: FM and PM waveform

2. (a) Estimate B_{FM} and B_{PM} for the modulating signal $m(t)$ in *Figure* 5.19 for $k_f = 2\pi \times 10^5$ and $k_p = 5\pi$. Assume that, the essential bandwidth of the periodic $m(t)$ has a frequency of its third harmonic.
(b) Repeat the problem if the amplitude of $m(t)$ is doubled [if $m(t)$ is multiplied by 2].

3. Discus the nature of distortion inherent in the Armstrong indirect FM generator.

4. State how the frequency of a carrier varies in an FM system when the modulating signal amplitude and frequency change.

5. State how the frequency of a carrier varies in a PM system when the modulating signal amplitude and frequency change.

6. When does maximum frequency deviation occur in an FM signal and PM signal?

7. State the conditions that must exist for a phase modulator to produce FM.

8. What do you call FM produced by PM techniques?

9. How must the nature of the modulating signal be modified to produce FM by PM techniques?

10. What is the difference between the modulation index and the deviation ratio?

11. Define narrowband FM. What criterion is used to indicate NBFM?

12. What is the name of the mathematical equation used to solve for the number and amplitude of sidebands in an FM signal?

13. List four major applications for FM.

14. The AM broadcast band consists of 107 channels for stations 10 kHz wide. The maximum permitted modulating frequency is 5 kHz. Could FM be used on this band? If so, explain what would be necessary to make it happen.

15. A carrier of 49 MHz is frequency-modulated by a 1.5 kHz square wave. The modulation index is 0.25. Sketch the spectrum of the resulting signal. (Assume that only harmonics less than the sixth are passed by the system)

16. The FM radio broadcast band is allocated the frequency spectrum from 88 to 108 MHz. There are 100 channels spaced 200 kHz apart. The first channel center frequency is 88.1 MHz; the last, or 100^{th}, channel center frequency is 107.9 MHz. Each 200 kHz channel has a 150 kHz modulation bandwidth with 25 kHz "guard bands" on either side of it to minimize the effects of overmodulation (over deviation). The FM broadcast band permits a maximum deviation of ± 75 kHz and a maximum modulating frequency of 15 kHz.

    i. Draw the frequency spectrum of the channel centered on 99.9 MHz, showing all relevant frequencies.
    ii. Draw the frequency spectrum of the FM band, showing details of the three lowest-frequency channels and the three highest-frequency channels.
    iii. Determine the bandwidth of the FM signal by using the deviation ratio and the Bessel table.
    iv. Determine the bandwidth of the FM signal by using Carson's rule.
    v. Which of the above bandwidth calculations best if ts the available channel bandwidth?

17. Assume that you could transmit digital data over the FM broadcast band radio station. The maximum allowed bandwidth is 200 kHz. The maximum deviation is 75 kHz, and the deviation ratio is 5. Assuming that you wished to preserve up to the third harmonic, what is the highest-frequency square wave you could transmit?

18. (a) An angle modulated signal is describe by $x(t) = 10\cos\left[2\pi(10^6)t + 0.1\sin(10^3)\pi t\right]$ considering $x(t)$ at a PM signal with $k_p = 10$, find $m(t)$.
    (b) Determine the bandwidth of FM signal, if the maximum value of frequency deviation Δf is fixed at 75 kHz for commercial FM broadcasting by radio and modulation frequency is $\omega = 15$ kHz.

19. An angle modulated signal with carrier frequency $f_c = 10^5$ Hz is described by the equation
$$u(t) = 10\cos(2\pi f_c t + 5\sin 3000\pi t + 10\sin 2000\pi t)$$
Find
    i) The power of the signal of the modulated signal
    ii) The frequency deviation
    iii) Modulation index

20. In an FM carrier, modulating signal frequency and voltages are 400 Hz and 2.4 V respectively. For a modulation index 60, calculate maximum deviation. What is the modulation index when the modulation frequency is reduced to 250 Hz and modulation voltage is simultaneously raised to 3.2 V?

21. Carrier $A_c \cos\omega_c t$ is modulated by a signal $f(t) = 2\cos 2\pi 10^4 t + 5\cos 2\pi 10^3 t + 3\ cos 4\pi 10^4 t$. Find the bandwidth of the FM signal by using Carson's rule. Assume $k_f = 15\times 10^3$ Hz/V. Also find the modulation index?

22. An angle modulated wave is described by equation
$$u(t) = 10\cos\left(2\times 10^6\pi t + 10\cos 2000\pi t\right)$$
Find
    (a) The power of the modulated signal
    (b) The maximum frequency deviation
    (c) Maximum phase deviation
    (d) The bandwidth of the signal

23. A single-tone modulating signal $\cos(15\pi 10^3 t)$ frequency modulates a carrier of 10 MHz and produces a frequency deviation of 75 kHz. Find
    (a) The modulation index
    (b) Phase deviation produced in the FM wave
    (c) If another modulating signal produces a modulation index of 100 while maintaining the same deviation, find the frequency and amplitude of the modulating signal, assuming $k_f = 15$ kHz per volt.

24. A modulating signal $5\cos 2\pi 15\times 10^3 t$, angle modulates a carrier $A\cos\omega_c t$.
    (a) Find the modulation index and the bandwidth for a) FM b)PM
    (b) Determine the change in the bandwidth and the modulation index for both FM and PM, if fm is reduced to 5 kHz.

25. The normalized signal $m(t)$ has a bandwidth of 10000 Hz and its power content is 0.5 W. The carrier $A\cos 2\pi f_c t$ has a power content of 200 W.

    (a) If $m(t)$ modulates the carrier using SSB-AM, what will be the bandwidth and the power content of the modulated signal?

    (b) If the modulation scheme is DSB-SC, what will the answer to part (a)?

    (c) If the modulation scheme is AM with modulation index of 0.6, what will be the answer to part (a)?

    (d) If the modulation is FM with $k_f = 50000$, what will be the answer to part (a)?

26. An angle modulated signal has the form

$$s(t) = 100\cos\left[2\pi f_c t + 4\sin 2000\pi t\right]$$

    Where $f_c = 10$ MHz.

    a) Determine the average transmitted power.

    b) Determine the peak-phase deviation.

    c) Determine the peak-frequency deviation.

    d) Is this an FM or a PM signal? Explain.

27. An angle-modulated signal has the form $u(t) = 100\cos\left[2\pi f_c t + 4\sin 2\pi f_m t\right]$ Where $f_c =$ 10 MHz and $f_m = 1000$ Hz.

    a) Assuming that this is an FM signal, determine the modulation index and the transmitted signal bandwidth.

    b) Repeat a) if f_m is doubled.

    c) Assuming that this is an PM signal, determine the modulation index and the transmitted signal bandwidth.

    d) Repeat c) if f_m is doubled.

28. An Armstrong FM modulator is required in order to transmit an audio signal of bandwidth 50 Hz to 15 kHz. The Narrowband (NB) phase modulator used for this purpose utilizes a crystal controlled oscillator to provide a carrier frequency of $f_{c1} = 0.2$ MHz. The output of the NB phase modulator is multiplied by n1 by a multiplier and passed to a mixer with a local oscillator frequency $f_{c2} = 10.925$ MHz. The desired FM wave at the transmitter output has a carrier frequency $f_c = 90$ MHz, and a frequency deviation $\Delta f = 75$ kHz, which is obtained by multiplying the mixer output frequency with n_2 using another multiplier. Find n_1 and n_2. Assume that NBFM produces deviation of 25 Hz for the lowest baseband signal.

29. Show that a non-linear square-law device used for frequency multiplication of an FM signal doubles the carrier frequency as well as the frequency deviation?

30. (a) Consider the RC network shown in *Figure* 5.20. For the values of R and C given, we will show that for frequencies around 1.0 MHz, this can act as a differentiator.

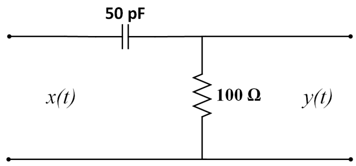


Figure 5.20: RC network

(b) Let us find the condition on R and C such that this network can act as a differentiator for frequencies around some frequency f_C

(c) For the above network, let $x(t) = s(t) = A_c \cos\left[2\pi \times 10^6 t + k_f m(t)\right]$ where $\left|k_f m(t)\right| \ll 10^6$ Hz. If $y(t)$ is envelope detected, show that, $m(t)$ can be recovered from the output.

31. For a modulating signal

$$m(t) = 2\cos 100t + 18\cos 2000\pi t$$

(a) Write expression for $s_{PM}(t)$ and $s_{FM}(t)$ when $A = 10$, $\omega_c = 10^6$, $k_f = 1000\pi$, and $k_p = 1$. For determining $s_{FM}(t)$, use the indefinite integral of $m(t)$, that is take the value of the integral at $t = -\infty$ to be 0.

(b) Estimate the bandwidth of $s_{PM}(t)$ and $s_{FM}(t)$

32. An angle modulated signal with carrier frequency $\omega_c = 2\pi \times 10^6$ is described by the equation

$$s(t) = 20\cos\left(\omega_c t + 0.1\sin 2000\pi t\right)$$

(a) Find the power of the modulated signal

(b) Find the frequency deviation (maximum) Δf

(c) Find the phase deviation $\Delta\phi$

(d) Estimate the bandwidth of $s(t)$

33. A signal $m(t)$ frequency modulates a 100 kHz carrier to produce NBFM signal as

$$s_{NBFM}(t) = 5\cos\left(2\pi \times 10^5 t + 0.0050\sin 2\pi \times 10^4 t\right)$$

Generate (block diagram design) the WBFM signal $s_{WBFM}(t)$ with a carrier frequency of 150 MHz and a peak frequency deviation of 100 kHz. Assume that the following are available for the design:

i) Frequency Multipliers of any (integer) value.

ii) A local oscillator whose frequency can be tuned to any value between 100 MHz to 300 MHz

iii) An ideal BPF with tunable centre frequency and bandwidth

Your block diagram must specify the carrier frequencies and frequency deviations at all points, as well as the centre frequency and bandwidth of the BPF.

34. An angle modulated signal is given by

$$s(t) = 5\cos(\omega_c t + 40\sin 500\pi t + 20\sin 1000\pi t + 10\sin 2000\pi t)$$

a) Determine the frequency deviation, Δf in Hz
b) Estimate the bandwidth in Hz of the angle modulated signal by Carson's rule.
c) If the angle modulated signal is a pulse modulated signal with the phase deviation constant, k_p is 5 radians per volt, determine the message signal $m(t)$.
d) If the angle modulated signal is a frequency modulated singal with a frequency deviation constant, k_f is 20,000π radians per volt, determine the message signal $m(t)$.

5.12 Textbooks and References

[1] Principles of Communication Systems by Taub and Schilling,2nd Edition. McGraw Hill.

[2] Communication Systems by Siman Haykin,4th Edition, John Wiley and Sons Inc.

[3] Modern Digital and Analog Communication System, by B. P. Lathi, 4th Edition, Oxford University Press.

[4] Principles of Communication Systems by Taub and Schilling,2nd Edition. McGraw Hill.

[5] Digital and analog communication systems, by L. W. Couch, 6th Edition, Pearson Education, Pvt. Ltd

[6] Analog Communications, First Edition 2009, A. P. Godse and U. A. Bakshi

[7] Analog Communication(Jntu) By Thomas-Chandrasekhar

[8] Communication Systems Engineering, Second Edition by John G. Proakis and Masoud Salehi

[9] Introduction to Analog and Digital Communications, Second Edition by Simon Haykin and Michael Moher

[10] Signals and Systems by Alan V. Oppenheim

[11] A Course in Electrical Technology (Electronic Devices and Circuits) Vol. III, Chapter21, 3rd Edition by J. B. Gupta

[12] Principles of Communication Lecture Notes, Indian Institutes of Technology, Madras by Prof V. Vankata Rao

[13] Electronic communications, 4th ed. by Dennis Roddy and John Coolen

Part III

Digital Communication Techniques

Chapter 6

Introduction to Digital Communications

***

One person can make a difference and every person should try.

