TELECOMMUNICATION ENGINEERING TECHNOLOGY

(BOOK OF BASIC LESSONS)

ISBN: 978-978-767-138-2

AFROPOLITAN PUBLICATIONS INT'L

KENNETH UGO UDEZE

Telecommunication Engineering Technology

(Book of Basic Lessons)

By

Kenneth Ugo Udeze

Telecommunication Engineering Technology (Book of Basic Lessons)

Published by Afropolitan Publications Int'l

200132, Institute of African Studies,
University of Ibadan, Nigeria.
https://books.afropolitanjournals.com/index.php/pub

The online version of this publication is at permanently available Afropolitan Books, The e-Book platform of Afropolitan Publications, https://books.afropolitanjournals.com/index.php/pub (Open Access).

ISBN: 978-978-767-138-2

CIP Data:
NATIONAL LIBRARY OF NIGERIA CATALOGUING-IN-PUBLICATION DATA
Telecommunication Engineering Technology: Book of Basic Lessons
Udeze, Kenneth Ugo, 1975-
1. Telecommunication—Study and teaching
I. Title

TK5101.A3 U19	2024	621.384
ISBN: 978-978-767-138-2	(pbk)	AACR2

Table of Contents

Preface

This book originates from notes used in teaching Digital Communication and Telecommunication courses in Electrical and Electronics Engineering Department, Federal Polytechnic, Oko, Anambra State, Nigeria. Along with other materials gathered by the author during his degree and post-degree years of academic pursuit, and over fifteen (15) years of teaching experience in accordance with course curriculum guidelines from the National Board for Technical Education (NBTE), this text, "**TELECOMMUNICATION**", was written.

The content of each chapter was designed to accommodate Higher National Diploma (HND) and Bachelor of Science/Engineering (B.Sc./B.Eng.) undergraduate students as the materials presented were made comprehensive enough to cover both classes of programs at their mid-course levels.

Chapters 1 and 2 cover the basic Background of Telecommunication Systems and Signal/Modulation.
Chapters 3 and 4 discuss Amplitude Modulation Theory and Frequency Modulation Theory.
Chapter 5 covers Phase Modulation Theory which is similar to the Modulation Theory covered in chapter 4.
Chapters 6 covers Different Types of Signals, Signals Systems, Fourier Series and Convolution Theory.
Chapter 7 and 8 cover Sampling Theory and Digital Modulation, respectively.
Chapter 9 and 10 discuss, in detail, Multiplexing and Detectors.

At the end of the chapters are enough review problems designed to help the students exercise their level of comprehension of the treated matters, and by so doing internalize the underlying principles of the lessons taught.

About the Author

Kenneth Ugo Udeze hails from Onicha Ugbo in Aniocha North Local Government Area in Delta State, Nigeria. He attended his primary school at Aniemeke Primary School Onicha Ugbo, Delta State, Nigeria and attended his secondary education at Model Secondary School Maitama, Abuja, FCT, Nigeria where he obtained his Senior School Certificate in 2003.

Between 2005 and 2010, he obtained his National Diploma (ND) and Higher National Diploma (HND) in Electrical and Electronics Engineering (Telecommunication and Electronics Options) with a CGPA of 3.68/4 i.e., Distinction Honors from Federal Polytechnic Oko, Anambra State, Nigeria. He also obtained his first degree in Electrical and Electronics Engineering (Power, Telecommunication and Electronics Options), in 2013 from the prestigious University of Ibadan, Oyo State, Nigeria with a CGPA of 5.8/7 i.e., a Second-Class Upper Division (2.1).

After his one year mandatory National Youth Service Corp (NYSC) in Electrical and Electronics Engineering Department in Federal Polytechnic Oko, Anambra State, Nigeria in 2015, he proceeded to obtain his Masters degree in Offshore Engineering in 2016, majored in Offshore Design and installations, Subsea umbilical cables designed, Installation and maintenance of offshore facilities, Submarine power cable design and maintenance, Subsea instrumentation and control system (E&I) from Offshore Technology Institute, School of Advance Engineering, University of Port-Harcourt, Rivers State, Nigeria. Then a second Masters degree in Electrical and Electronics Engineering and majored in Power System Engineering, from University of Lagos, Lagos State, Nigeria in 2023. He graduated with a CGPA of 4.7/5 i.e., Distinction Honors.

He is currently a staff of Federal Polytechnic Oko, Anambra State, Nigeria attached to Electrical and Electronics Engineering Department. He teaches Mathematics and Electrical Engineering courses.

He is presently prospecting for PhD admission overseas for researches in Renewable Energy.

CHAPTER 1

HISTORY OF COMMUNICATION

1.0 In the Beginning

One paramount characteristic which distinguishes man from other animals is the ability to communicate with his fellow men at a very high level of complexity and speed. David Attenborough entitled the final programme of his *Life on Earth* TV series “Man the compulsive communicator” – and that is by no means an exaggeration.

In animals, communication is nevertheless an essential part of behaviour. Without it enemies and rivals could not be warned; herds could not show group behaviour; mating could not occur. However, most of it is at a very instinctive level (e.g. scent marking of territory, plumage displays in matting, bird song, rump marking in deer, etc.); even the limited range of voiced signals in higher mammals comes into this category.

By contrast man’s survival and dominance has depended on the ability to recognize individuals and to work together when hunting; to discuss complex ideas and actions; and, above all, to pass on information once gained. These methods became possible once mankind had developed a large versatile brain enabling speech and, eventually, writing as means of instant communication and permanent memory, respectively.

Both these techniques have the three main parts which any communication system must have to be useful. These parts are shown in Fig. 1.1.

Figure 1.1. Basic parts of a communication system

This is the simplest communication system man uses, yet it has all the three main elements quite obvious to see.

There is a **sending end –** the person speaking converts thoughts into muscular movements which operate a transmitter (the vocal cords and voicing system), converting them into pressure variations in the **transmission medium** (the air). These pressure variations travel outwards as a signal through the transmission channel (more air) to reach the **receiving end** (the listener), where a receiver (the ear drum, etc.) converts the pressure variations back into first movement, then electrical signals, and finally thoughts.

Thus, the listener had planted into his or her brain a replica of the speaker’s **message** and **information** has been conveyed.

Some other aspects of communication systems are also illustrated by this trivial example. They are as follows:

1. **Coding**

The speaker's message is conveyed in a language the receiver can understand. If it were in, say, Swahili a few of us could decode the message and so would receive little information. Morse code is a simple engineering example.

2. **Noise**

Unwanted random signals added to the original message may mean the listener has difficulty decoding it or it can equally be seen as unwanted random signals added to the original information signal. It corrupts the original message. Noise can occur in the channel or any part of the block. Try talking across a crowded room at a party and you will get the point. Noise can occur in all three parts of the system. At the party, it came in the transmission channel, but it could be generated by the sender (wheezy breathing) or by the receiver (tinnitus).

3. **Distortion**

Here the signal is changed but nothing is added. This change may be so great that the receiver cannot decode the message correctly. This is an unwanted change in the information signal. Try talking across a large empty hall. Some frequencies resonate. Reverberation occurs, introducing time delays and attenuation which cause the instantaneous spectrum of the signal to become so distorted as to become unintelligible. Distortion usually occurs in the transmission channel, although it may occur elsewhere (e.g. if the speaker has a strong accent). Now, let us concentrate on the sort of system this course is all about.

1.1 Communication at a Distance

We must now restrict ourselves to artificial methods of communication which enable us to send information over longer distances than can be spanned by the human voice. At our noisy party, hand signals or lips reading might help for simple ideas. Carrying memorized or written messages has always worked and, for centuries, the greatest speed with which a message could be conveyed was that of the galloping horse (as in famous events like Paul Revere, Ghent-to-Aix and the Pony Express but usually it was much slower (as at Thermopylae).

Even then there was one way of transmitting messages much more quickly by sight. For centuries these ways were very simple indeed (e.g. bonfires heralding the approach of the Armada). Native smoke signals were at one stage better, as were flags, but the most advanced techniques were those of the heliograph and the semaphore. Actual words could be conveyed, although incredibly slowly by modern standards. The need for sending information quickly over long distances was so great that in UK for instance, governments went to the lengths of building large wooden semaphore towers on hilltops along important routes. A well-known English example occurred during the Napoleonic wars when the Admiralty built such a chain from London to Portsmouth. Their sites are often still called Telegraph Hill.

Not until electrical signals were used in the nineteenth century can we say that we were getting anywhere near the type of signal with which this course deals. The spur to it all was the development of the railways which this course deals. The spur to it all was the development of the railways which needed to let their operators know, beforehand, that a train was coming down the tract. They were able to do this by an electric telegraph which used pulses of current along copper wire with a variety of electromechanical indicators at the terminals. Transmission was virtually instantaneous but only one letter or number could be conveyed at a time, making all but the shortest messages as slow as by semaphore. Of course greater distances could be covered and suitable coding (e.g. Morse) speeded things up.

But the telephone, invented in 1869, marked the real breakthrough. It must have seemed miraculous at the time. You could now talk to someone miles away as easily and as quickly as if you were standing next to them (subject, of course, to noise, distortion, breakdown, etc.). by using repeaters, exchanges and multiplexing, vast complex telephone networks were built up culminating in long submarine cable links. Much of the content of this course applies to these wired systems as well as to the next development.

Wireless communication completed the breakthrough. It all started in 1895 when Hertz carried out some experiments using spark gaps in resonant loops. He transmitted near-microwave frequencies. Many workers (e.g. Lodge, Popov, Kelvin) tried to develop the use of these new electromagnetic waves to provide telecommunications, but eventual success went to Marconi who was initially supported by the British Admiralty in its race to communicate with its fleets at sea. He founded the Marconi Wireless Telegraphy Company Ltd and, by sheer hard work, developed an enormous spark transmitter which, on 12 December 1901, was able to signal across the Atlantic using very long wavelengths - and without any amplification at all.

After this developments occurred in two main areas – hardware and software. The main hardware changes have been the introduction of amplifiers (vacuum valves for a long time, then semiconductors); the use of higher frequencies; improvements in aerial design (and an understanding of propagation mechanisms); optical fibres and cables; the use of satellites. Software changes have been mainly concerned with developing methods of modulation and coding to reduce bandwidth and power and to improve range, speed, and reliability. That is what this book is about.

1.2 Elements of Communication

Tele means distance. Telecommunication is the process of passing information energy over long distance by electrical means. The information energy is passed to the destination either over suitable insulated conducting wires called transmission lines, or through the atmosphere without the use of wires by a radio link. Consider these:

Telecommunication	communication at a distance
Telephone	speaking at a distance
Television	seeing at a distance
Telegraph	writing at a distance

A communication system is a combination of circuits and devices put together to accomplish the task of transmission of information from one point to another.

The source and the destination: The origin of the information or message is referred to as the source and the destination as the sink. The source and the sink can be either be man or machine. The original form of the information may be sound, light pattern (picture), temperature etc. The information is first converted into electrical form to produce electronic information signal. This is achieved by a suitable transducer or encoder.

Channel: This is the medium through which the information passes. The channel may be transmission lines, like in telephone systems, or radio link (space). In the channel, the signal undergoes degradation although this may occur at any point in the communication system block, it is customarily associated with the channel alone. This degradation often results from noise and other undesired signals or interferences, but may also include distortion effects due to fading signal level, multiple transmission paths and filtering.