–John F. Kennedy

6.1 Elements of Digital Communication System

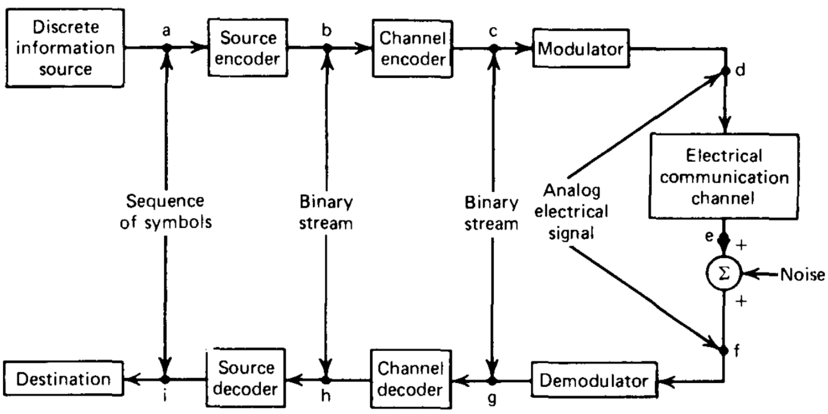


Figure 6.1: Elements of Digital Communication System

Figure 6.1 shows the functional elements of a digital communication system. The overall purpose of the system is to transmit the messages (or sequences of symbols) coming out of a source to a destination point at as high a rate and accuracy as possible. The source and the destination point are physically separated by communication channel that smears and distort the information bearing signal from the source. The information bearing signal is also corrupted by unpredictable electrical signals (noise) from both man-made and natural causes. The smearing and noise introduce errors in the information being transmitted and limits the rate at which information can be communicated from the source to the destination. The probability of incorrectly decoding a message symbol at the receiver is often used as a measure of performance of digital communication systems. The main function of the coder, the modulator, the demodulator, and the decoder is to combat the degrading effects of the channel on the signal and maximize the information rate and accuracy.

6.1.1 Information Source

As noted earlier, information sources can be classified into two categories based on the nature of their outputs: analog information sources and discrete information sources. Analog information sources, such as a microphone actuated by speech, or a camera capturing a scene, emit one or more continuous amplitude signals (or functions of time). The output of discrete information sources such -as a teletype or the numerical output of a computer consists of a sequence of discrete symbols or letters.

6.1.2 Source Encoder / Decoder

The input to the source encoder is a string of symbols occurring at a rate of R_s symbols/sec. The source coder converts the symbol (mostly statistically dependent) sequence into a binary sequence of 0's and 1's by assigning code words to the symbols in the input sequence. The Source encoder assigns a fixed-length binary code word to each symbol in the input sequence in order to achieve this – encoding. The optimum source encoder is designed to produce an output data rate approaching R, the source information rate, but due to practical constraints, the actual output rate of source is greater than the source information rate, R. The important parameters of a source coder are block size, code word lengths, average data rate, and the efficiency of the coder.

The decoder at the receiver converts the binary output of the channel decoder into a symbol sequence.

6.1.3 Modulator and Demodulator

The modulator accepts a bit stream as its input and converts it to an electrical waveform suitable for transmission over the communication channel. Modulation is basically used to minimize the effects of channel noise, to match the frequency spectrum of the transmitted signal with channel characteristics, to provide the capability to multiplex many signals, and to overcome some equipment limitations. The important parameters of the modulator are the types of waveforms used, the duration of the waveforms, the power level, and the bandwidth used. The modulator minimizes the effects of channel noise by the use of large signal power and bandwidth, and by the use of waveforms that last for longer durations.

Demodulator extracts message from the information bearing waveform produced by the modulator. Given the type and duration of waveforms used by the modulator, the power level at the modulator, the physical and noise characteristics of the channel, and the type of demodulation, a unique relationships between data rate, power bandwidth requirements, and the probability of incorrectly decoding a message bit can be derived. The characteristics of the modulator, the demodulator, and the channel establish an average bit error rate between points c and g in *Figure* 6.1.

6.1.4 Channel Encoder / Decoder

Digital channel coding is a practical method of realizing high transmission reliability and efficiency that otherwise may be achieved only by the use of signals of longer duration in the modulation/demodulation process. There are two methods of performing the channel coding operation. In the first method, called the block coding method, the encoder takes a block of k information bits from the source encoder and adds r error control bits. The number of error control bits added will depend on the value of k and the error control capabilities desired. In the second method, called the *convolutional coding* method, the information bearing message stream is encoded in it continuous fashion by continuously interleaving information bits and error control bits. These two methods require storage and processing of binary data at the encoder and decoder.

Important parameters of a channel encoder are the method of coding, rate or efficiency of the coder (as measured by the ratio of data rate at input to the data rate at the output), error control capabilities, and complexity of the encoder. The channel decoder recovers the information bearing bits from the coded binary stream. Error detection and possible correction is also performed by the channel decoder.

6.2 Sampling Theorem and Quantization

Continuous signal are made amenable for digital transmission by transforming the analog information source into digital symbols compatible with digital processing and transmission. The first step in the transformation involves discretization of the time axis, which sampling the continuous time signal at discrete values of time. Though time has been discretized by the sampling process the sample values are still analog, i.e., they are continuous variables. To represent the sample value by a digital symbol chosen from a finite set necessitates the choice of a discrete set of amplitudes to represent the continuous range of possible amplitudes. This process is known as **quantization** and unlike discretization of the time axis; this results in a distortion of the original signal since it is a many-to-one mapping. The measure of this distortion is commonly expressed by the signal power to quantization noise power ratio, SNR_q.

Sampling process converts an analog waveform into a sequence of discrete samples that are usually spaced uniformly in time. The basis of sampling is the Nyquist sampling theorem as outlined below.
Let a signal $x(t)$ be bandlimited to W Hz; that is $X(f) = 0$ for $|f| > W$. Let $x(nT_s) = x(t)|_{t=nT_s}$, $-\infty < n < \infty$ represent the samples of $x(t)$ at uniform intervals of T_s seconds. If $T_s \leq 1/2W$, then it is possible to reconstruct $x(t)$ exactly from the set of samples, $\{x(nT_s)\}$.

Alternatively, the sequence of $\{x(nT_s)\}$ can provide the complete time behavior of $x(t)$.
Now, let $f_s = 1/T_s$, then $f_s = 2W$ is the minimum sampling rate for $x(t)$. This minimum sampling rate is called *Nyquist rate*. **Nyquist criterion** is given by $f_s \geq 2W$.

If $x(t)$ is a sinusoidal signal with frequency, f_o, then $f_s > 2f_o$, $f_s = f_o$ is not adequate because if the two samples per cycle are at the zero crossing of the tone, then all the samples are zero.

6.2.1 Ideal or Impulse Sampling

A sampled waveform $m_s(t)$ is obtained by multiplying an analog signal, $m(t)$ by a periodic train of unit impulse function $s(t)$ (See *Figure* 6.2).
The impulse function is given by

$$s(t) = \sum_{n=-\infty}^{\infty} \delta\left(t - nT_s\right) \tag{6.2.1}$$

Thus, the resulting sampled waveform $m_s(t)$ is expressed as

$$\begin{aligned} m_s(t) &= m(t)s(t) \\ &= \sum_{n=-\infty}^{\infty} m(t)\delta\left(t - nT_s\right) = \sum_{n=-\infty}^{\infty} m(nT_s)\delta\left(t - nT_s\right) \end{aligned} \tag{6.2.2}$$

The period of the impulse train or *sampling period* in *Eqn.* 6.2.1 and 6.2.2 is a key parameter in the sampled signal. The inverse of the the sampling period is sampling frequency or sampling rate given by $f_s = 1/T_s$.

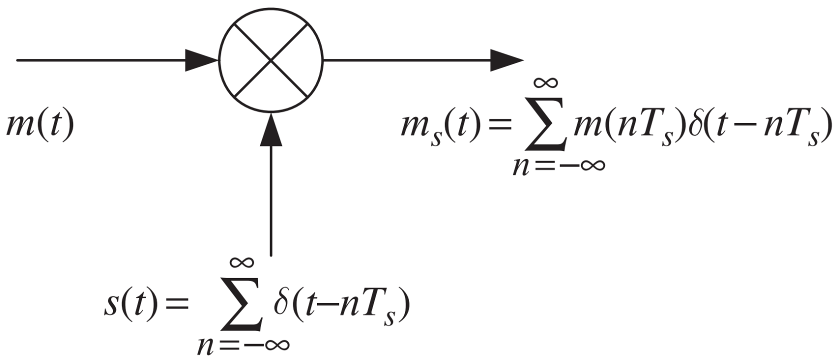


Figure 6.2: Ideal sampling – mathematical model

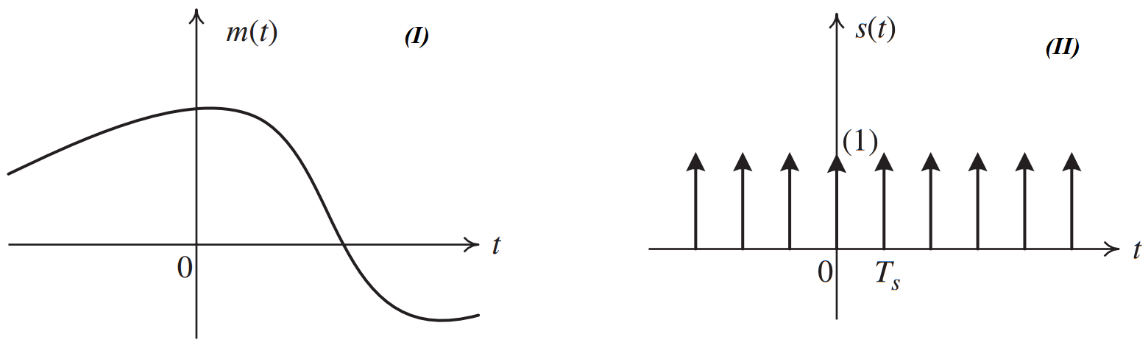
Figure 6.3a–b(I) graphically illustrate the ideal sampling process. It is intuitive that the higher the sampling rate, the more accurate the representation of $m(t)$ by $m_s(t)$. However, to achieve a high efficiency, it is desired to use low sampling rate as possible. Consider the Fourier transform of the sampled waveform $m_s(t)$. Since $m_s(t)$ is the product of $m(t)$ and $s(t)$, the Fourier transform of $m_s(t)$ is the convolution of the Fourier transforms of $m(t)$ and $s(t)$. The Fourier transform of $s(t)$ is given by

$$S(f) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} \delta\left(f - nf_s\right) \tag{6.2.3}$$

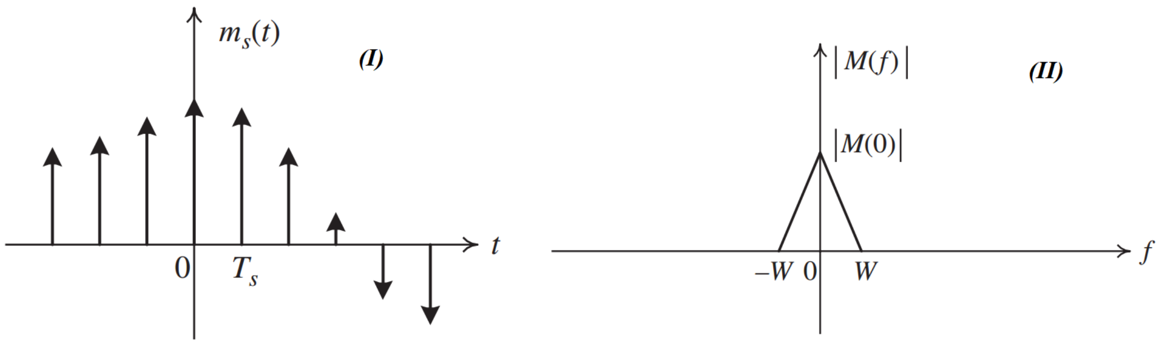
The transform of the sampled waveform $m_s(t)$ is given by

$$\begin{aligned} M_s(f) = M(f) * S(f) = M(f) * \left[\frac{1}{T_s} \sum_{n=-\infty}^{\infty} \delta\left(f - nf_s\right)\right] \\ = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} M\left(f - nf_s\right) \end{aligned} \tag{6.2.4}$$