Receiver: The receiver's function is to extract the desired message from the received signal at the channel output and to convert it to a form suitable to the output transducer. The main functions are amplification, filtering and demodulation.

Output transducer: The output transducer completes the communication system. This device converts the electric signal at its output circuit into forms desired by the system user. Examples of output devices are loudspeakers, tape recorders, PCs, meters and CRT, etc.

Communication (Information) signals: An examples of an information signal is the voltage waveform produced by a microphone in response to a spoken message. The voltage varies continuously with time in an unpredictable manner and it Is the task of the communication link to product a close replica of the waveform at the receiver.

The signal is passed along a pair of wires. For communication over long distances, some modification of the signal is required to make it compatible with an available channel such as a radio link. This modification is referred to as ENCODING and/or MODULATION and a reverse process is required at the receiver to undo the modification and recover the original message waveform.

The communication signal has a bandwidth which is the range of significant frequency component in the signal.

Signal types	Frequency ranges
Speech	300 Hz to 34 kHz
Music	20 Hz to 20 kHz
Television (video)	0 Hz to 5.5 MHz

These signals are low pass. It may be necessary to transform a low pass signal to a band pass signal to enable the message to be conveyed over a band pass channel.

Perhaps a more complete general representation of a communication system should be as shown in Fig. 1.2.

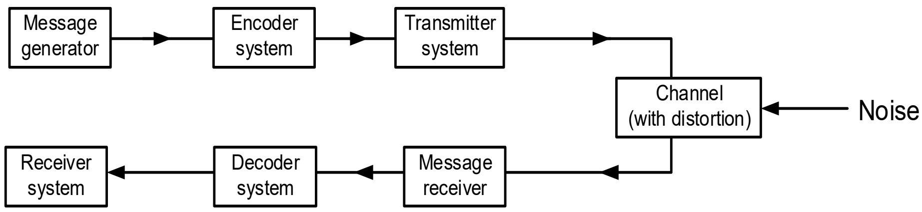


Figure 1.2. Parts of a communication system

Ideal communication channel

An ideal communication channel would convey unaffected or impaired communication signal from source to destination. It Is expected to pass all frequencies in the bandwidth equally, and also to be free from unwanted signals and pass the signal with minimum attenuation. However, communication channels can be divided into 2 broad classes namely:

1. **Guided wave systems**: In which the signal is conveyed via some constraining physical medium such as a pair of wires.
2. **Radio systems:** In which signal transfer is effected in a freely propagating electromagnetic wave. The range of frequencies available is very wide ranging from d.c. (0 Hz) to optical frequencies of the order 10^{14} Hz. A message signal can be translated from its bandwidth to another frequency for transmission and is recoverable at remote locations.

1.3 Telecommunications Systems Design Considerations

The basic parts of a communication system are very simple. The complexity comes when trying to create practical communication systems to do specific tasks to exact specifications. The following considerations have to be taken into account in designing a communication system with detailed considerations.

a. Range
b. Power
c. Cost
d. Bandwidth
e. Speed
f. Reliability
g. Convenience
h. Accuracy/Quality

a. Range

The longer the range, the more difficult it is to get the message through uncorrupted. Wire links work well at low frequency and require repeaters for longer distances. High frequencies require special cables-w axial cable, wave guides and optical fibres. Terrestial radio links require different frequencies for different purposes. Microwave for line of sight links, HF for long distance ionospheric communication across the world, VHF for a variety of shorter range users, UHF to give large bandwidth for TV, medium wave for local broadcasters.

Satellite links can be used for both small and large terrestrial distances but always have long path lengths with the consequent problems of attenuation and noise.

b. Power

The less the power required at the sending end, the simpler and cheaper is the transmitting installation required (but the receiver has to be more sophisticated). Where other considerations allow it, transmitted power kept to a minimum. Compare Marconi's megawatts of peak power with the milliwatts used by the early satellites.

Some of the factors involved are that higher frequencies produce a higher proportion of radiated power from aerials; radiated power may need to be kept high to allow the use of cheap receivers (as in radio broadcasting); pulse coding techniques increase signal accuracy and enable very weak signals to be recovered from noise (e.g. on space probes); directional aerials increase effective use of radiated power (e.g. on microwave links); multiplexing enables more information to be sent for the same power.

c. Cost

This obviously has kept as low as is compatible with achieving the desired system performance. What is economic could depend on the application. For example, it is worth while spending many millions of pounds on the ground installations for transatlantic communication satellite reception, but 10 s of thousands naira would be too much for a roof-top aerial for direct reception of educational TV from satellites in Nigeria. Cost is almost irrelevant for stringent military requirements, such as in the guidance systems of nuclear missiles, whereas it becomes the major factor for mass commercial systems such as CB radio. Much of the ingenuity in developing systems goes into doing same thing but cheaper (e.g. optical fibres not microwaves).

d. Bandwidth

No information will be obtained at all unless the signal received contains at least a small range of frequencies (its bandwidth) and unchanging single carrier frequency tells you very little, except that the transmitter is on. But using a larger bandwidth than is essential increases cost, and complexity unnecessarily. The narrower the individual channel, the more channels can be sent simultaneously over the same link bandwidth. Thus considerable ingenuity goes into reducing channel bandwidth whilst retaining acceptable quality of information. Often, a compromise is reached between the two.

For example, in a telephone channel the bandwidth needed is halved at the start by using sideband (SSB) techniques, and then making what is left just wide enough so that voices are recognizable. The result is the standard 4 kHz bandwidth (B/W) voice frequency channel for telephones. On the other hand, where virtual 100% accuracy is required (as in many data transmission requirements) digital techniques must be used. These require larger bandwidths than other modulation techniques but this is acceptable because much higher carrier frequencies can be used. This is one reason for the present shift to optical fibre transmission.

e. Speed

Real-time transmission is very common (e.g. telephones, TV). If you send information more slowly, you save bandwidth in expense of time. This may be acceptable and cheaper (e.g. teleprinters, facsimile) and so is used. On the hand, by sending information faster more bandwidth is required but less time is taken. This may be necessary (e.g. high-speed data) or even cheaper. Again, the modulation and encoding methods used can be designed to get the best out of the system – digital methods in particular.

f. Reliability

Reliability of the received message. The aim is to use the cheapest and simplest system which will give acceptable re-productivity of the signal which will reduce signal degradation to the minimum.

How much does it matter if your signal arrives corrupted? Most of this is considered above but it is worth summarizing separately here. Obviously reliability of equipment is a factor but not the one meant here. We mean factors causing signal degradation in a working system. The aim is to use the cheapest and simplest system which will give acceptable reproducibility of signal. A telephone channel is a good example of an uncomplicated system which has acceptable quality without achieving complete accuracy. Narrow bandwidths, low frequencies and intense multiplexing can be used.

Digital data streams obviously need much greater accuracy and have to have much larger bandwidths to get it. But even here a very low error rate (e.g. 1 in 10^5) may be achieved by the use of coding techniques. There always seems to be a trade-off between bandwidth and accuracy.

g. Convenience

This encompasses a multitude of factors of which the most restricting technically is the need for new systems to be compatible with older existing systems. A well-known example occured when colour was introduced into TV. It had both to keep the same bandwidth and to be compatible with existing black and white sets. It is now ocurring again with direct broadcast satellites. A more recent example has been the digitizing of the telephone system.

Other aspects of convenience occur with the growth of multi-facility networks (they must be digital); the use of larger and more comprehensive integrated circuits wherever possible (perhaps) fixing details of the modulation method); the need for ease of production and cheaper repair (modular design); and so on.

h. Accuracy/Quality

These considerations have been mentioned in most of the others above. The more accurate the received information signal must be compared with the original, the more complex and expensive the communication system has to be, in general. Thus, economics can be made because there is absolutely no need to recover the signal at the receiver more accurately than it needs to be to serve the purpose for which it was sent. Take speech signals as an example. Here, accuracy is a fairly subjective matter better described by the term quality and is something we automatically take into account every day when telephoning or listening to the radio knowing that what we hear is understandable and recognizable but not perfect. Qualitative considerations of acceptability of received sound have led to quantitative requirements for system specifications which differ for different purposes. The 4 kHz standard telephone channel bandwidth has been mentioned already, but in "steam" radio 3 kHz is regarded as enough with even less HF communications. On the other hand, a minimum of 15 kHz is considered necessary on VHF radio to reproduce music at acceptable quality. By contrast digital signals need much higher accuracy, partly because a single missed bit can cause a much more serious error than the loss of a short piece of analogue signal, but also because digital data is often of no use unless very accurate indeed.

Table 1.3. Communication frequency and ranges

Ranges	Definition	Frequency	Wavelength	Application
VLF	Very low frequency	<30 kHz	>10 km	long distance
LF	Low frequency	30 kHz-300 K Hz	10 km-1 km	telegraphy point-to-point service, navigation sound broadcasting
MF	Medium frequency	300 kHz-3 MHz	1 km-100 m	sound broadcasting, ship shore service
HF	High frequency	3 MHz-30 MHz	100 m-10 m	point-to-point, sound broadcasting
VHF	Very high frequency	30 MHz-300 MHz	10 m-1 m	T.V, radio (sound broadcasting)

UHF	Ultra High frequency	300 MHz-3 GHz	1 m-10 cm	radio, air-to-air and air-to-land
SHF	Super High frequency	3 GHz-30 GHz	10 cm-1 cm	series microwave communication radio
EHF	Extra High frequency	>30 GHz		
	Optical frequency	≥10^{19} Hz		

1.4 Conclusion

However interesting they may be, these generalized preliminaries are for now enough. Our aim in the rest of the book is to make a detailed examination of the basics of the various methods of modulation, their theories, and something about their implementations and uses.

CHAPTER 2

SIGNALS AND MODULATION

2.0 Introduction

There are four aims of this chapter. The first is to give an overview of the various types of baseband signal which can form part of the information to be sent down a communication system. The second is allied to this and looks at the baseband bandwidths involved and how they come about. The third discusses the need for modulation to enable these basebands to be sent, giving the various types of modulation used and showing how they relate to each other. The fourth looks at the general basis of the two classes of modulation technique.

2.1 Baseband Signals of Types

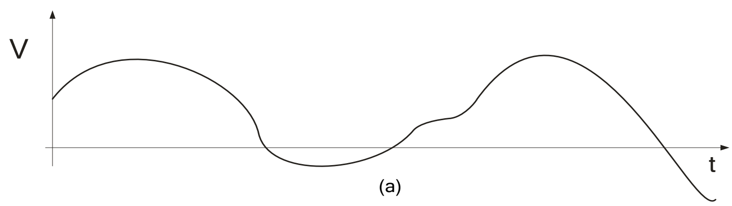
The information to be transmitted are considered signal only after they have been converted or become electrical signals which are either voltage or current. Many will start off in non-electrical form (e.g. voice, temperature) but will have been converted to voltage using some form of transducer (e.g. microphone, thermocouple etc). These electrical signals form the original information which is to be sent is therefore known as the baseband signals. Advancement of communication has been able to divide these into main two classes, namely; analogue signals, and digital signals.

1. **Analogue signals**

An analogue signal is a type of electrical signal that shows a continuous variation in time at a wide range of magnitudes. Often this time variation is the same as that of an original non-electrical signal (e.g. air pressure and voltage for a microphone) so that the two are said to be analogous; hence the name analogue signal.

2. **Digital signals**

In the digital signals, electrical signals consist of discontinuous pulses (or digits) each constant in value but changing abruptly from one digit (level) to the next. They are usually coded signals as in a teleprinter. The most common type is, of course, binary coding where only one type of pulse occurs but they can, in general, be multilevel (M-level) digits. Often a signal will start out analogue and be converted to digital (Fig. 2.1).



(a)

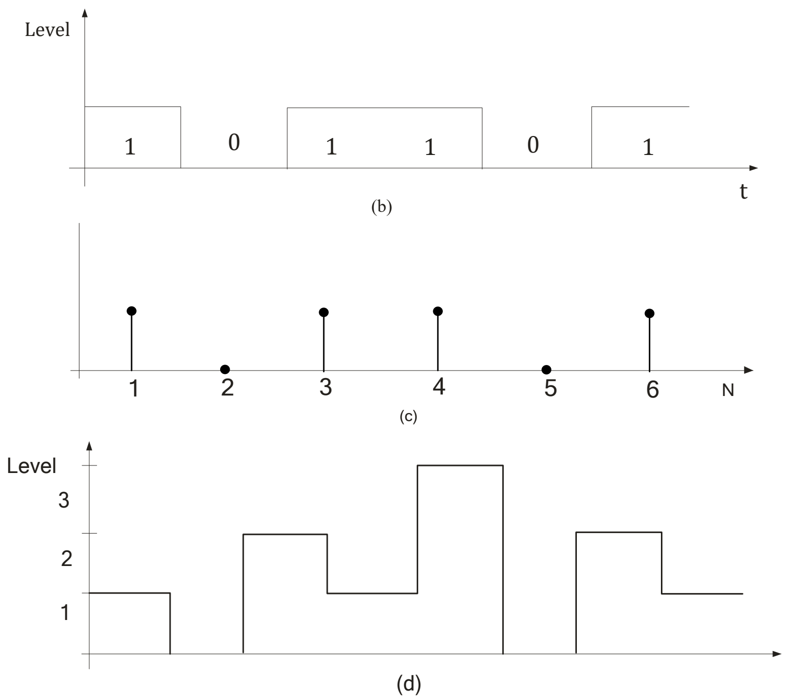


Figure 2.1. Classes of signal: (a) analogue (continuous); (b) analogue (discrete) or digital (binary); (c) functional representation of binary; (d) M-level digital (4-PAM: M=4=2^2)

Within these two broad classes, there are only a few distinct signal types, and the main ones are listed in Table 2.1 although the boundaries are somewhat blurred in applications.

For some types, the baseband bandwidths required depend greatly on the application (e.g. telemetry), but for others it is much more closely defined (e.g. television). The next section shows how some of these standard bandwidths are obtained.

2.2 Baseband and Bandwidth Terminology

The term **baseband** has both a general and a specific meaning. Generally, it is used as above, to refer to the original information signal. Specifically, it means the band (or range) of frequencies occupied by the baseband signal. The actual limits of this range need to be specified separately.

Table 2.1. Types of baseband signal

Signal Type	Nature	Class	Baseband Bandwidth	Usual Modulation type
Morse	Pulsed c.w.	Digital	0-50 Hz	ASK (OOK)
	Pulses from keyboard	Digital	0-120 Hz	FSK/PSK
Facsimile	Still copies	Digital	0-9.6 kHz	FSK/PSK
Telephone	Voice frequencies	Analogue	0-4 kHz	SSB/FDM
Audio	Music	Analogue	0-15 kHz	FM
Radio	AM (LF-HF)	Analogue	0-4.5 kHz	AM
Broadcast	FM (VHF)	Analogue	0-15 kHz	FM
Radio	Amateur	Analogue	0-3 kHz	SSB/NBFM
Radio	CB	Analogue	0-4 kHz	NBFM
Radio comms	Mobile	Analogue	0-3 kHz	AM/FM
	HF	Analogue	0-3 kHz	AM
PCM	Digitized audio	Digital	0-64 kHz	PSK
Telemetry	Data	Digital	To MHz	ASK/PSK
Television	Moving pictures	Analogue	0-6.5 MHz (UK)	VSB
			0-5.5 MHz (US)	VSB
Radar	Pulsed c.w	Digital	To GHz	ASK

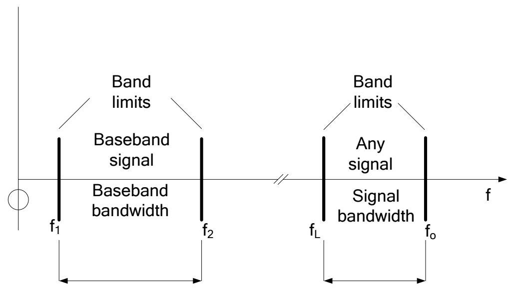


Figure 2.2. Bandwidths, bandwidths, baseband, and band limits

The term **bandwidth** refers merely to the frequency range within any band, without specifying the limits. It is a term which can be used at any frequency and not just at baseband (e.g. 9 kHz bandwidth for AM radio) and therefore, strictly speaking, we should use the term **baseband bandwidth** when specifically referring to baseband signals. Fig. 2.2 illustrates these terms.

The two terms are often used loosely as if they were synonymous, but there is an important distinction between them. For instance, a voice frequency channel for telephone always has a nominal bandwidth of 4 kHz, but the actual band it occupies at various stages of a telephone network may be very much higher, even megahertz. It is only at the beginning that it is a baseband.

2.3 Baseband Bandwidth Evaluation

Most baseband signals used in commercial communication systems are assigned to standard bandwidths fixed by international agreement. This is done by URSI (Union Radio Scientific International) through its two consultative bodies, CCITT (Comité Consultatif International Téléphonique et Télégraphique) and CCIR (Comite Consultatif International des Radiocommunications).

The reasons for the choices of bandwidth are partly subjective (i.e. how much needs to be sent to get the information across in acceptable form) and partly technical based on the way in which the information is sent. In this next section the reasons for some of the choices are discussed and values obtained for the relevant baseband bandwidths.

2.3.1 Telephony (Analogue)

Here the reasons for choice of bandwidth are mostly subjective. Human hearing can extend from about 20 to 20,000 Hz, but speech is quite clear and individual voices quite recognizable, whether pitched high or low, if only the middle frequencies are transmitted. The exact limits vary a little from country to country (300-3400 Hz in UK) but the addition of guard bands to allow for filtering and multiplexing means that we can have international agreement on a standard voice frequency channel (Fig. 2.3) of 0-4 kHz – standard throughout the world.

2.3.2 Telegraphy (Digital) (i.e. teleprinters)

Here the criteria used are only partly subjective. Operators can type fast but only with continuously varying symbol rates, so, to standardize matters, buffer amplifiers are used to allow data to be transmitted at one of several standard baud rates (e.g. 50 or 100 symbols per second) into the teleprinter. Below we shall look at 50 baud transmission but expressed as an information rate of 66 words per minute.

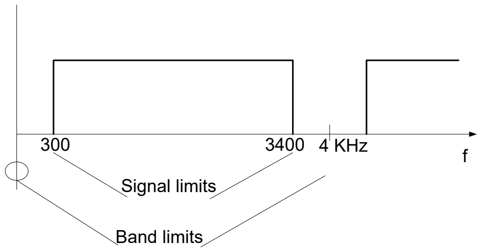


Figure 2.3. Standard voice frequency (telephone) channel

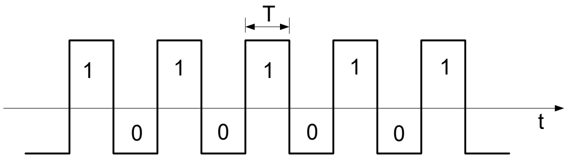


Figure 2.4. "Worst case" bit stream

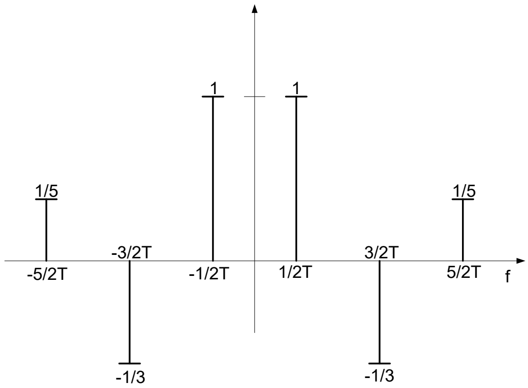


Figure 2.5. Spectrum of "worst case" square wave

Some assumptions have to be made about the average number of characters per word (5 here) and the number of symbols per character used by the code chosen (8.5 here). Hence, we take the particular case of 66 words per minute;
6 + 1 = characters per word (1 for space)
6 + 1 + 1.5 = 8.5 symbols per word (1.5 for "start" and 1 for "stop"
Number of words per minute fixed at = 66
Number of characters per minute = 66 x 7 = 462
Number symbols per minute = 66 x 7 x 8.5 = 3927
Number of symbols per second

$$= \frac{66 \times 7 \times 8.5}{60} = \frac{2970}{60} = 65.45$$

Therefore, symbol length

$$= \frac{1}{\text{Number of symbols per second}}$$

$$= \frac{1}{65.45} \cong \mathbf{15.28\ ms} = \text{T}.$$

So, to translate this into bandwidth, it is necessary to consider the "worst case" situation. This is the one for which the signal is changing the most rapidly, which is when the symbol stream consists of alternate zeros and ones as shown in Fig. 2.4 – drawn bipolar.

Fig. 2.4 is equivalent to a square pulse voltage stream of period 2T which will have standard spectrum as shown in Fig. 2.5, and a fundamental frequency f_0 of 1/2T or half the symbol rate.

The question then is, how much of this spectrum needs to be retained as it comes out of the teleprinter still in baseband form (e.g. to operate a printer). Here, there is another somewhat subjective criterion that the symbols may need to retain a reasonably "square" shape. To do this, it is usually regarded as adequate to send up to the third harmonic only. Thus,

$$\text{Baseband bandwidth } + \ 3\text{f}_0 = 3 + \frac{1}{2\text{T}} = \frac{3}{2\text{T}}$$
$$= 3(2 \text{ x } 0.2) = \mathbf{75\ Hz}$$

.

But for an actual teleprinter communication channel, some allowance must be made for guard bands and multiplexing so that this bandwidth is increased in practice to a standard value of 120 Hz. That is, teleprinter channel is between 0-120 Hz. Note that the d.c. term has to be retained in the spectrum.

2.3.3 Television (Analogue)

What baseband bandwidth is needed to send the luminance (light intensity) information in the standard UK TV signal? Very similar reasoning is used to that above for a teleprinter. Subjective criteria are used to decide values for the factors controlling the signal. Then, the symbol length is calculated and the bandwidth obtained from this. Here the subjective factors are more involved, and there are four of them:

1. Sufficient picture detail obtained by using 625-line scans per frame (525 in USA).
2. Vertical and horizontal resolutions kept the same by assuming smallest areas of constant intensity (picture elements or **pixels**) are square.
3. Most scenes framed adequately by using a rectangular picture shape of aspect ratio 4:3.
4. Unacceptable flickering avoided by making use of the eye's persistence of vision at the slowest convenient picture refresh rate of 25 frames per second (30 in Japan and North America).