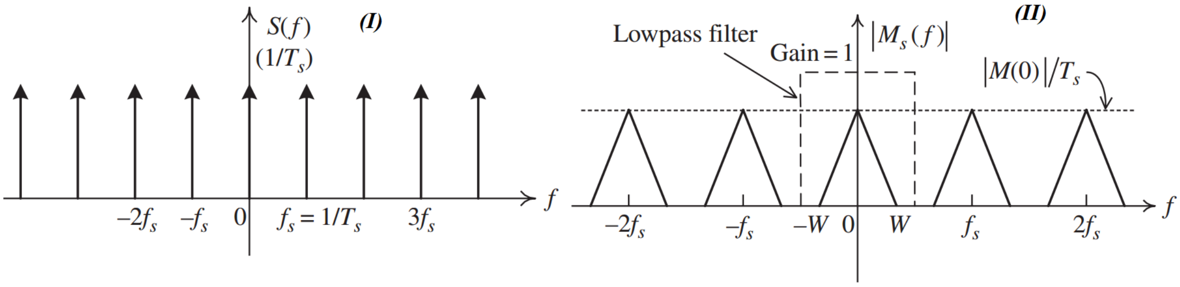
Eqn. 6.2.4 shows that, the spectrum of the sampled waveform consists of an infinite number of scaled and shifted copies of the spectrum of the original signal $m(t)$.



(a) I. analog signal, II. impulse train



(b) I. sampled version of the analog signal, II. spectrum of bandlimited signal



(c) I. spectrum of the impulse train, II. spectrum of the sampled waveform.

Figure 6.3: Ideal sampling process

Since within the original bandwidth (around zero frequency) the spectrum of the sampled waveform is the same as that of the original signal, it suggests that the original waveform $m(t)$ can be completely recovered from $m_s(t)$ by an ideal lowpass filter (LPF) of bandwidth W as shown in *Figure* 6.3c(II). It is important to note that, the condition for no overlapping of the copies of $M(f)$ is $f_s \geq 2W$, therefore the minimum sampling rate is $f_s = 2W$. When the sampling rate $f_s < 2W$ (**under-sampling**), then the copies of $M(f)$ overlap in the frequency domain and it is not possible to recover the original signal $m(t)$ by filtering. The distortion of the recovered signal due to under-sampling is referred to as **aliasing**. To avoid aliasing, the sample frequency must satisfy the Nyquist criterion. Then, the original signal $m(t)$ can be completely recovered from the sampled signal $m_s(t)$.

When the Nyquist criterion is satisfied, $m(t)$ can be recovered by passing $m_s(t)$ through an ideal LPF with bandwidth B, where B satisfies

$$W \leq B \leq f_s - W \tag{6.2.5}$$

Furthermore, the original signal can be reconstructed from the sampled signal by taking inverse Fourier transform of the signal. The reconstructed signal is then given by

$$m(t) = \sum_{n=-\infty}^{\infty} m(nT_s) \frac{\sin\left[2\pi W(t - nT_s)\right]}{\pi f_s (t - nT_s)} = \sum_{n=-\infty}^{\infty} m\left(\frac{n}{2W}\right) \sin c\,(2Wt - n) \tag{6.2.6}$$

6.2.2 Natural Sampling

A bandlimited analog signal $m(t)$ is sampled by multiplying the signal with a pulse train wave $x_p(t)$ that yields sampled-data sequence, $x_s(t)$. This gives as

$$\begin{aligned} m_s(t) &= m(t)x_p(t) \\ &= m(t) \sum_{n=-\infty}^{\infty} D_n e^{j2\pi n f_s t} \end{aligned} \tag{6.2.7}$$

The sampling rate f_s is equal to the inverse of the period, T_s of the pulse train, i.e. $f_s = 1/T_s$. In natural sampling, a bandlimited waveform can also be reconstructed from its sampled version as long as the sampling rate satisfies the *Nyquist criterion*. The Fourier transform of *Eqn.* 6.2.7 is

$$\begin{aligned} M_s(f) = \mathscr{F}\left\{m_s(t)\right\} &= \sum_{n=-\infty}^{\infty} D_n \mathscr{F}\left\{m(t)e^{j2\pi n f_s t}\right\} \\ &= \sum_{n=-\infty}^{\infty} D_n M\left(f - n f_s\right) \end{aligned} \tag{6.2.8}$$

$M_s(f)$ consists of an infinite number of copies of $M(f)$, which are periodically shifted in frequency every f_s hertz. However, the copies of $M(f)$ are not uniformly weighted (scaled) as in the case of ideal sampling, but rather they are weighted by the Fourier series coefficients of the pulse train.

Figure 6.4 shows various waveforms of natural sampling

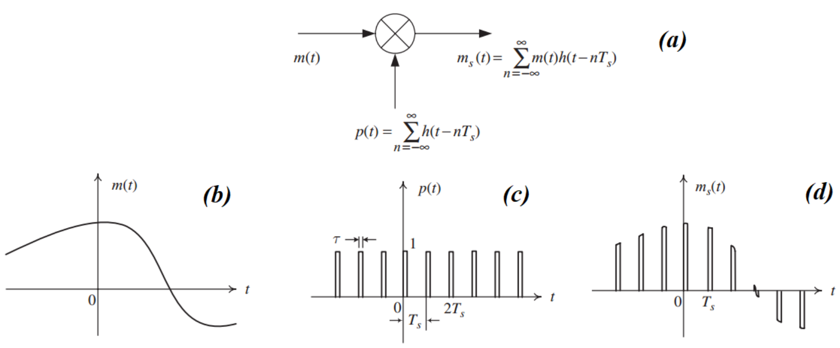


Figure 6.4: The natural sampling process: (a) mathematical model, (b) analog signal, (c) rectangular pulse train, (d) sampled version of the analog signal

6.2.3 Flat-Top Sampling

Flat-top sampling process involves two simple operations:

1. Instantaneous sampling of the analog signal $m(t)$ every T_s seconds. As in the ideal and natural sampling cases, the sampling rate, $f_s = 1/T_s$ must satisfies Nyquist criterion in order to reconstruct the original signal $m(t)$ from the sampled version.
2. Maintaining the value of each sample for a duration of τ seconds.

The expression of interest is $m_s(t)$, given by

$$m_s(t) = \left[m(t) \sum_{n=-\infty}^{\infty} \delta\left(t - nT_s\right) \right] * h(t) \tag{6.2.9}$$

The process is outlined in *Figure* 6.5 below

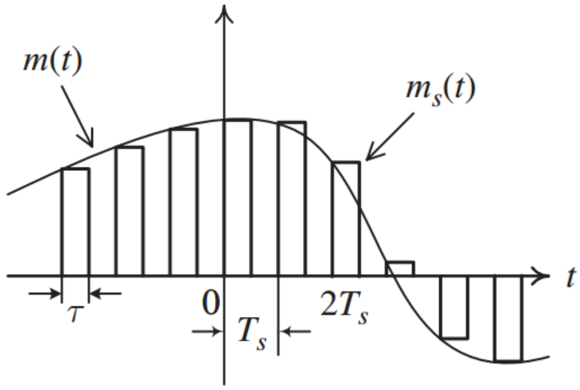
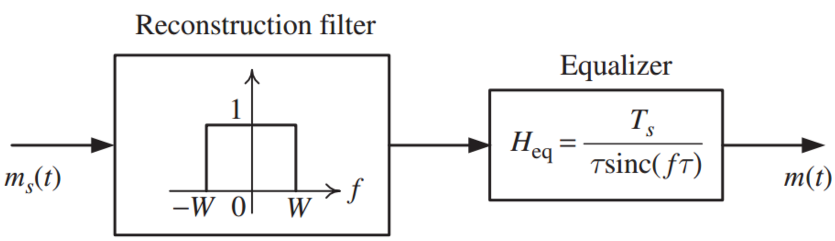
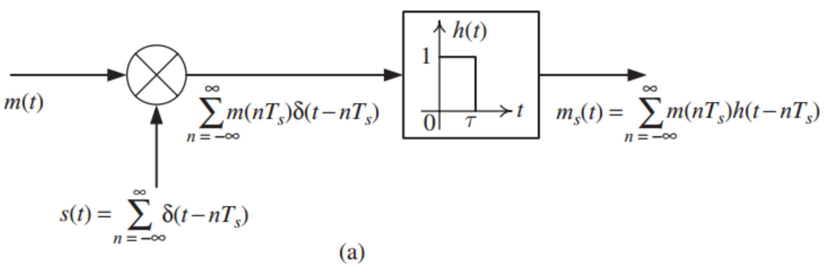


Figure 6.5: Flat-top sampling:- $m(t)$ is the original analog signal and $m_s(t)$ is the sampled signal.



(a) The flat-top sampling process



(b) Reconstruction

Figure 6.6: Reconstruction

The Fourier transform of $m_s(t)$ is given by

$$\begin{aligned} M_s(f) &= \mathscr{F}\left\{m(t)\sum_{n=-\infty}^{\infty}\delta\left(t-nT_s\right)\right\}\mathscr{F}\left\{h(t)\right\} \\ &= \frac{1}{T_s}H(f)\sum_{n=-\infty}^{\infty}M\left(f-nf_s\right) \end{aligned} \tag{6.2.10}$$

If the Nyquist criterion is satisfied, then passing the flat-top sampled signal through an LPF (with a bandwidth of W) produces the signal whose Fourier transform is $(1/T_s)M(f)H(f)$. Thus, the distortion due to $H(f)$ can be corrected by connecting an equalizer in cascade with the lowpass reconstruction filter.

6.3 Quantization

To obtain fully digital representation of a continuous signal, requires quantization of the amplitude of the sampled signal and encoding of the quantized values as illustrate in *Figure* 6.7

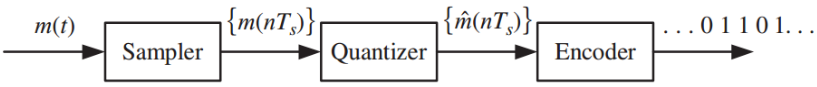


Figure 6.7: Quantization and encoding operations

By definition, amplitude quantization is the process of transforming the sample amplitude $m(nT_s)$ of a message signal $m(t)$ at time $t = nT_s$ into a discrete amplitude $\hat{m}(nT_s)$ taken from a finite set of possible amplitudes. Approximated (or quantized) signal, $\hat{m}(nT_s)$, can be made indistinguishable from the continuous sampled signal, $m(nT_s)$, if the finite set of amplitudes is chosen such that the spacing between two adjacent amplitude levels is sufficiently small.

Figures 6.8(a) and 6.8(b) display two uniform quantizer characteristics, called midtread and midrise. As can be seen from these figures, the classification whether a characteristic is midtread or midrise depends on whether the origin lies in the middle of a tread, or a rise of the staircase characteristic. For both characteristics, the decision levels are equally spaced and

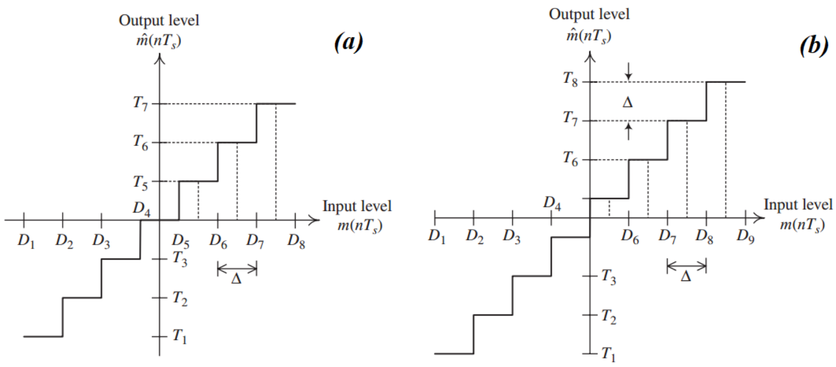


Figure 6.8: Two types of uniform quantization: (a) midtread and (b) midrise

the lth target level is the midpoint of the lth interval. The amplitude level is given by

$$A_l = \frac{D_l + D_{l+1}}{2} \tag{6.3.1}$$

It should be noted that quantization process always introduces an error called *quantization error*, thus, the performance of a quantizer is often evaluated in terms of its *SNR*. Considering memoryless (current state is independent of earlier or later state) quantization, the input and output will be denoted by m and $\hat{m}$. Let m represent the amplitude of the output of uniform quantizer of midrise type within the range $-m_{max} \leq m \leq m_{max}$ and L as the number of quantization levels. The quantization step-size is given by

$$\Delta = \frac{2m_{\max}}{L} \tag{6.3.2}$$

Figure show a plot of the input and output waveforms of a midrse uniform quantizer.

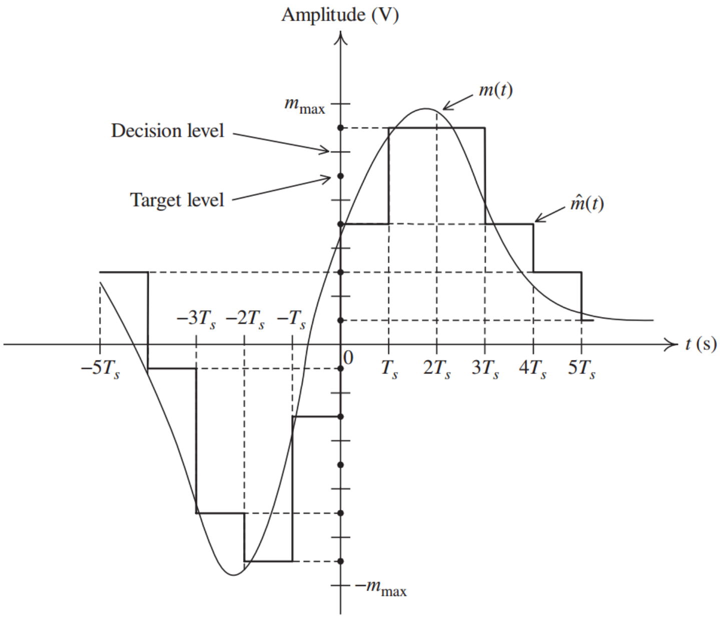


Figure 6.9: Input and output waveform of a midrise uniform quantizer

The number of quantization levels is usually chosen to be a power of 2, i.e., $L = 2^R$, where R is the number of bits needed to represent each target level. The SNR_q is expressed as

$$\begin{aligned} SNR_q &= \left(\frac{2\sigma_m^2}{m_{\max}^2}\right) 2^{2R} \\ &= \frac{3 \times 2^{2R}}{F^2} \end{aligned} \tag{6.3.3}$$

Where

$$F = \frac{\text{peak value of the signal}(m_{\max})}{\text{RMS value of the signal}}$$

.

Eqn. 6.3.3 shows that the SNR_q of a uniform quantizer increases exponentially with the number of bits per sample R and decreases with the square of the message's crest factor. Quantization error introduced by the quantizer is given by

$$q = m - \hat{m} \tag{6.3.4}$$

Eqn. 6.3.4 lies in the interval $\Delta/2 \leq q \leq \Delta/2$. If the step-size Δ is sufficiently small (i.e., the number of quantization intervals L is sufficiently large).

6.4 Pulse Code Modulation (PCM)

6.5 Baseband Data Transmission

6.6 Pulse Amplitude Modulation

6.7 Basic Digital Carrier Modulation

6.8 Review Questions

6.9 Textbook and References

Part IV

Random Processes and Noise in Communication

Chapter 7

Random Processes

***

Strive not to be a success, but rather to be of value.

– Albert Einstein

Random variable is used to describe process of assigning a number to the outcome of a random experiment. In general, a function whose domain is a sample space and whose range is a set of real numbers is called a **random variable** of the experiment. That is, for events in a random variable assigns a subset of the real line.

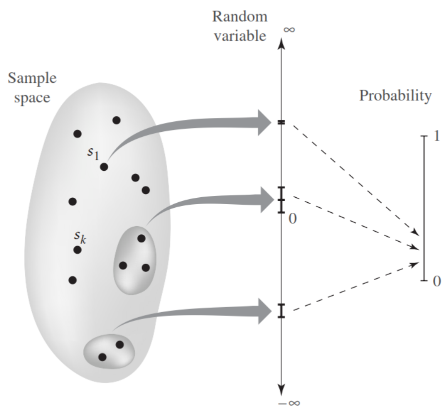


Figure 7.1: Illustration of the relationship between sample space, random variables, and probability.

Thus, if the outcome of the experiment is t, we denote the random variable as $X(t)$ or just X. Where X is a function called a random variable.

The concept of a random variable is illustrated in Figure. 8.3, where we have suppressed the events but show subsets of the sample space being mapped directly to a subset of the real

line. The probability function applies to this random variable in exactly the same manner that it applies to the underlying events. The benefit of using random variables is that probability analysis can now be developed in terms of real-valued quantities regardless of the form or shape of the underlying events of the random experiment. Random variables may be discrete and take only a finite number of values, such as in the coin-tossing experiment. Alternatively, random variables may be continuous and take a range of real values.
The cumulative probability distribution function or distribution function of a random variable is defined as

$$F_X(x) = P[X \leq x] \tag{7.0.1}$$

The function is a function of x, not of the random variable X. However, it depends on the assignment of the random variable X, which accounts for the use of X as subscript. For any point x, the distribution function expresses the probability that the value of X is less than x. The distribution function has two basic properties,

1. The distribution function is bounded between zero and one.

2. The distribution function is a monotone nondecreasing function of x

$$F_X(x_1) \leq F_X(x_2), \qquad \text{if} x_1 \leq x_2 \tag{7.0.2}$$

If X is a continuous-valued random variable and $F_X(x)$ is differentiable with respect to x, then a third commonly used function is the **probability density function**, denoted by $f_X(x)$, where

$$f_X(x) = \frac{\partial}{\partial x} F_X(x) \tag{7.0.3}$$

A probability density function has three basic properties:

1. Since the distribution function is monotone nondecreasing, it follows that the density function is nonnegative for all values of x.

2. The distribution function may be recovered from the density function by integration, as shown by

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt \tag{7.0.4}$$

$$\Rightarrow \qquad P(a < X \leq b) = F(b) - F(a) = \int_{a}^{b} f_X(t)dt$$

3. Property 2 implies that the total area under the curve of the density function is unity.

The PDF is always a nonnegative function with total area = 1.
Consider a strictly stationary random process $X(t)$. We define the mean of the process $X(t)$ as the expectation of the random variable obtained by observing the process at some time t, as shown by

$$\begin{aligned} \mu_X(t) &= E[X(t)] \\ &= \int_{-\infty}^{\infty} x f_{X(t)}(x)dx \end{aligned} \tag{7.0.5}$$

where $f_{x(t)}(t)$ is the first-order probability density function of the process. As $f_{x(t)}$ is independent of time it means the mean of strictly stationary is a constant

$$\mu_x(t) = \mu_x \qquad \text{for all} \;\; t$$

From the definition of auto-correlation function, we can write

$$\begin{aligned} R_x(t_1,t_2) &= E\left[X(t_1)X(t_2)\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 f_{X(t_1)X(t_2)}(x_1,x_2)dx_1dx_2 \end{aligned} \tag{7.0.6}$$

where $f_{X(t_1)X(t_2)}(x_1,x_2)$ is the second-order probability density function of the process. The auto-correlation function of a strictly stationary process depends only on the time difference $t_2 - t_1$ as

$$R_X(t_1,t_2) = R_X(t_2 - t_1) \qquad \text{for all } t_1 \text{ and } t_2 \tag{7.0.7}$$

The autocorrelation function for strictly stationary process $X(t)$ is

$$\begin{aligned} C_X(t_1,t_2) &= E\left[(X(t_1)-\mu_X)(X(t_2-\mu_X))\right] \\ &= R_X(t_2-t_1)-\mu_X^2 \end{aligned} \tag{7.0.8}$$

Example 7.40

Consider a sinusoidal signal with phase, defined by $X(t) = A\cos(2\pi f_c t + \vartheta)$, where A and f_c are constants and ϑ is a random variable that is uniformly distributed over the interval $[-\pi,\ \pi]$, thus

$$f_\vartheta(\theta) = \begin{cases} \frac{1}{2\pi}, & -\pi \le \theta \le \pi \\ 0, & \text{elsewhere} \end{cases}$$

The random variable denotes the phase difference between the locally generated carrier and the sinusoidal carrier wave used to modulate the message signal in the transmitter. The auto-correlation function of $X(t)$ is

$$\begin{aligned} R_X(\tau) &= E\left[X(t+\tau)X(t)\right] \\ &= \frac{A^2}{2}E\left[A^2\cos(2\pi f_c t + 2\pi f_c\tau + \vartheta)\cos(2\pi f_c t + \vartheta)\right] \\ &= \frac{A^2}{2}\int_{-\pi}^{\pi}\frac{1}{2\pi}\cos(4\pi f_c t + 2\pi f_c\tau + 2\vartheta)\,d\theta + \frac{A^2}{2}\cos(2\pi f_c\tau) \end{aligned}$$

$$R_X(\tau) = \frac{A^2}{2}\cos(2\pi f_c\tau)$$

We see therefore that the autocorrelation function of a sinusoidal wave with random phase is another sinusoid at the same frequency in the ”τ domain” rather than the original time domain.

7.1 Statistical Tools

7.1.1 Variance

The variance of a random variable is an estimate of the spread of the probability distribution about the mean. For discrete random variables, the variance, σ_X^2 is given by the expectation of the squared distance of each outcome from the mean value of the distribution.

$$\begin{aligned}\sigma_X^2 &= Var(X)\\ &= E\left[(X-\mu_X)^2\right]\\ &= \sum_X (x-\mu_X)^2 f_X(x)dx\end{aligned} \tag{7.1.1}$$

For a continuous random variable with density function $f_X(x)$, the analogous definition of variance is given by

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)dx \tag{7.1.2}$$

Example 7.41

Roll a dice: $k = 1:6$ Assign to the event $A_k(t)$ a random process given by the function

$$x_(t) = \cos k\omega_o t$$

Evaluate the time statistics:
MEAN

$$M^t\{x_k(t)\} = \lim_{T\to\infty}\frac{1}{T}\int_0^T a\cos k\omega_o t dt = 0$$

VARIANCE

$$V^t\{x_k(t)\} = \lim_{T\to\infty}\frac{1}{T}\int_0^T a^2\cos^2 k\omega_o t dt = \frac{a^2}{2}$$

CORRELATION

$$R^t\{x_k(t)\} = \lim_{T\to\infty}\frac{1}{T}\int_0^T a^2\cos^2(k\omega_o t)\cos(k\omega_o(t+\tau))dt$$

$$= \frac{a^2}{2}\cos k\omega_o\tau$$

Looking at the correlation function then we see that if $k\omega_o t = \pi/2$ then the correlation is zero – for this example it would be the same as taking the correlation of a sine with cosine, since cosine is simply the sine function phase-shifted by $\pi/2$, and cosine and sine are not correlated.