Three of these criteria are illustrated in Fig. 2.6. The calculation of bandwidth now proceeds as follows:

Each line has 625 x 4/3 pixels in it = **833 pixels**

Each scan has 625 x 625 x 4/3 pixels in it = 520,000 pixels (520 kpixels)
Each scan takes 1/25 s to complete = **40 ms**
Each pixel takes 40 ms/520 kpixels to scan = **76.8 ns**
That is, pixel length (t_p) = **76.8 ns**

Again, we take the "worst case" situation in which the pixel intensity is changing most often and by the largest amount each time. This obviously occurs when alternate black and white elements occur as in Fig. 2.6 the result is a sort of "square wave" of light intensity which must be supplied by a square wave voltage signal of period $2t_p$ or 153 ns. This square wave therefore has a fundamental frequency of $f_0 = 1/153$ ns = 6.54 MHz

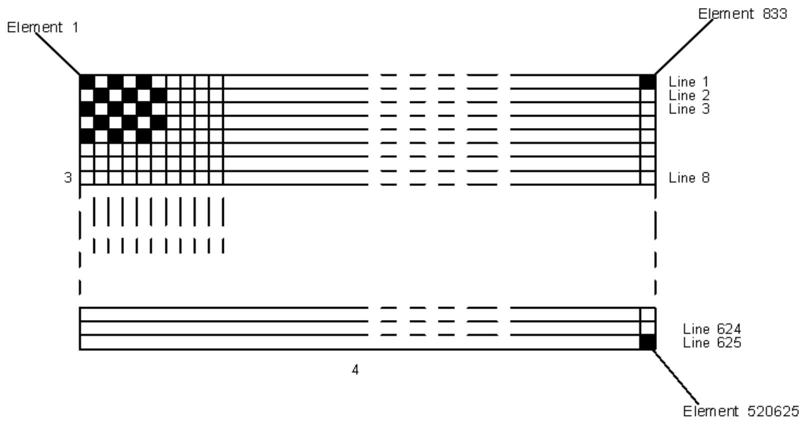


Figure 2.6. Scanning lines and pixels on a UK TV screen

But, because it is visually quite acceptable for pixel intensity to fall off at the edges, it is now only necessary to send f_0 itself without any of its harmonics. There is a need, however, to use more bandwidth to allow for channel separation, VSB (vestigial sideband) modulation, and sound signals. Thus, the actual r.f. frequency range allowed is UK TV bandwidth = 8.0 MHz (6.5 in USA - just calculated).

The baseband bandwidths of other signals in Table 2.1 can be calculated in a similar way – partly subjective and partly technical.

2.3.5 Signals and modulation

The electrical signal from the message or information is converted into electrical signals known as information or baseband signal. There are 2 broad classes:

1. Digital
2. Analogue

2.3.6 Baseband and Bandwidth Terminology

Signal type	Nature	Class	Baseband width	Usual modulation
Morse	Pulsed (telegraphy)	Digital C/V	0.50 Hz	ASK
Teletype	Pulse from keyboard	,,	0.120 Hz	FSK/PSK
Facsimile	still copies	,,	0-9.5 kHz	FSK/PSK
Telephone	voice	analog	0-9 kHz	SSB/FDM
Audio	Music	,,	0-15 kHz	FM
TV	Moving pictures	,,	0.6.5 MHz	VSB

2.4 The Need for Modulation

In general a baseband signal cannot be transmitted usefully without modification. The exceptions are some simple dedicated systems such as an internal telephone system or the line from a teletype machine to a printer. But for the vast mjority of systems communication would be either prohibitively expensive or actually impossible without the kind of changes which come under the heading of modulation. For example, a public telephone system would need a separate wire connection for each conversation and a radio link would require huge aerials and enormous power. Even then, only one station could operate at a time to avoid interference.

The answer is to change the baseband signal in some way to enable efficient economic communication methods to be used. Two distinct classes of methods come under these headings:

1. **Frequency translation:** moving the whole baseband up to a much higher frequency range.
2. **Digitizing:** changing the baseband to digital form, usually binary, by sampling.

Both are given the name modulation although the second keeps the signal in a baseband frequency region of difference region or different bandwidth. The baseband is changed in nature by a process which is strictly a form of source coding and not really by **"modulation** - alteration in amplitude or frequency of a wave by a frequency of different order." This is a good description of most frequency translation methods but not of digitizing ones. Perhaps we can forgive such a non-technical publication for not being entirely up to date in our speciality.

2.5 Classification of Modulation Types

All the methods of modulation described later in this book fall into one of the two categories mentioned in the last section. First we need a slightly fuller definition of each category:

Frequency translation: The baseband is moved to a higher frequency range by arranging for it to alter some property of a higher frequency carrier.

Sampling: The baseband waveform voltage is allowed through for short periods of time at regular intervals and these values only are sent, either coded or uncoded. The signal, although much changed in form, still remains essentially baseband in nature.

These categories now subdivided into the specific modulation types related by the techniques used, as shown in Fig. 2.7(b).

In addition, all the methods resulting from sampling can themselves be further modulated by frequency translation methods related to the keying methods already included for binary signals.

1. To transmit or radiate information energy as an electromagnetic wave into space, the length of the transmitting aerial needs to approach at least ¼ of a wave-length at the working frequency.

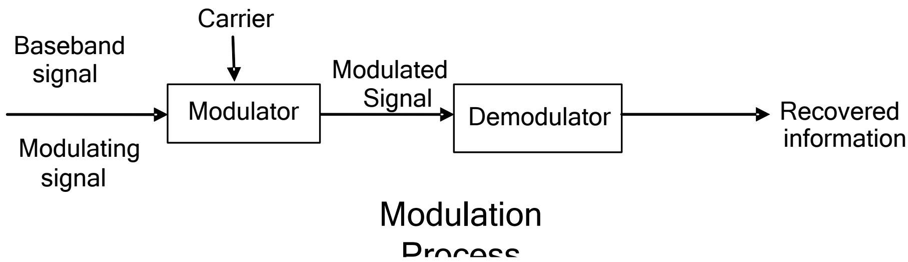
e.g. if $\text{f} = 10\ \text{kHz}$

$$\text{length} \simeq \frac{\lambda}{4\text{f}} = \frac{3 \times 10^8}{4 \times 10 \times 10^3} \simeq \mathbf{7,500\ m}$$

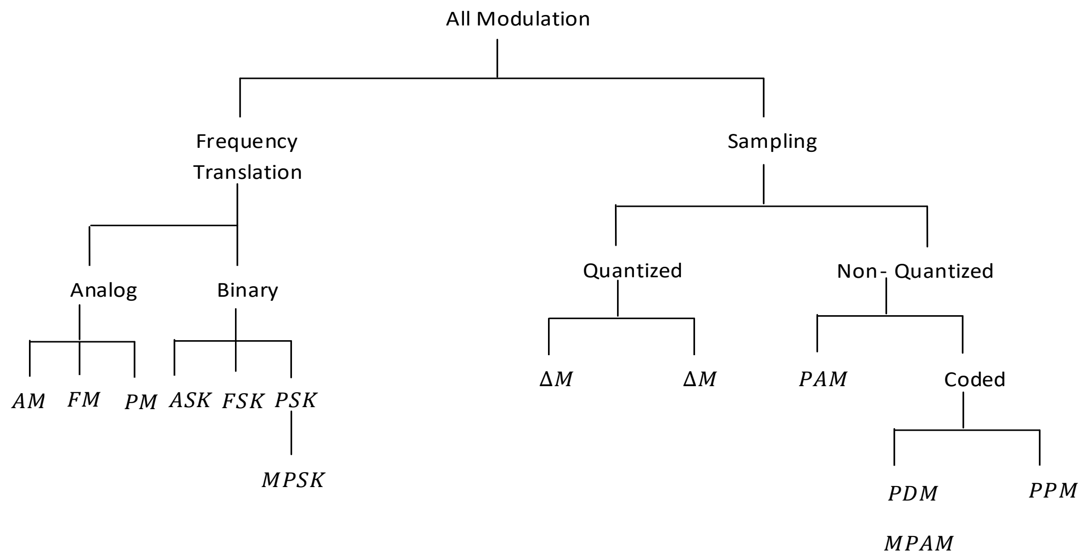
This is clearly not practicable, and its cost will be very prohibitive. It's difficult to transmit low frequency speech and music information signals directly as radio wave. Radio systems use high frequencies to 'carry' the low frequency information signals to the destination and the process that makes such possible is called MODULATION.

2. Often, there's need to send different information signals between 2 points. By transmitting the 2 voice information signals into different higher frequencies (MODULATION), the two can be sent over the same channel without interference. This technique is called MULTIPLEXING.

A shift of the range of frequencies in a signal is accomplished by using modulation. Thus, **modulation is defined as a process by which some characteristics of a carrier signal is varied in accordance with amplitude of a modulating or information signal.** The baseband signal is referred to as the modulating signal. The signal being modulated is the carrier. The result of the modulation process is referred to as the MODULATED signal. See Fig. 2.6(a)



(a)



(b)

Figure 2.7. Modulation process and family tree of modulation methods

The full meaning of the type abbreviations in Fig. 2.7(b) are as follows:

AM	Amplitude modulation
PM	Phase modulation
FM	Frequency modulation
ASK	Amplitude shift keying
PSK	Phase shift keying
FSK	Frequency shift keying
PAM	Pulse amplitude modulation
PDM	Pulse duration modulation
PPM	Pulse position modulation
PWM	Pulse width modulation, another name for PDM
ΔM	Delta modulation
PCM	Pulse code modulation
$A\Delta$PCM	Adaptive delta PCM
QAM	Quadrature amplitude modulation
M-	Multilevel signal (e.g. M-QAM).

2.6 Advantages of using modulation

There are several general reasons for modulating a baseband signal. Some have been mentioned already but are included again in this summary.

1. **Advantages produced by frequency translation**

(i) Use of frequency division multiplexing (FDM): This allows many signals to be sent simultaneously down the same communication channel. It gives economic use of equipment and enables systems to be designed.

(ii) Use of correct transmission frequency to give best transmission conditions: This is especially important in radio links where efficiency increases with frequency, and the best frequencies may need to be selected for propagation through the troposphere or via the ionosphere.

2. Advantage produced by sampling

(iii) Use of time division multiplexing (TDM): This allows many signals to be sent simultaneously along the same communication link by interleaving them in time. Gives similar economic and design advantages as (i).

3. Advantages produced by coding

(iv) Reliability of transmission greatly increased. Noise corruption very much less likely. Received baseband reproduces original signal very accurately.

(v) Signal processing much easier using standard logic and computing techniques. Facilities design and production of complex systems at their most economic and reliable as in modern telephone systems.

In general, it can be said that electrical and electronic communication systems would be virtually impossible without modulation. It is a vital part of most systems. The exact method used will depend on the application, and the next few chapters will look in detail at some of the methods. But first we must highlight some important features or characteristics.