Now if we look at the STATISTICS of the random process, for some time $t = t_o$,

$$x_k(\zeta,t_o) = a\cos(k\omega_o t_o) = y_k(\zeta)$$

where k is the random variable ($k = 1, 2, 3, 4, 5, 6$) and each event has probability, $p_i = 1/6$.

7.1.2 Statistical or Expected Average

Consider the signal $m(t)$, the expected average can be found by

$$\mu_x(t) = \frac{1}{T}\int_0^T E\left[X(t)\right]dt \tag{7.1.3}$$

7.1.3 Ergodic Process

Consider $x(t)$, the expectations or ensemble averages of a random process $X(t)$ are averages across the process. The relationship between time average and ensemble average as is usually given as

$$\mu_x(t) = \frac{1}{2T}\int_{-T}^{T} x(t)dt \tag{7.1.4}$$

Since $X(t)$ is assumed to be stationary, the mean of the time average $\mu_x(t)$ is given by

$$\begin{aligned} E\left[\mu_x(t)\right] &= \frac{1}{2T}\int_{-T}^{T} x(t)dt \\ &= \frac{1}{2T}\int_{-T}^{T} \mu_x dt = \mu_x \end{aligned} \tag{7.1.5}$$

We can say that the process $X(t)$ is ergodic in the mean if

- $\lim_{T\to\infty} \mu_x(t) = \mu_x$
- $\lim_{T\to\infty} Var\left[\mu_x(t)\right] = 0$

Time average auto-correlation function of $x(t)$ is

$$R_x(\tau,t) = \frac{1}{2T}\int_{-T}^{T} x(t+\tau)x(t)dt, \qquad -T \le t \le T \tag{7.1.6}$$

Hence, $x(t)$ is ergodic in the autocorrelation function if the following limiting conditions are satisfied

$$\lim_{T\to\infty} R_x(\tau,t) = R_x(\tau)$$
$$\lim_{T\to\infty} Var\left[R_x(\tau,t)\right] = 0$$

7.2 Gaussian Random Variables

The Gaussian random variable (also called normal random variables) plays an important role in many applications and is by far the most commonly encountered random variable in the statistical analysis of communication systems. A Gaussian random variable is a continuous random variable with a density function given by

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right\} \tag{7.2.1}$$

where the Gaussian random variable X has mean and variance σ_X^2. This density function extends from to and is symmetric about the mean μ_X.

For special case of a Gaussian random variable with mean of zero, $\mu_X = 0$, and a variance of unity, $\sigma_X^2 = 1$, the density function is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \qquad -\infty < x < \infty \tag{7.2.2}$$

Eqn. 7.2.2 is depicted in *Figure* 7.2

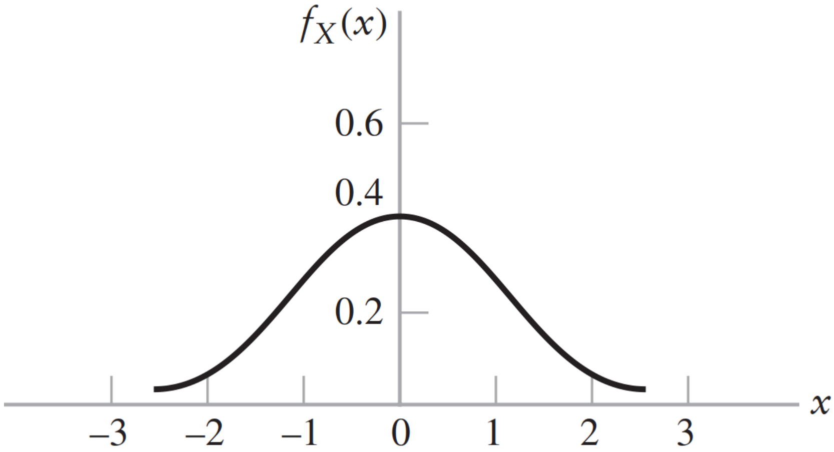


Figure 7.2: The normalized Gaussian distribution – the probability density function.

7.3 White Noise

The noise analysis of communication systems is often based on an idealized noise process called white noise. The power spectral density of white noise is independent of frequency. White noise is analogous to the term ”white light” in the sense that all frequency components are present in equal amounts or equal power, i.e. the power-spectral density is a constant for all frequencies.

A process $X(t)$ is called **white process** if it has flat spectral density, that is if $S_X(f)$ is constant for all f The importance of white processes in practice stems from the fact that thermal noise can be closely modeled as a white process over a wide range of frequencies. Also, a wide range of processes used to describe a variety of information sources can be modeled as the output of LTI systems driven by a white process.
The power spectral density of a white noise process is given by

$$S_X(f) = \frac{N_o}{2} \tag{7.3.1}$$

where the factor $1/2$ has been included to indicate that half the power is associated with positive frequencies and half with negative frequencies, as illustrated in *Figure* 7.3. The power content of a white process is

$$P_X = \int_{-\infty}^{\infty} S_X(f)df = \int_{-\infty}^{\infty} \frac{N_o}{2} df = \infty \tag{7.3.2}$$

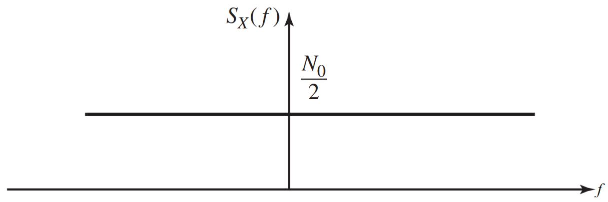


Figure 7.3: Power-spectral density of a white process

The autocorrelation function function of white noise is

$$R_X(\tau) = \frac{N_o}{2}\delta(\tau) \tag{7.3.3}$$

The autocorrelation function of white noise consists of a delta function weighted by the factor $N_o/2$ located at $\tau = 0$ as in *Figure* 7.4.

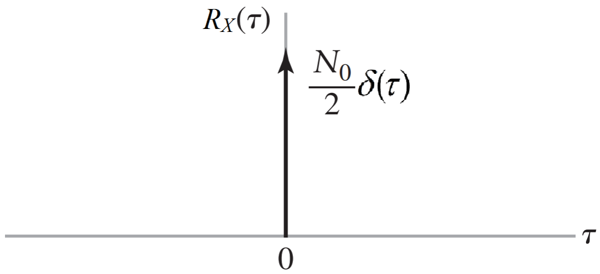


Figure 7.4: Autocorrelation function of white noise

7.4 Noise-Equivalent Bandwidth

When white Gaussian noise passes through a filter, the output process, although still Gaussian, will not be white anymore. The filter characteristics determine the spectral properties of the output process, and we have

$$S_y(f) = S_x(f)|H(f)|^2 = \frac{N_o}{2}|H(f)|^2 \tag{7.4.1}$$

where $|H(f)|$ is the response of the filter.
The power content (noise power) of the output process can be found by integration as

$$\begin{aligned} P_N &= \int_{-\infty}^{\infty} S_y(f)df \\ &= \int_{-\infty}^{\infty} \frac{N_o}{2}|H(f)|^2 df \\ &= N_o \int_0^{\infty} |H(f)|^2 df, \quad \Rightarrow \text{ If } |H(f)| \text{ is even fxn} \end{aligned} \tag{7.4.2}$$

Note: $S_x(f) = \dfrac{N_o}{2}$ is the PSD of the input signal.
Now, consider the same noise source connected to the input of ideal LPF of zero-frequency

response $H(0)$ or with maximum gain H_o, that is maximum transfer, H_{max} of $|H(f)|$ and bandwidth B_{Neq} as illustrate in *Figure* 7.5a. In this case the average output noise power is

$$\begin{aligned} P_N &= \frac{N_o}{2}\int_{-\infty}^{\infty} |H(f)|^2 df \\ &= \frac{N_o}{2} \times 2B_{Neq}H_{\max}^2 \\ &= N_o B_{Neq} H_{\max}^2 \end{aligned} \tag{7.4.3}$$

Thus, the **noise-equivalent bandwidth** of the filter with frequency response $H(f)$ as

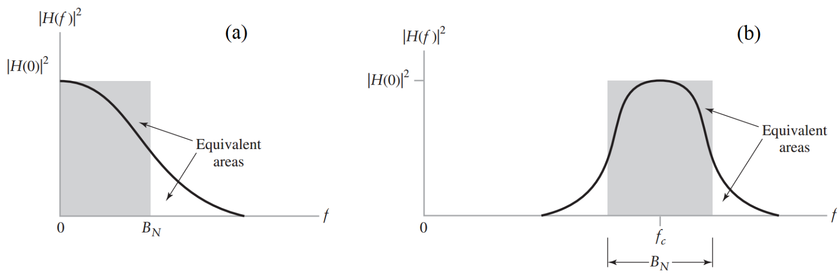


Figure 7.5: (ab) Illustration of arbitrary lowpass filter and ideal low-pass filter $H(f)$ of bandwidth B_{Neq}. (b) Illustration of arbitrary bandpass filter and ideal bandpass filter of bandwidth B_{Neq}

$$\begin{aligned} B_{Neq} &= \frac{\int_{-\infty}^{\infty} |H(f)|^2 df}{2H_{\max}^2} = \frac{1}{2H_{\max}^2}\int_{-\infty}^{\infty} |H(f)|^2 df \\ &= \frac{1}{H_{\max}^2}\int_0^{\infty} |H(f)|^2 df \end{aligned} \tag{7.4.4}$$

In a similar way, we may define the noise-equivalent bandwidth for a band-pass filter, as illustrated in *Figure* 7.5b. Thus, the noise-equivalent bandwidth for a band-pass filter may be defined as

$$B_{Neq} = \frac{\int_0^{\infty} |H(f)|^2 df}{|H(f_c)|^2} \tag{7.4.5}$$

where $|H(f_c)|^2$ is the center-frequency amplitude response of the filter. In fact, it can be seen *Eqn.* 7.4.5 can be used to represent both cases by setting $f_c = 0$ for LPF. Then, generally

$$P_N = N_o |H(f_c)|^2 B_{Neq} \tag{7.4.6}$$

and the effect of passing white noise through a filter may be separated in two parts:

1. The center-frequency power gain $|H(f_c)|^2$
2. The noise-equivalent bandwidth B_{Neq}, representing the *frequency selectivity* of the filter.

This separation applies whether the filter is low-pass or band-pass.

Example 7.42

Consider the single-pole LPF shown *Figure* 7.6. Find the noise-equivalent bandwidth of the filter.

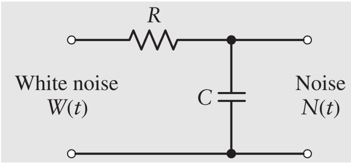


Figure 7.6: Single-pole LPF

Solution
The transfer function of the RC filter is

$$H(f) = \frac{1}{1 + j2\pi fRC}$$

The noise-equivalent bandwidth filter is

$$\begin{aligned} B_{Neq} &= \frac{1}{|H(f_c)|^2} \int_0^\infty |H(f)|^2 df \\ &= 1 \times \int_0^\infty \frac{df}{1 + (2\pi fRC)^2} \end{aligned}$$

Recognising the integrand as the scaled derivative of $\tan^{-1}(f)$, we obtain

$$\begin{aligned} B_{Neq} &= \frac{1}{2\pi RC} \tan^{-1}(2\pi fRC)\Big|_0^\infty \\ &= \frac{1}{2\pi RC}\left(\frac{\pi}{2} - 0\right) \\ &= \frac{1}{4RC} \end{aligned}$$

7.5 Textbooks and References

[1] Principles of Communication Systems by Taub and Schilling,2nd Edition. McGraw Hill.

[2] Communication Systems by Siman Haykin,4th Edition, John Wiley and Sons Inc.

[3] Modern digital and analog communication system, by B. P. Lathi, 3rd Edition, Oxford University Press.

[4] Digital and analog communication systems, by L. W. Couch, 6th Edition, Pearson Education, Pvt. Ltd

[5] Introduction to Analog and Digital Communications, Second Edition, by Simon Haykin and Michael Moher.

Chapter 8

Noise in Communication Systems

***

> Dreaming, after all, is a form of planning.
>
> – Gloria Steinem

The following assumption are made in the treatment of noise effect on communication system

Channel

i) Distortionless

ii) Additive white Gaussian noise (AWGN)

Receiver

i) Ideal bandpass filter centre freq. f_c

ii) Ideal demodulator

iii) A detector

iv) Ideal lowpass filter

Component of noise

$$n(t) = \underbrace{n_I(t)\cos\omega_c t}_{\text{in}-\text{phasecomponent}} - \underbrace{n_Q(t)\sin\omega_c t}_{\text{quadraturecomponent}} \tag{8.0.1}$$

The quadrature component is normally filtered out.
Let the message $x(t)$ be a LP signal generated from an ergodic process with bandwidth W. $x(t)$ can be normalized such that

$$|x(t)| \leq 1, P_x = \overline{x^2} = \left\langle x^2(t) \right\rangle \leq 1 \tag{8.0.2}$$

Let L be the transmission loss of a channel which provide nearly distortionless and negligible time delay.
Let $x_c(t)$ be carrier with amplitude A_c represent the output of the channel, the received signal power is then given as

$$P_R = \frac{P_T}{L} = \overline{x_c^2} \tag{8.0.3}$$

whereas, the transmit power is

$$P_T = \sqrt{L}x_c(t) = \sqrt{L}A_c \tag{8.0.4}$$

Let model the predetection stage as a bandpass filter with $H_R(f)$ having unity gain over the transmission bandwidth B_T. Under this conditions, we can assume that the addictive noise is at the receiver's input, the total signal-plus-noise at the detector is given as

$$v(t) = x_c(t) + n(t) \tag{8.0.5}$$

Where $n(t)$ represent the predetection noise.

The last stage in the diagram is LPF. This has a transfer function of $H_D(f)$ that gives output of $y_D(t)$ at the destination. The LPF stage is called post-detection stage.

8.1 Bandpass Noise

Although we must give individual attention to the effects of noise on specific types of analog modulation, all analog continuous wave communication systems have the same general structure and suffer from bandpass noise. We will assume that the noise comes from an additive white Gaussian noise (AWGN) process.

8.1.1 System Models

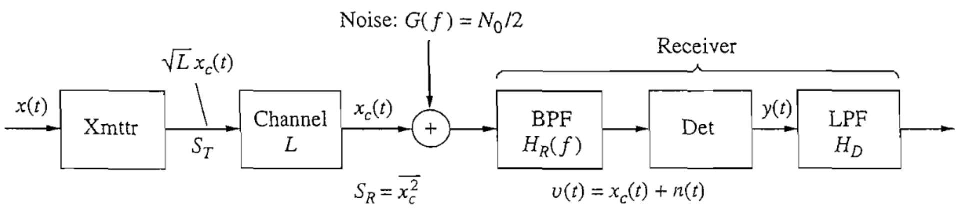


Figure 8.1: Model of a continuous wave communication system with noise

Eqn. 8.0.5 can be written as

$$\overline{v^2} = \overline{x_c^2} + \overline{n^2} = P_R + N_R \tag{8.1.1}$$

With the assumption that the signal and the noise are independent, then $N_R = \overline{n^2}$ is the predetection noise power. Here (*Figure* 8.2), let's treat the channel noise plus any noise generated in the predetection portion of the receiver as being equivalent to white noise.

The filtered output $n(t)$ has spectral density

$$S_n(f) = \frac{N_o}{2}|H_R(f)|^2 \tag{8.1.2}$$

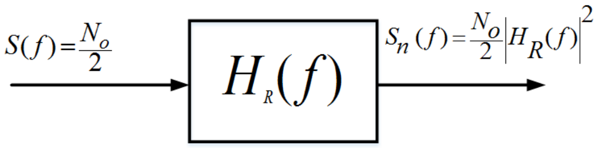


Figure 8.2: Block diagram of BPF filtered white noise

It is important to note that a bandpass noise is obtained when a white noise is passed through a BPF.

The spectrum diagram (*Figure* 8.3) is based upon a predetection filter with nearly square frequency response, so its noise bandwidth equals B_T and noise power of

$$P_{no} = N_R = \int_{-\infty}^{\infty} S_n(f)df = N_o B_T \tag{8.1.3}$$

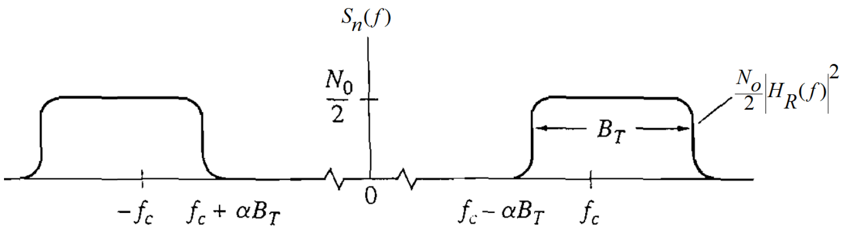


Figure 8.3: PSD of filtered white noise

Note that the carrier freq.f_c does not necessarily fall at the center of the passband of the noise spectrum. To consider symmetry, we have the lower cutoff freq as $f_c - \alpha B_T$ take care of symmetric-sideband case ($\alpha = 1/2$) and the suppressed-sideband case ($\alpha = 0$ *or* 1). It should be note again that the value of B_T depends upon the message bandwidth W and the type of modulation.

Eqn. 8.1.3 can be written to consider certain freq. range as

$$P_{no} = \int_{-B_T}^{B_T} \frac{N_o}{2} df = N_o B_T \tag{8.1.4}$$

Let define the predetection signal-to-noise ratio from the analysis above

$$\left(\frac{S}{N}\right)_R \triangleq \frac{P_R}{N_R} = \frac{P_R}{N_o B_T} \tag{8.1.5}$$

Signal-to-noise at the destination is of a baseband transmission system

$$\left(\frac{S}{N}\right)_b \triangleq \frac{P_R}{N_R} = \frac{P_R}{N_o W} \tag{8.1.6}$$

Where W is the modulating signal bandwidth, and $B_T \geq W$

Note well that the interpretation that λ equals the maximum destination *S/N* for analog baseband transmission with identical values of P_R and N_o at the receiver.

The relationship between predetection SNR and destination SNR is given by

$$\left(\frac{S}{N}\right)_R = \frac{W}{N_R}\gamma, \qquad \gamma = \frac{P_R}{N_o W} \tag{8.1.7}$$

this implies $\left(S/N\right)_R \leq \gamma$ since $B_T \geq W$, and γ is system definition parameter.

8.2 Noise on Linear Modulation

8.2.1 Effect of Noise on DSB-SC AM

In DSB, the modulated signal is

$$s(t) = A_c \cos\left(2\pi f_c t + \phi_c\right) \tag{8.2.1}$$

The received signal at the output of the receiver noise-limiting filter is

$$\begin{aligned} r(t) &= s(t) + n(t) \\ &= A_c m(t)\cos\left(2\pi f_c t + \phi_c\right) + n_c(t)\cos 2\pi f_c t - n_s(t)\sin 2\pi f_c t \end{aligned} \tag{8.2.2}$$

The lowpass filter rejects the double-frequency components and passes only the lowpass components. Hence, the output is

$$y(t) = \frac{1}{2}A_c m(t)\cos\left(\phi_c - \phi\right) + \frac{1}{2}\left[n_c(t)\cos\phi + n_s(t)\sin\phi\right] \tag{8.2.3}$$

The output of coherent or synchronous demodulator when noise is considered is given by

$$y(t) = \frac{1}{2}\left[A_c m(t) + n_c(t)\right] \tag{8.2.4}$$

Therefore, at the receiver output the message signal and the noise components are additive and we are able to define a meaningful SNR. The message signal power is given by

$$P_o = \frac{A_c^2}{4}P_m \tag{8.2.5}$$

and the noise power is given by

$$P_{n_o} = \frac{1}{4}P_n \tag{8.2.6}$$

Where P_n is the noise, $n(t)$, power, which is $P_n = N_o B_T$, if the bandwidth of the predetection filter is $B_T = 2W$, then

$$P_n = 2N_o W \tag{8.2.7}$$

Now, the output SNR is

$$\begin{aligned} \left(\frac{S}{N}\right)_o &= \frac{P_o}{P_{n_o}} \\ &= \frac{\frac{A_c^2}{4}P_m}{\frac{1}{4}2WN_o} \\ &= \frac{A_c^2 P_m}{2WN_o} \end{aligned} \tag{8.2.8}$$

In this case, the received signal power is $P_m = \frac{A_c^2 P_m}{2}$. Therefore the output signal-to-noise ratio of DSB system can be written as

$$\left(\frac{S}{N}\right)_{oDSB} = \frac{P_R}{N_o W} = \left(\frac{S}{N}\right)_b \tag{8.2.9}$$

It can be seen that the output SNR is the same as the SNR for a baseband system. Therefore, DSB-SC AM does not provide any SNR improvement over a simple baseband communication system.

8.2.2 Effect of Noise on SSB AM

The SSB modulated signal is

$$s(t) = A_c m(t)\cos 2\pi f_c t \pm A_c \hat{m}(t)\sin 2\pi f_c t$$

Therefore, the output of the demodulator is

$$r(t) = A_c m(t)\cos 2\pi f_c t \pm A_c \hat{m}(t)\sin 2\pi f_c t + n(t) \tag{8.2.10}$$

Assuming an ideal LPF, the output of the lowpass filter is

$$y(t) = \frac{A_c}{2} m(t) + \frac{1}{2} n_c(t) \tag{8.2.11}$$

It is observed that, in this case, again the signal and the noise components are additive and a meaningful SNR at the receiver output can be defined. The output noise power for SSB is

$$P_n = W N_o \tag{8.2.12}$$

Therefore,

$$\left(\frac{S}{N}\right)_{oSSB} = \frac{P_o}{P_{n_o}} = \frac{A_c^2 P_m}{N_o W} \tag{8.2.13}$$

But in SSB,

$$P_R = P_U = A_c^2 P_m \tag{8.2.14}$$

and hence,

$$\left(\frac{S}{N}\right)_{oSSB} = \frac{P_R}{N_o W} = \left(\frac{S}{N}\right)_b \tag{8.2.15}$$

From *Eqn.* 8.2.15 the SNR in a SSB system is equivalent to that of DSB-SC system.

8.2.3 Effect of Noise on Conventional AM

The received signal at the input to the demodulator is

$$r(t) = [A_c[+\mu m_n(t)] + n_c(t)]\cos 2\pi f_c t - n_s(t)\sin 2\pi f_c t \tag{8.2.16}$$

where $m_n(t)$ is normalized so that its minimum value is -1. The output of the lowpass filter after mixing is

$$y_1(t) = \frac{1}{2}\{A_c[1 + \mu m_n(t)] + n_c(t)\} \tag{8.2.17}$$

The desired component is $m(t)$, not $1+\mu m_n(t)$. The DC component in the demodulated waveform is removed by a dc blocking device and, hence, the lowpass filter output is

$$y(t) = \frac{1}{2}A_c \mu m_n(t) + \frac{n_c(t)}{2} \tag{8.2.18}$$

Hence, the received power is

$$P_R = \frac{A_c^2}{2}\left[1+\mu^2 P_{m_n}\right] \tag{8.2.19}$$

where we have assumed that the message signal is zero mean.