2.7 Analogue Modulation-General

In analogue modulation, a continuously varying analogue signal changes some aspect of a carrier signal so that when sent through a transmission system, the baseband can be recovered intact from the carrier. The carrier is a single frequency which can be represented generally as

$$v_c = E_c \cos(\omega_c t + \emptyset_c) \qquad (2.1)$$

where E_c is the signal amplitude
ω_c is frequecy and $\emptyset_c$ is phase

To modulate a signal, there are only three things which must be changed by the baseband signal as shown above: amplitude, frequency or phase. Each of these leads to one class of analogue modulation:

changing $E_c \rightarrow$ Amplitude modulation
changing ω_c(i.e. f_c) $\rightarrow$ Frequency modulation
changing $\phi_c \rightarrow$ Phase modulation

A further general classification occurs if the carrier is rewritten as

$$v_c = E_c \cos(\theta_c) \quad (2.2)$$

From this you can see that both FM and PM alter the total phase angle in different ways. Thus, they are closely related and are often classified together as angle modulation**.** Hence, the three A's of continuous modulation: **Analogue = Amplitude + Angle**

Finally, note that many books write the equation for the carrier in sine form as

$$v_c = E_c \sin(\omega_c t + \phi_c) \quad (2.3)$$

Do not be confused by the Eq (2.2); It is merely the same form as used in (2.3) but with a phase difference of $\pi/2$ at time zero as $\cos\theta = \sin(\theta + \frac{\pi}{2})$. The cosine form seems to make the algebra marginally easier.

2.8 Digital Modulation-General

For the digital modulation techniques, the single common feature is the process of sampling – sending short bits of the analogue baseband at regular intervals. In the analogue sense, there is no carrier, and the baseband not only changes greatly in form but also remains baseband in nature.

In another sense, the sampling signal (the one which chops the bits out of the analogue baseband) acts as the carrier, as we shall see in the Chapter 8.

2.9 Summary

Signals are of two broad classes – analogue and digital. They vary greatly in waveform and bandwidth, and modulation used are depending on purpose (see Table 2.1).

A baseband is the original information signal in a communication system. The band of frequencies baseband occupies is the **baseband bandwidth** which usually starts from 0 Hz and is often quoted as a single figure in Hz. Some very common ones are:

Telephone channel	4 kHz
Hi-fi music on VHF	15 kHz
TV in USA	6.5 MHz
TV in UK	8 MHz

Modulation refers to changes produced by/in the baseband to facilitate its transmission. There are two broad classes of modulation:

Frequency translation	AM, FM, PM, etc.
Sampling and coding	PAM, PCM, etc.

With a third group combining the two in

Binary (or keying) modulation	ASK, FSK, PSK, etc

Analogue modulations are frequency translation methods caused by changing the appropriate quantity in a carrier signal as

$$v_c = E_c \cos(\omega_c t + \phi_c)$$

Digital modulation is the result of changing analogue signals into binary pulses by sampling and coding.

Keying modulations are digital signals subsequently modulated by frequency translation using one or either of analogue methods.

2.10 Conclusions

This chapter gives an overall picture of signals, uses and the purpose, and types of modulation they can undergo. Now, we start the long process of looking at these modulation methods in detail, with amplitude modulation in the next chapter.

2.11 Chapter Review Problems

2.1. For standard telephone channels answer the following questions:

(i) Give the baaseband frequency limits for a single telephone channel, then give its bandwidth and band limits if frequency by 60 kHz.

(ii) A standard telephone group consists of 12 adjacent channels. Give its bandwidth and its limits id its lowest frequency channel is that in (i) above.

(iii) What is the width of each guard band in (ii)?

(iv) Give the bandwidth of a telephone supergroup (= five groups).

(v) How many standard teleprinter channels can be sent down one telephone channel?

2.2. Morse code is usually transmitted with a dash equal to three dots; a space of one dot length between symbols; three dot lengths between characters and five between words. A skilled operator can send at 25 words per minute, assuming an average of 5 characters per word. Making reasonable assumptions calculate the overall symbol rate and assign a suitable baseband bandwidth.

2.3. A facsimile transmission system uses a cylinder 15.2 cm in diameter. The picture, on the outer surface of the cylinder, is scanned at 90 rpm with a lateral speed of 38 rev per cm. Each picture element is square. Calculate the minimum baseband bandwidth.

2.4. An image measuring 100 x 50 cm^2 in area is made up of alternate "light" and "dark" elements 1.0mm square. Calculate how long it will take to transmit this picture, by facsimile, through a channel band limited to 10 kHz.

2.5. One modern data Facsimile machine claims to be able to transmit an A4 page in 20 s using an information rate of 9600 baud. By making reasonable assumption show that the scan rate is around 6 lines per second. What line width does that represent?

2.6. Assign a suitable baseband bandwidth for a 525 line TV signal using alternate line scanning at 60 Hz. Use 4:3 aspect ratio and square pixels.

2.7. The width-to-height ratio of 405 line TV screen is 1.6:1. The horizontal definition is half that of the vertical definition. If the TV operates at 32 frames per second, determine the highest baseband frequency.

2.8. A facsimile document transmission system will send an A4 size picture in 6.5 s using pixels 1 mm square. Calculate the signal bandwidth required.

CHAPTER 3

AMPLITUDE MODULATION THEORY

3.0 Introduction

In an amplitude-modulated signal, the baseband information which is to be conveyed is impressed on to the carrier by varying its instantaneous amplitude. This leads to extra frequency components (the sidebands) which actually contain the information.

To economize power and bandwidth, some frequency components may be removed to give other forms of AM signal which do not show their AM nature as clearly. That is discuss is this chapter.

3.1 Types of Amplitude Modulation

There are three types of amplitude modulation:

"Full" amplitude modulation	(AM)
Double sideband with suppressed carrier	(DSBSC)
Single sideband	(SSB)

All three types are analysed below and their waveform with spectra illustrated as well as other aspects.

3.2 "Full" Amplitude Modulation

This type of modulation was the first in use in the early days of broadcasting in the 1920s and has developed several names in the course of time. It is often called just simply amplitude modulation (AM), but sometimes envelope modulation or even double sideband with carrier (DSBWC). Also the term full AM is often used to mean maximum amplitude modulation.

Whatever it is called, it is obtained by taking a single frequency carrier and altering its amplitude instantaneously in proportion to the instantaneous magnitude of a baseband signal. First, take the simple situation where the baseband is also a signal frequency sinusoidal. Then we can represent them both mathematically as

$$Carrier\ v_c = E_c cos\omega_c\, t \qquad (3.1)$$

$$Baseband\ v_m = E_m cos\omega_m\, t \qquad (3.2)$$

[Note that you may also find v_m written as $v_m(t)$ or $m(t)$ in different books. It simply means that the expression is time dependent or is a function of time.]

Now, we can arrange for E_c (carrier amplitude) to vary with
v_m (baseband instanteneous magnitude) but with the condition that it never becomes negative because that would lead to over-modulation. This means that E_c is replaced by $E_c + E_m \cos \omega_m t$ with $E_m \leq E_c$, giving a modulated signal which is now written as;

$$v_{AM} = (E_c + v_m) \cos \omega_c t$$

$$v_{AM} = (E_c + E_m \cos \omega_m t\ ) cos\omega_c t$$

$$= E_c(1 + \frac{E_m}{E_c} cos\omega_m t)\ cos\omega_c t$$
$$= E_c(1 + m\ cos\omega_m t)cos\omega_c t$$

Where the m is **modulation factor** and must have a value between 0 and 1 (to avoid E_c becoming negative as stipulated above). m is a very important characteristic of any system which uses full AM. It is often quoted as a percentage and in different thesis or occasions represented with m_a, β_m, or even K_m. Some people use full AM to mean m= 100% modulation only and not any level of envelope modulation as here. In the eqs (3.4 and 3.5), we extracted modulation factor to be

$$m = \frac{E_m}{E_c}$$

However, we must beware of using this definition as a basis for measurement of m. The two amplitudes are those at the point of modulation and not at the input terminals of the modulator or anywhere else.

Thus, the modulated signal as from expression (3.5) is

$$v_{AM} = E_c(1 + m\cos\omega_m t)\cos\omega_c t$$

This produces waveform as shown in Fig. 3.1 for two values of m.

See how clearly this diagram shows the idea of this method as envelope modulation. Note, however, that the modulating waveform is not actually there; it just looks as if it is because the locus of the peaks of the modulated carrier follows it. It forms an envelope containing the carrier, and its shape is very clearly defined, especially if $f_c \geq f_m$ so that the carrier peaks are very close together.

m can be measured directly from this waveform by measuring the peak-to-peak voltages A and B in Fig. 3.2.
this can be seen as follows:

$$m = \frac{E_m}{E_c} = \frac{\frac{1}{2}(E_c + E_m) - \frac{1}{2}(E_c - E_m)}{\frac{1}{2}(E_c + E_m) + \frac{1}{2}(E_c - E_m)} = \frac{\frac{1}{4}A - \frac{1}{4}B}{\frac{1}{4}A + \frac{1}{4}B}$$

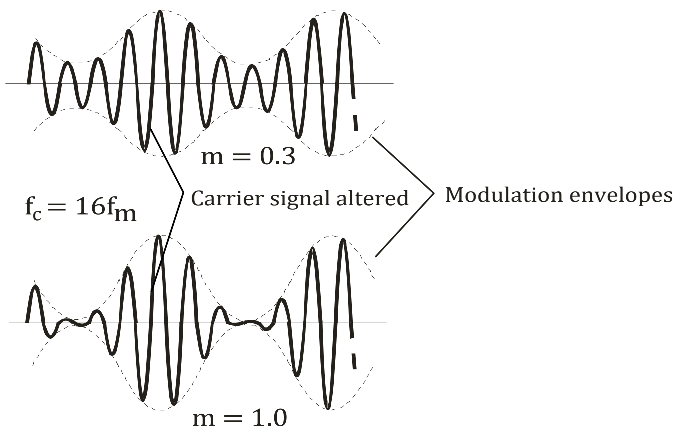


Figure 3.1. Full Am waveforms

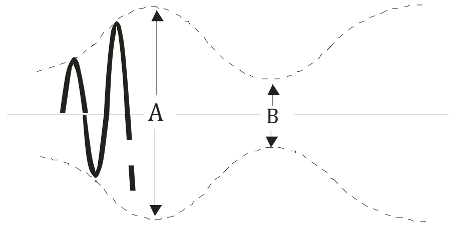


Figure 3.2. Measuring modulation factor

Therefore

$$m = \frac{A - B}{A + B}$$

This modulated carrier is obviously not a single frequency. Its spectrum can be obtained from the following analysis:

v_{AM} $= E_c(1 + m\cos\omega_m t)\cos\omega_c t$

$= E_c\cos\omega_c t + mE_c\cos\omega_m t\cos\omega_c t$

$= E_c\cos\omega_c t + \frac{1}{2}mE_c\,[\cos(\omega_c - \omega_m)t + \cos(\omega_c + \omega_m)t]$

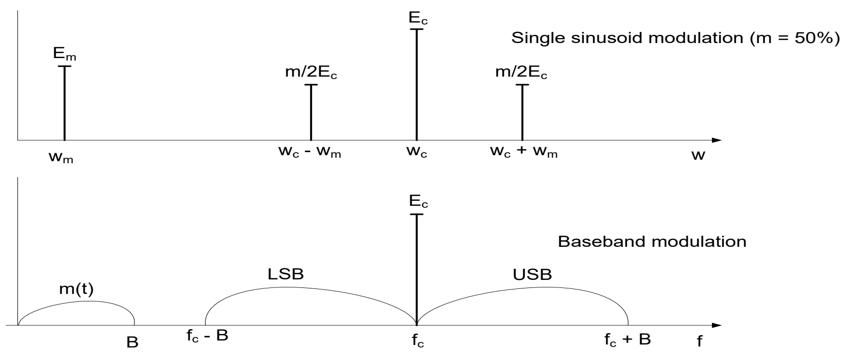


Figure 3.3. Full AM spectra

or

$$v_{AM} = E_c\cos\omega_c t + \frac{mE_c}{2}\cos(\omega_c - \omega_m)t + \frac{mE_c}{2}\cos(\omega_c + \omega_m)t$$

3.3 Amplitude Modulation General

$$v_c = E_c\sin(\omega_c t + \emptyset_c)$$
$$v_c = E_c\sin(2\pi f_c t + \emptyset_c)$$

$\omega_c = 2\pi f_c$

$$v_m = E_m\sin 2\pi f_m t$$

From trigonometry, $E_c\cos 2\pi f_c t = E_c\sin(2\pi f_c t + \emptyset_c)$

Alter E_c in accordance with v_m

Alter f_c in accordance with f_m

$\emptyset_c$ in accordance with P_m

Amplitude modulation is obtained by taking a single frequency carrier and altering its amplitude instantaneously.

One frequency component modulation signal

$$v_{AM} = (E_c + v_m)\cos 2\pi f_c t$$
$$= (E_c + E_m \cos 2\pi f_m)\cos 2\pi f_c t$$
$$= E_c \cos 2\pi f_c t + E_m \cos 2\pi f_m t \cos 2\pi f_c t$$
$$v_{AM} = E_c \cos 2\pi f_c t + \tfrac{1}{2} E_m \cos 2\pi\,(f_c - f_m)t + \tfrac{1}{2} E_m \cos 2\pi\,(f_c + f_m)t$$
$$E_m = mE_c$$
$$m = \frac{E_m}{E_c}$$
$$v_{AM} = E_c \cos 2\pi\, f_c + \frac{1}{2} mE_c \cos 2\pi\,(f_c + f_m) + \frac{1}{2} m\, E_c \cos 2\pi\,(f_c - f_m)t$$

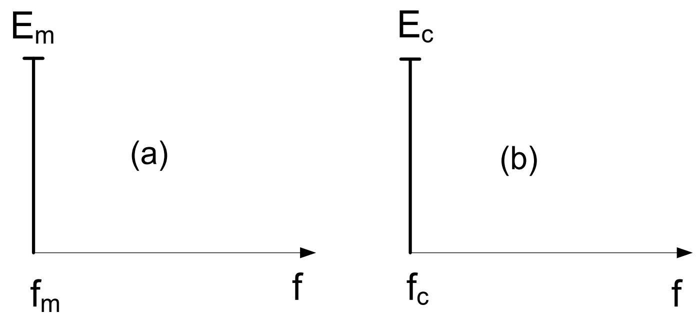


Figure 3.4 Frequency spectrum of (a) basedband and (b) carrier signal

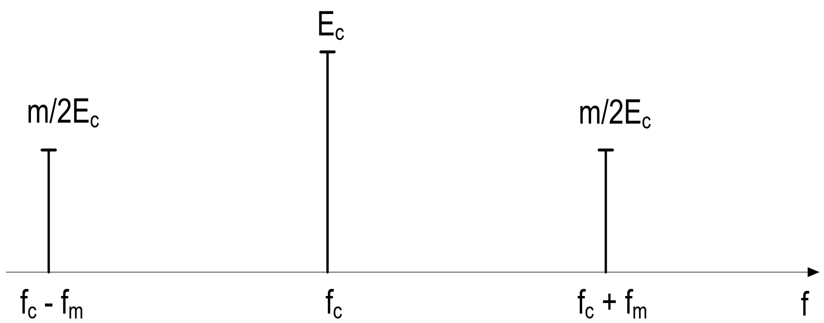


Figure 3.5 Frequency spectrum of the signal v_{AM}

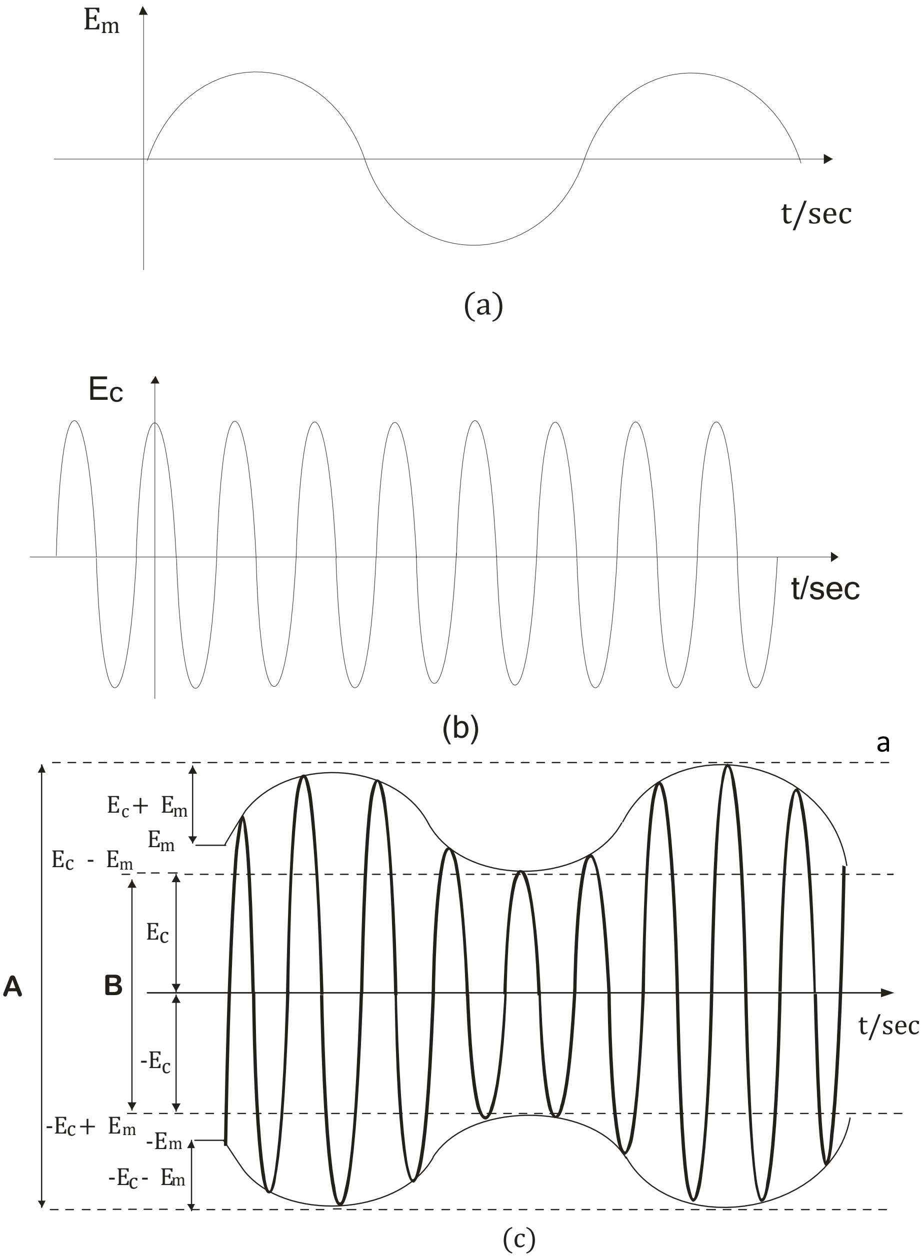


Figure 3.6 Information signal (a) Carrier signal (b) Modulating Voltage signal (c)

Where;

$$A = 2(E_c + E_m)$$

$$B = 2(E_c - E_m)$$

Now, in addition to the original unmodulated carrier, there are two new frequencies present, one greater that f_c by the amount of the m modulating frequency, f_m, and one less than it by the same amount. These are the **upper and lower side**

frequencies, $(f_c - f_m)$ and $(f_c + f_m)$. Each of amplitude m/2 less than Ec, and are shown in Fig. 3.3 together with the analogous situation for a baseband which is a band of frequencies and not just a single sinusoidal. For both cases, the baseband is shown as well but is not, of course, part of the modulated spectrum.

The lower spectrum is the more usual practical situation because most actual basebands have a spread of frequencies and often contain a continuous band of frequencies from near zero up to some band limit of B Hz (e.g. voices). Then, each frequency of the band produces its own side frequency pair above and below f_c. The effect is to produce sidebands, the **upper** one (USB) of which has the same spectral shape as v_m whilst the **lower** one (LSB) has this shape reversed in mirror-image form.

But be warned that although widely used, there is some confusion in this second spectrum type because the baseband is now a spectral density distribution with a vertical axis of volts per hertz (V Hz^{-1}). The single frequency carrier is still plotted with a vertical axis of volts only so that the modulated signal has two different vertical scales on the same plot. Thus, this spectrum is really a hybrid which must only be used for what it is – a useful diagrammatic way of illustrating what happens. Do not use it quantitatively – for example, the upper spectrum can show clearly that the side frequency cannot have amplitude greater than $\frac{E_c}{2}$, but this cannot be done with the lower one.

So, let us safely take the upper single frequency situation and show how little power there is in the sidebands, which carry the baseband information. This is done as follows:

Carrier power

$$\left(\frac{E_c}{\sqrt{2}}\right)^2 = \frac{{E_c}^2}{2}$$

Side frequency power

$$= 2\left(\frac{\mathrm{mE_c}}{2\sqrt{2}}\right)^2 = \frac{\mathrm{m}^2\,{E_c}^2}{4}$$

$$\frac{\text{carrier power}}{\text{s.f. power}} = \frac{\frac{{E_c}^2}{2}}{\frac{\mathrm{m}^2{E_c}^2}{4}} = \frac{2}{\mathrm{m}^2} \geq \mathbf{2}$$

for all m

Thus, at least two-thirds of the transmitted power is in the carrier and conveys no information. Therefore, it can be dispensed with (see Section 3.4).

m can be measured from the modulated spectrum (as well as from the waveform Figure 3.2) merely by taking the ratio of the sideband amplitude to the carrier amplitude (for a single sinusoidal baseband, of course). That is

$$\frac{\text{sideband amplitude}}{\text{carrier amplitude}} = \frac{\frac{\mathrm{mE_c}}{2}}{\mathrm{E_c}} = \frac{\mathrm{m}}{2}$$

In practice, this particular measurement is done on a spectrum analyser using the log amplitude (dB) scale so that the ratio is the difference (in dB) as

Difference (in dB) = $20\log_{10}$(ratio) = $20\log_{10}(2/m)$

$$\mathrm{m} = \frac{2}{\mathrm{antilog(diff.\,20)}}$$

[For example, 6 dB makes m= 2/antilog (0.3) = 2/2 = 1 or 100%.]

The bandwidth needed to transmit a full AM signal can also be seen from the spectrum. It is 2B Hz which is twice that of a baseband. That is

B/W =2B

A full AM signal, with single sinusoid baseband, can be well described graphically using a quasi-stationary phasor as in Fig.3.10. Let's work out the summary of all our discussions so far, an example.