Now, the output SNR for coherent demodulator is

$$\begin{aligned}
\left(\frac{S}{N}\right)_{oAM} &= \frac{\frac{1}{4}A_c^2\mu^2 P_{m_n}}{\frac{1}{4}P_{n_c}} \\
&= \frac{A_c^2\mu^2 P_{m_n}}{2N_oW} \\
&= \frac{\mu^2 P_{m_n}}{1+\mu^2 P_{m_n}} \frac{\frac{A_c^2}{2}\left[1+\mu^2 P_{m_n}\right]}{N_oW} \\
&= \frac{\mu^2 P_{m_n}}{1+\mu^2 P_{m_n}} \frac{P_R}{N_oW} \\
&= \frac{\mu^2 P_{m_n}}{1+\mu^2 P_{m_n}} \left(\frac{S}{N}\right)_b \\
&= \eta \left(\frac{S}{N}\right)_b
\end{aligned} \tag{8.2.20}$$

Note that $\left(\frac{S}{N}\right)_b = \frac{P_R}{N_oW}$ From *Eqn.* 8.2.20, since $\mu^2 P_{m_n} < 1+\mu^2 P_{m_n}$, the SNR in conventional AM is always smaller than the SNR in a baseband system.

The power content of the normalized message process depends on the message source. For speech signals that usually have a large dynamic range, P_{m_n} is in the neighborhood of 0.1. This means that the overall loss in SNR compared to a baseband system is a factor of 0.075 or equivalent to a loss of 11 dB. The reason for this loss is that a large part of the transmitter power is used to send the carrier component of the modulated signal and not the desired signal.

8.3 Noise on Angle Modulation

Figure 8.4 is the effect of additive noise on zero crossings of two FM signals, one with high power and the other with low power. From the previous discussion on noise on AM signals and *Figure* 8.4, it should be clear that the effect of noise in an FM system is different from that for an AM system. IT can also be observed that the effect of noise in a low-power FM system is more severe than in a high-power FM system. In a low power signal, noise causes more changes in the zero crossings.

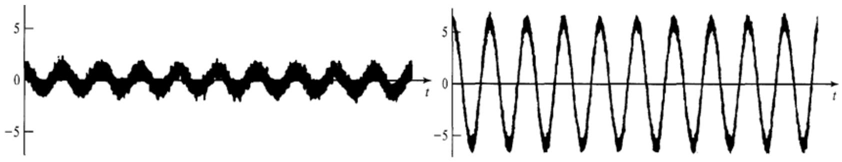


Figure 8.4: Effect of noise on FM

The receiver for a general angle-modulated signal is shown in Figure 8.5

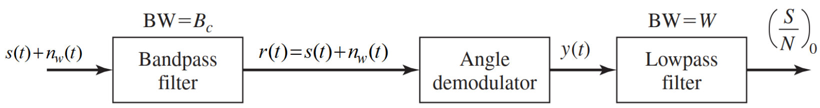


Figure 8.5: Block diagram of receiver for a general angle-demodulated signal

The angle-modulated signal is represented as

$$s(t) = A_c \cos(2\pi f_c t + \phi(t)) = \begin{cases} A_c \cos\left(2\pi f_c t + 2\pi k_f \int_{-\infty}^{\infty} m(\tau) d\tau\right), & \text{FM} \\ A_c \cos(2\pi f_c t + k_p m(t)), & \text{PM} \end{cases} \tag{8.3.1}$$

The AWGN $n_w(t)$ is added to $s(t)$, and the result is passed through a noise-limiting filter whose role is to remove the out-of-band noise. The bandwidth of this filter is equal to that of the modulated signal, therefore, it passes the modulated signal without distortion.

However, it eliminates the out-of-band noise. Hence, the noise output of the filter is a filtered noise denoted by $n(t)$.
The output of the filter is

$$r(t) = s(t) + n(t) = s(t) + n_c(t)\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t) \tag{8.3.2}$$

In effect, angle modulated signal performs better in the presence of noise than linear modulated signal.

Example 8.43

Design an FM system that achieves an SNR at the receiver equal to 40 dB and requires the minimum amount of transmitter power. The bandwidth of the channel is 120 KHz, the message bandwidth is 10 KHz, the average-to-peak-power ratio for the message, $P_{m_n} = \frac{P_m}{(\max|m(t)|)^2}$ is 1/2, and the (one-sided) noise power-spectral density is $N_o = 10^{-8}$ W/Hz.
What is the required transmitter power if the signal is attenuated by 40 dB in transmission through the channel?

Solution
Using Carson's rule it can be checked whether the bandwidth or the threshold impose a more restrictive bound on the modulation index.

$$B_T = 2(\beta + 1)W$$
$$120000 = 2(\beta + 1) \times 10000$$

from which we obtain $\beta = 5$.
By the relation

$$\left(\frac{S}{N}\right)_o = 60\beta^2(\beta+1)P_{m_n} \tag{8.3.3}$$

with $\left(\frac{S}{N}\right)_o = 10^4$, $\beta \approx 6.6$ is obtained. Since the value of β given by the bandwidth constraint is less than the value of β given by the power constraint, we are limited in bandwidth opposed to being limited in power. Therefore, we can choose $\beta = 5$ which, when substituted in the output SNR given by

$$\left(\frac{S}{N}\right)_o = \frac{3}{2}\beta^2\left(\frac{S}{N}\right)_b \tag{8.3.4}$$

Eqn. 8.3.4 yields

$$\left(\frac{S}{N}\right)_o = \frac{800}{3} = 266.6 \approx 24.26 \text{ dB}$$

Since $\left(\frac{S}{N}\right)_b = \dfrac{P_R}{N_o W}$ with $W10000$ and $N_o = 10^{-8}$, we obtain

$$P_R = \frac{8}{300} = 0.0266 \approx -15.74 \text{ dB}$$

and

$$P_T = -15.74 + 40 = 24.26 \text{ dB } \approx 266.66 \text{ W}$$

Had there been no bandwidth constraint, we could have chosen $\beta = 6.6$, which results in $\left(\frac{S}{N}\right)_b \approx 153$. In turn, we have $P_R \approx 0.0153$ and $P_T = \approx 153$ W.

8.4 Review Questions

1. In commercial FM broadcasting, $W = 15$ kHz, $f_o = 2100$ Hz, and $\beta = 5$. Assuming that the average-to-peak-power ratio of the message signal is 0.5, find the improvement in the output SNR of FM with pre-emphasis and de-emphasis filtering compared to a baseband system.

2. A signal transmitted through a 10 km coaxial line channel which exhibits a loss of 2 dB/km. The transmitted signal power is $P_{T\text{dB}} = -30$ dBW (-30 dBW means 30 dB below 1 W or simply, 1 mW). Determine the received signal power and the power at the output of an amplifier which has a gain of $G_{\text{dB}} = 15$ dB.

3. The received signal $r(t) = s(t) + n(t)$ in a communication system is passed through an ideal LPF with bandwidth W and unity gain. The signal component $s(t)$ has power-spectral density

$$S_s(f) = \frac{P_o}{1 + (f/B)^2}$$

where B is the 3-dB bandwidth.
The noise component $n(t)$ has a PSD $N_o/2$ for all frequencies. Determine and plot the SNR as a function of ratio W/B. What is the filter bandwidth W that yields a maximum SNR?

4. A certain communication channel is characterized by 90-dB attenuation and additive white noise with power-spectral density of $\frac{N_o}{2} = 0.5 \times 10^{-14}$ W/Hz. The bandwidth of the message signal is 1.5 MHz and its amplitude is uniformly distributed in the interval $[-1,\ 1]$. If we require that the SNR after demodulation be 30 dB, in each of the following cases find the necessary transmitter power.

   1. USSB modulation.
   2. Conventional AM with a modulation index of 0.5.
   3. DSB-SC modulation

5. In an analog communication system, demodulation gain is defined as the ratio of the SNR at the output of the demodulator to the SNR at the output of the noise- limiting filter at the receiver front end. Find expressions for the demodulation gain in each of the following cases:

   (a) DSB
   (b) SSB
   (c) Conventional AM with a modulation index of μ. What is the largest possible demodulation gain in this case?
   (d) FM with modulation index β_f
   (e) PM with modulation index β_p

6. is applied to a low-pass RC filter. The amplitude A_c and frequency f_c of the sinusoidal component are constants, and $w(t)$ is white noise of zero mean and power spectral density $\frac{N_o}{2}$. Find an expression for the output signal-to-noise ratio with the sinusoidal component of $x(t)$ regarded as the signal of interest.

7. A 10 kilowatt transmitter amplitude modulates a carrier with a tone $m(t) = \sin(2000\pi t)$, using 50 percent modulation. Propagation losses between the transmitter and the receiver attenuate the signal by 90 dB. The receiver has a front-end noise with spectral density $N - o = -113$ W/Hz and includes a bandpass filter with bandwidth $B_T = 2$ W $= 10$ kHz. What is the post-detection signal-to-noise ratio, assuming that the receiver uses an envelope detector?

8. The average noise power per unit bandwidth measured at the front end of an AM receiver is 10^-6 W/Hz. The modulating signal is sinusoidal, with a carrier power of 80 watts and a sideband power of 10 watts per sideband. The message bandwidth is 4 kHz. Assuming the use of an envelope detector in the receiver, determine the output signal-to-noise ratio of the system. By how many decibels is this system inferior to a DSB-SC modulation system?

8.5 Books and References

[1] Principles of Communication Systems by Taub and Schilling,2nd Edition. McGraw Hill.

[2] Communication Systems by Siman Haykin,4th Edition, John Wiley and Sons Inc.

[3] Modern Digital and Analog Communication System, by B. P. Lathi, 4th Edition, Oxford University Press.

[4] Principles of Communication Systems by Taub and Schilling,2nd Edition. McGraw Hill.

[5] Digital and analog communication systems, by L. W. Couch, 6th Edition, Pearson Education, Pvt. Ltd

[6] Analog Communications, First Edition 2009, A. P. Godse and U. A. Bakshi

[7] Analog Communication(Jntu) By Thomas-Chandrasekhar

[8] Communication Systems Engineering, Second Edition by John G. Proakis and Masoud Salehi

[9] Introduction to Analog and Digital Communications, Second Edition by Simon Haykin and Michael Moher

[10] Signals and Systems by Alan V. Oppenheim

[11] A Course in Electrical Technology (Electronic Devices and Circuits) Vol. III, Chapter21, 3rd Edition by J. B. Gupta

[12] Principles of Communication Lecture Notes, Indian Institutes of Technology, Madras by Prof V. Vankata Rao

[13] Electronic communications, 4th ed. by Dennis Roddy and John Coolen

Appendix A

Basic Communications Terminologies

***

If it must be done, it must be done well

-A Presbyterian Theme

A Signal:- is a function that specifies how a specific variable changes versus an independent variable such as time, location, height (examples: the age of people versus their coordinates on Earth, the amount of money in your bank account versus time).

Message:- the physical manifestation of information as produced by the source. Whatever form the message takes, the goal of a communication system is to reproduce at the destination an acceptable replica of the source message. There are two distinct message catergories, *analog* and *digital*

Distortion:- is waveform perturbation caused by imperfect response of the system to the desired signal itself. Unlike noise and interference, distortion disappears when the signal is turned off. If the channel has a linear but distorting response, then distortion may be corrected, or at least reduced, with the help of special filters called equalizers.

Interference:- is contamination by extraneous signals from human sources—other transmitters, power lines and machinery, switching circuits, and so on. Interference occurs most often in radio systems whose receiving antennas usually intercept several signals at the same time. *Radio-frequency interference* (RFI) also appears in cable systems if the transmission wires or receiver circuitry pick up signals radiated from nearby sources.

Noise:- refers to random and unpredictable electrical signals produced by natural processes both internal and external to the system. When such random variations are superimposed on an information-bearing signal, the message may be partially corrupted or totally obliterated.

Filtering reduces noise contamination, but there inevitably remains some amount of noise that cannot be eliminated. It is an undesired signal that gets added to (or sometimes multiplied with) a desired transmitted signal at the receiver. The source of noise may be external to the communication system (noise resulting from electric machines, other communication systems, and noise from outer space) or internal to the communication system (noise resulting from the collision of electrons with atoms in wires and ICs).

Signal to Noise Ratio (SNR):-is the ratio of the power of the desired signal to the power of the noise signal. That's noise is measured in relation to an information signal

Bandwidth (BW):- is the width of the frequency range that the signal occupies. For example the bandwidth of a radio channel in the AM is around 10 kHz and the bandwidth of a radio channel in the FM band is 150 kHz. The concept of bandwidth applies to both signals and systems as a measure of speed. When a signal changes rapidly with time, its frequency content, or spectrum, extends over a wide range, and we say that the signal has a *large* bandwidth. Similarly, the ability of a system to follow signal variations is reflected in its usable *frequency response*, or *transmission bandwidth*. Now all electrical systems contain energy-storage elements, and stored energy cannot be changed instantaneously. Consequently, every communication system has a finite bandwidth B that limits the rate of signal variations. Communication under real-time conditions requires sufficient transmission band- width to accommodate the signal spectrum; otherwise, severe distortion will result.

A System:- operates on an input signal in a predefined way to generate an output signal.

Analog Signals:- are signals with amplitudes that may take any real value out of an infinite number of values in a specific range (examples: the height of mercury in a 10cm–long thermometer over a period of time is a function of time that may take any value between 0 and 10cm, the weight of people setting in a class room is a function of space (x and y coordinates) that may take any real value between 30 kg to 200 kg (typically)).

Digital Signals:- are signals with amplitudes that may take only a specific number of values (number of possible values is less than infinite) (examples: the number of days in a year versus the year is a function that takes one of two values of 365 or 366 days, number of people sitting on a one-person chair at any instant of time is either 0 or 1, the number of students registered in different classes at DCEE is an integer number between 1 and 100).

Rate of Communication:- is the speed at which DIGITAL information is transmitted. The maximum rate at which most of today's modems receive digital information is around 56 k bits/second and transmit digital information is around 33 k bits/second. A Local Area Network (LAN) can theoretically receive/transmit information at a rate of 100 M bits/s. Gigabit networks would be able to receive/transmit information at least 10 times that rate.

Modulation:- is changing one or more of the characteristics of a signal (known as the carrier signal) based on the value of another signal (known as the information or modulating signal) to produce a modulated signal.

Symbol (digital message):- A symbol is a group of k bits considered as a unit. This unit is referred to as a message symbol m_i $(i = 1,...,M)$ from finite symbol set or alphabet.

The size of the alphabet, M, is $M = 2^k$, where k is the number of bits in the symbol. For baseband transmission, each m_i, symbol is represented by one of a set of baseband pulse waveforms $g_1(t),\ g_2(t),...,g_M(t)$. When transmitting a sequence of such pulses, the unit *Baud* is sometimes used to express pulse rate (symbol rate). For typical band-pass transmission, each $g_i(t)$ pulse is represented by one of a set of bandpass waveforms $s_1(t),\ s_2(t),...,s_M(t)$. Thus, for wireless systems, the symbol m, is sent by transmitting the digital waveform $s_i(t)$ for T seconds, the symbol-time duration. The next symbol is sent during the next time interval, T.

Data rate:- This quantity in bits per second (bits/sec) is given by $R = k/T = (1/T)log_2M$ bits/sec, where k bits identify a symbol from an $M = 2^k$ symbol alphabet, and T is the k-bit symbol duration.

Appendix B

Some Selected Past Questions

***

B.1 December, 2015

Section B – Answer only TWO (2) questions from this section. Each question carries 10 marks

---

42. a) The message signal $m(t)$ has a bandwidth of 10 kHz, a power of 16 W and a maximum amplitude of 6. It is desirable to transmit this message to a destination via a channel with 80 dB attenuation and additive white noise with power-spectral density $S_n(f) = N_o/2 = 10^{-12}$ W/Hz and achieve a SNR at the modulator output of at least 50 dB. What is the required transmitter power and channel bandwidth if the following modulation schemes are employed?
    i. DSB AM
    ii. SSB AM
    iii. Conventional AM with modulation index equal to 0.8

    b) State three advantages DSB-SC systems have over conventional AM systems. Support your answer with relevant equations if necessary.

43. a) A signal $m(t)$ bandlimited to 5 kHz is DSB modulated by $cos(2\pi f_c t + \phi_c)$ to yield $s(t)$. For what value of carrier frequency f_c, will the bandwidth of $s(t)$ be 2% of f_c.

    b) Name three key ways that a higher-frequency signal called the carrier can be varied to transmit intelligence.

    c) With the help of a diagram describe the operations of the envelope detector.

44. a) An angle-modulated signal with carrier frequency $\omega_c = 2\pi \times 10^5$is described by the

$$\mathscr{S}(t) = 10\cos(\omega_c t + 5\sin 3000t + 10\sin 2000\pi t)$$

    i. Find the power of the modulated signal
    ii. Find the frequency deviation Δf

iii. Find the deviation ratio β
iv. Find the phase deviation $\Delta\phi$
v. Estimate the bandwidth of the $\mathscr{S}(t)$

b) State four advantages FM systems have over AM systems

45. a. **[Top-up]** Find the correlation coefficient and the energy of the signal pair as described below

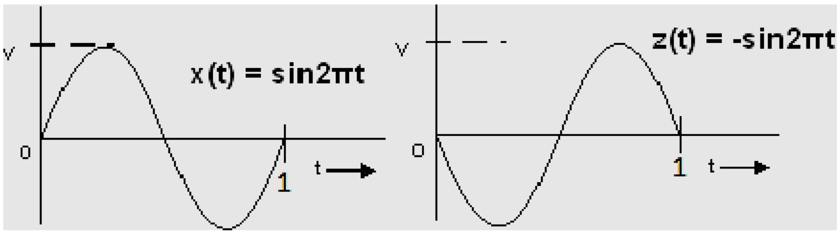


Figure B.1: Question 45

b. **[Top-up]** Compute the Fourier transform of $x(t) = Asin(\omega t)$

c. **[Top-up]** State Nyquist criterion based on the frequency and sampling interval.

46. a. **[Top-up]** Draw a block diagram of a typical communication system and label each unit.

b. **[Top-up]** Determine η and the percentage of the total power carried by the sidebands of an AM wave for tone modulation when

(i) $\mu = 0.5$ and
(ii) $\mu = 0.3$

c. **[Top-up]** Assume a carrier signal is given by $c(t) = 10cos(\omega_c t)$ and the message signal as $cos(\omega_m t)$. Find the carrier power.

B.2 December, 2016

Section B - Answer question 37 and any other TWO (2) questions in this section. ***Some selected questions.***

---

37. (a) A single-toned message of 25 kHz is frequency modulated onto a carrier of 75 kHz. The peak frequency deviation is set to be 10 kHz. What is the modulated index of the FM signal thus obtained? **[1 mark]**

    (b) Write the time-domain expression for the FM signal in (a). **[2 marks]**

    (c) What is the peak phase deviation? **[1 mark]**

    (d) What is the approximate bandwidth of the FM signal? **[1 mark]**

    (e) Write an expression for the spectrum of the FM signal in (a). **[2 marks]**

    (f) Sketch the magnitude spectrum of the FM signal in (a) for positive frequencies based on (e) using Bessel function table of values. **[2 marks]**

38. a) Briefly describe two methods that can be used to generate Wideband FM. **[2 marks]**

    b) Explain with appropriate spectrum diagram(s) how VSB is generated from DSB-SC waveform. **[3 marks]**

    c) A DSB-SC signal, $s(t) = A_c m(t) cos(\omega_c t)$, with a message bandwidth of W, is transmitted over AWGN channel with two-sided noise power spectral density of $N_o/2$:

       i. What is AWGN channel? **[1 mark]**

       ii. Write an expression for the transmitted power. **[2 marks]**

       iii. Derive an expression for the signal-to-noise ratio (SNR) on the channel. **[2 marks]**

39. a) An AM signal is represented by the equation below and has a total output power of 15 W.

$$v(t) = [25 + 10\sin(12566t)]\sin\left(1.2566 \times 10^6 t\right)$$

       i. What are the values of the carrier and modulating frequencies? **[1 mark]**

       ii. What are the amplitudes of the carrier and of the sideband frequency? **[1 mark]**

       iii. What is the modulation index of the AM signal? **[1 mark]**

       iv. What is the bandwidth of the signal? **[1 mark]**

    b) What digital modulation scheme is closely related to amplitude modulation? Explain your answer. **[2 marks]**

    c) An analog signal of bandwidth 5 kHz is to be sampled and encoded as a 6-bit PCM signal.

       i. What is the minimum sampling rate for the PCM signal? **[1 mark]**

       ii. If the signal is limited to the voltage range -2.5 to 2.5 V, what is the smallest quantizer step size? **[1 mark]**

       iii. What is the signal-to-quantization noise in decibel? **[1 mark]**

       iv. What is the bit-rate of the PCM signal? **[1 mark]**

B.3 February, 2017

Instruction: Answer any THREE (3) questions. ***Some selected questions.***

Question Three [Top-up]

a. Given a carrier waveform, $c(t) = A_c \cos(\omega_c t)$ and a modulating signal of $m(t) = A_m \cos(\omega_m t)$, explain briefly with the help of a diagram, the generation of basic DSB-SC waveform. **[4 marks]**

b. If the response of the diode in a Switching Modulator of DSB-FC (Ordinary AM) is given by

$$p(t) = \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} \cos\left[2\pi f_c t\,(2n-1)\right]$$

i. Obtain the first three terms of the diode response. **[4 marks]**

ii. Given that, the input of the modulator is $v_m(t) = m(t) + A_c \cos 2\pi f_c t$ and the output is $v_o(t) = v_m(t) \cdot p(t)$, show that after bandpass filtering using BPF centred at f_c of the output, $v_o(t)$ is given by **[10 marks]**

$$s(t) = \frac{A_c}{2} \cos\left(2\pi f_c t\right) \left[a + \frac{4}{\pi} \frac{m(t)}{A_c}\right]$$

iii. If $m(t) = A_m \cos(\omega_m t)$, what is the modulation index of $s(t)$ in (ii)? **[2 marks]**

Question Four [Top-up]
An FM signal has a deviation of 3 kHz and a modulating frequency of 1 kHz. Its total power is 5 W, developed across a 50 Ω resistive load. The carrier frequency is 160 MHz.

i. Calculate the RMS value of the total signal voltage **[1.5 marks]**

ii. Calculate the RMS voltage at the carrier frequency and each of the first three sets of sidebands. **[4 marks]**

iii. Calculate the frequency of each sideband for the first three sideband pairs. **[4 marks]**

iv. Calculate the power at the carrier frequency, and in each sideband, for the first three pairs. **[3 marks]**

v. Determine what percentage of the total signal power is unaccounted for by the components described in (iv). **[2 marks]**

vi. Sketch on a graph sheet the signal in the frequency domain, as it would appear on a spectrum analyzer. The vertical scale should be power in dBm, and the horizontal scale should be frequency. **[4 marks]**

vii. Find the bandwidth of the FM signal. **[1.5 marks]**