Example 3.1

The signal $\mathrm{i_1} = 30\cos(2\pi \mathrm{f_1 t})$ is used to amplitude modulate the signal $\mathrm{i_2} = 40\cos(2\pi \mathrm{f_2 t})$, $\mathrm{f_1} = 40$ kHz and $\mathrm{f_2} = 4$ MHz.

i) State the name of $\mathrm{i_1}$ and $\mathrm{i_2}$.

ii) Obtain the modulation index.

iii) Sketch a well labelled waveform of $\mathrm{i_1}$, $\mathrm{i_2}$ and the resulting amplitude modulated signal.

iv) Sketch the frequency spectral of $\mathrm{i_1}$, $\mathrm{i_2}$ and the resulting amplitude modulated signal.

v) Is it right to replace the first signal with $\mathrm{i_1} = 50\cos(2\pi \mathrm{f_1 t})$. If no, explain. State the consequence (if any).

vi) Suppose the first signal is replaced with $\mathrm{i_1} = 40 + 30\cos(2\pi \mathrm{f_1 t}) + 20\cos(20\pi \mathrm{f_1 t})$, sketch the frequency spectrum of the resulting amplitude modulated signal.

(Elect/Elect TEL 412, University of Ibadan, 2004/2005).

Solution

i) $\mathrm{i_1} = 30\cos(2\pi \mathrm{f_1 t})$, is the **baseband signal or information signal**

$\mathrm{i_2} = 40\cos(2\pi\, \mathrm{f_2 t})$ is the **carrier signal**.

ii) Modulation index $(\mathrm{m}) = \frac{\mathrm{I_m}}{\mathrm{I_c}} = \frac{30}{40} = \mathbf{0.75}$.

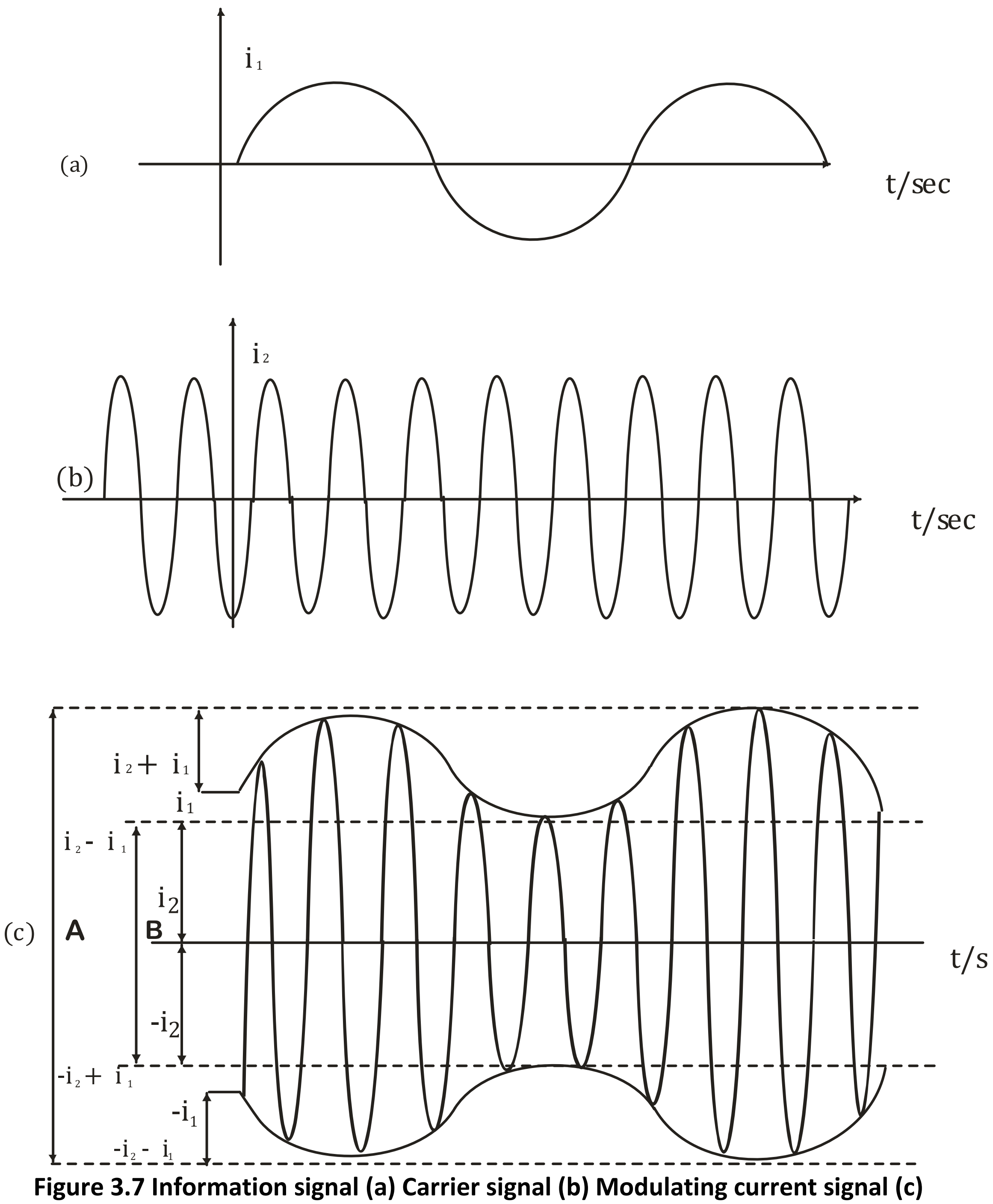


Figure 3.7 Information signal (a) Carrier signal (b) Modulating current signal (c)

iii) $i_1 = 30\,cos(2\pi\, f_1 t),\ i_2\ = 40\,cos(2\pi f_2 t),\ f_1 = 4\ \text{kHz}\ \text{f}_2 = 4\ \text{MHz}$

$$i_{AM} = [40 + 30\,cos(2\pi\, f_1 t)]\,cos\,2\pi f_2 t$$

$$= 40\,cos\,2\pi f_2 t + 30\,cos\,2\pi\, f_1 t\,cos\,2\pi f_2 t$$

But $cos\,A\,cos\,B = \frac{1}{2}[cos(A + B) + cos(A - B)\,]$

$$i_{AM} = 40\, cos\, 2\pi\, f_2 t + 15\, cos\, 2\pi(f_2 + f_1)t + 15\, cos\, 2\pi(f_2 - f_1)t$$
$$= 40\, cos\, 8\pi t + 15\, cos\, 8.008\pi t + cos\, 8.008\pi t + 15\, cos\, 7.992\pi t$$
$$\boldsymbol{i_{AM} = 40\, cos\, 2\pi(4)t + 15\, cos\, 2\pi(4.004)t + 15\, cos\, 2\pi(3.996)t}$$

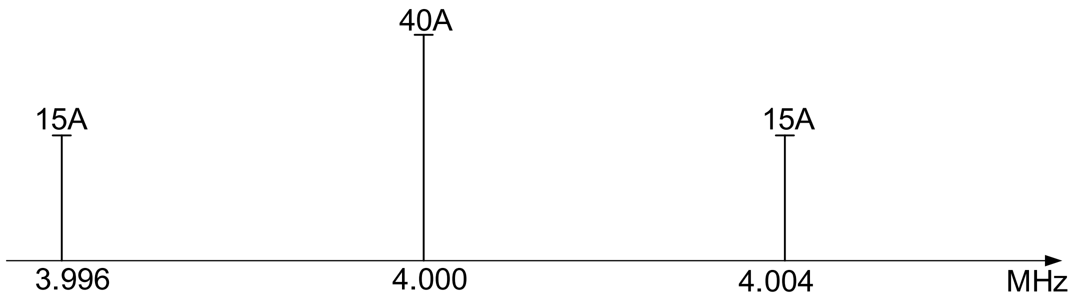


Figure 3.8 Frequency spectrum of the modulated current signal

It is not right to replace the first signal with $i_1 = 40\cos(2\pi f_1 t)$ because the modulation index will be greater that one. The consequence is that the information signal is not recoverable.

If $i_1 = 40 + 30\, cos\, 2\pi f_1 t + 20\, cos\, 20\, \pi f_1 t$

$$i_{Am} = i_1\, x\, i_2 = [40 + 40 + 30\, cos\, 2\pi f_1 + 20\, cos(20\pi f_1)]\, cos(2\pi f_2 t)$$
$$= 80\, cos(2\pi\, f_2 t) + 30\, cos\, 2\pi f_1 t\, cos\, 2\pi f_2 t + 20\, cos\, 20\pi f_1 t\, cos\, 2\pi f_2 t)$$

$but\ f_1 = 0.004$ MHz $and\ f_2 = 4$ MHz $and\ \ cos\, A\, cos\, B = \frac{1}{2}[cos(A + B) + cos(A - B)]$

$$i_{Am} = 80\, cos\, 2\pi\, f_2 t + \tfrac{30}{2} cos\, 2\pi\, (f_2 - f_1)t + \tfrac{30}{2} cos\, 2\pi(f_2 + f_1)t + \tfrac{20}{2} cos\, 2\pi(f_2 + 10f_1)t + \tfrac{20}{2} cos\, 2\pi(f_2 - 10f_1)t$$

$$i_{Am} = 80\, cos\, 2\pi\, f_2 t + 15\, cos\, 2\pi(f_1 + f_2)t + 15\, cos\, 2\pi(f_2 - f_1)t + 10\, cos\, 2\pi(f_2 + 10f_1)t + 10\, cos\, 2\pi(f_2 - 10f_1)t$$
$$= 80\, cos\, 8\pi t + 15\, cos\, 8.008\pi t + 15\, cos\, 7.992\pi t + 10\, cos\, 8.08\pi t + 10\, cos\, 7.92\pi t$$
$$\boldsymbol{i_{Am} = 80\, cos\, 2\pi(4)t + 15\, cos\, 2\pi(4.004)t + 15\, cos\, 2\pi(3.996)t}$$
$$\boldsymbol{+ 10\, cos\, 2\pi(4.04)t + 10\, cos\, 2\pi(3.96)t}$$

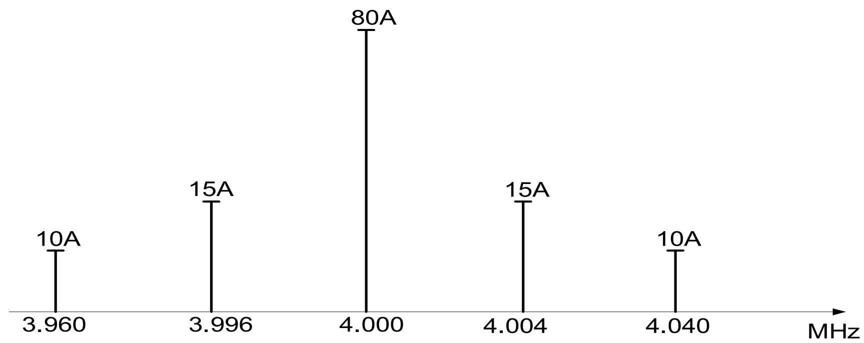


Figure 3.9 Frequency spectrum of the resulting amplitude modulated signal

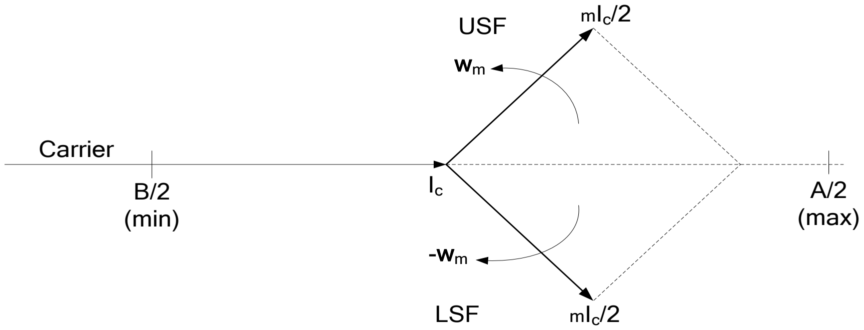


Figure 3.10 Phasor representation of full AM

The carrier itself is represented by a stationary phasor of constant amplitude E_c and at phase zero. Then the side frequencies are smaller phasors rotating symmetrically about the end of E_c at the same speed but in opposite directions, that is at ω_m and - ω_m as shown. Each will be of length $\frac{1}{2}$ mE_c and of opposite phase relative to E_c. At any instant the actual signal amplitude will be the phasor sum of these three so that the points marked A/2 and B/2 correspond to the maximum and minimum amplitude already shown in Fig. 3.2. Compare this with the phasor diagram for NBPM in chapter 7.

3.4 Double Sideband Suppressed Carrier (DSBSC) Modulation

As have seen, full AM is fairly easy to visualize and, as it turned out very easy to demodulate. Conversely, it does have the two disadvantages mentioned earlier: it wastes both power and bandwidth. Power sent as carrier contains no information independently giving unnecessary duplication. The modulated signal being DSBs is containing same message on both sides therefore placing double pressure both on equipment, cost, and medium with bandwidth being wasted.

Solutions to this limitation: Part of this can be overcome by transmitting only the sidebands. One way to do this would be to remove the carrier from a full AM signal, leading to the usual name for this type of modulation – **double sidedband suppressed carrier** or DSBSC for short. In practice, it is usually obtained more directly by multiplying the carrier and baseband signal together in a balanced modulator. That is

$$v_{DSBSC} = v_{AM} - v_c$$

Also

$$\begin{aligned} v_{DSBSC} &= v_m \times v_c \\ &= E_m \cos \omega_m t \times E_c \cos\omega_c t \\ &= \frac{1}{2} E_m E_c[\cos(\omega_c - \omega_m)t] \end{aligned}$$

This is simply written as

$$v_{DSBSC} = E_D \cos(\omega_c - \omega_m)\, t + E_D \cos(\omega_c + \omega_m)\, t$$

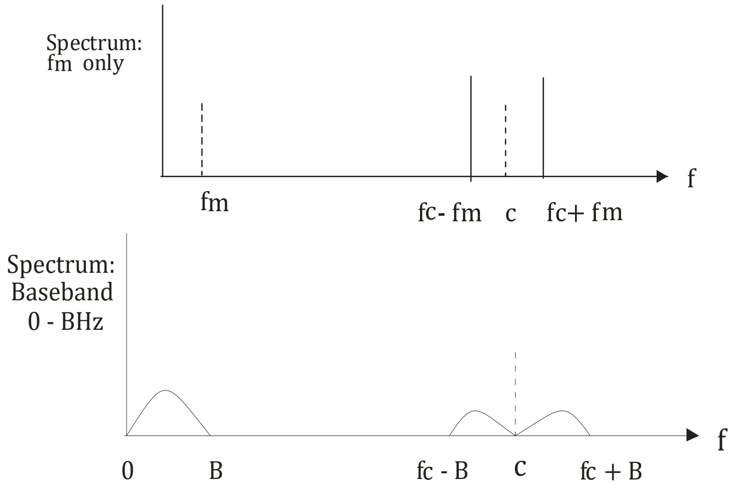


Figure 3.11 DSBSC spectra

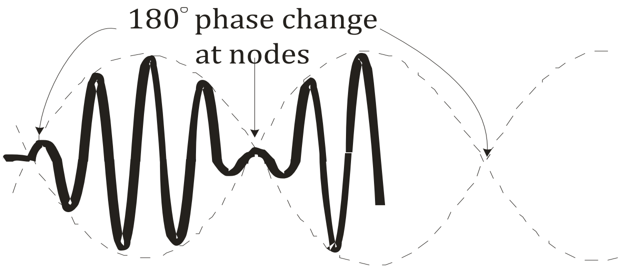


Figure 3.12 DSBSC waveform for single sinusoidal baseband

E_D is merely an amplitude factor proportional to both E_C and E_M with the dimensions of voltage not voltage squared (a popular misconception). This gives the characteristic double sideband spectrum as in the upper part of Fig. 3.11.
The waveform itself is even more characteristic and is shown in Fig. 3.12 for a single sinusoidal baseband.

This waveform is merely the carrier signal with its instantaneous amplitude fixed by the baseband voltage. The zero crossing points stay fixed but the peaks vary in height and position because the instantaneous signal amplitude is the baseband voltage. Thus, the modulated signal fits within an envelope of the baseband waveform, positive and negative. One effect of this is that the signal has an abrupt 180^0 phase change when the baseband passes through zero and changes sign. This occurs at the nodes of the waveform and is illustrated in Figs. 3.12 and 3.13. Another way to look at the characteristic "blobby" waveform in Fig. 3.7 is that it is the beat signal between two signals close together in frequency – the sidebands.

Also included in Fig. 3.13 is the zero signal part of a waveform of a full AM signal for which m = 1. Both that and the DSBSC nodes are very similar in appearance but with two important differences.

a) Envelope zeros occur every $\lambda_m/2$ for DSBSC but only every λ_m for full AM.
b) Envelope lines actually cross for DSBSC but touch asymptotically for full AM.

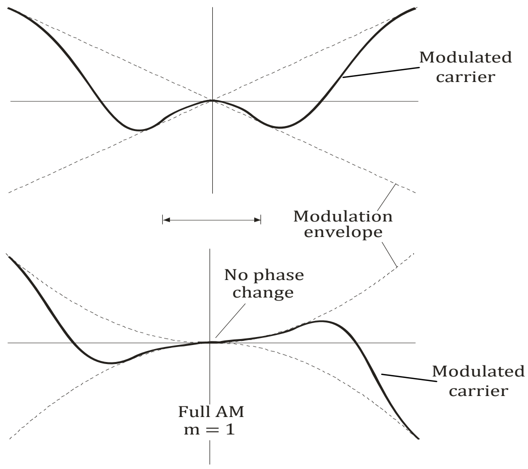


Figure 3.13 Phase changes at nodes

Of course, for a non-sinusoidal baseband the spectrum and waveform are more complicated in a very similar way to that for full AM. Fig. 3.11 has already shown a spectrum and now Fig. 3.14 shows a waveform.

As can be seen from the spectra, the bandwidth of a DSBSC signal is exactly the same as that for a full AM one, that is

$$B/W = 2f_{max}$$

One familiar use of this method of modulation is in submariner modulation of the L=R signal in stereo VHF FM radio.

3.5 Single Sideband Modulation

The final simplification in amplitude modulation is to send one sideband only. This is possible because it contains all the signal information (E and f_m) with less signal power and half the bandwidth.

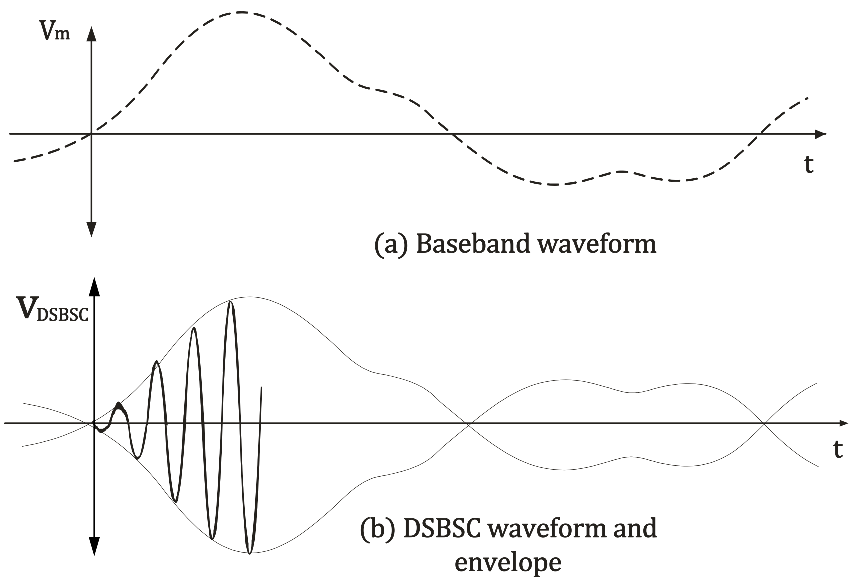


Figure 3.14 General spectra and waveform for DSBSC modulation

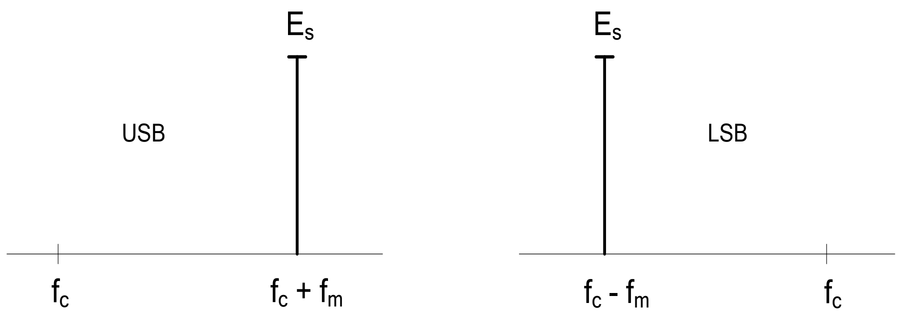


Figure 3.15 SSB spectra (after Holds worth and Martin, 1991)

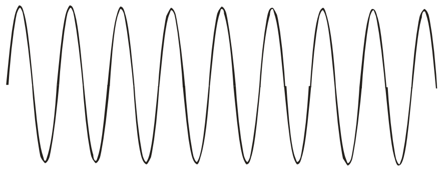
The simplest way to obtain an SSB signal is in principle at least, to take a DSBSC signal and remove one sideband by filtering. This leaves the other sideband only- either the upper or lower side, hence, producing from the two a SSB signal. For a single sinusoidal baseband, these can be written simply as

$E_s \cos\ (\omega_c - \omega_m)$ the **lower sideband** (LSB)

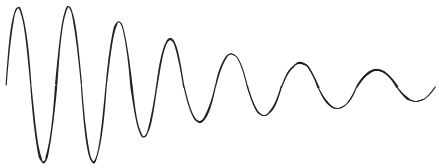
$E_s \cos\ (\omega_m + \omega_m)$ the **upper sideband** (USB)

Where E_s is proportional to E_c and E_m as for DSBSC: these produce spectra as in Fig. 3.15.

Relevant waveform are given in Fig. 3.16 and at first sight seem to have little meaning, looking just like the unmodulated carrier. On close inspection, in a suitable experimental set-up, it is possible to see that the frequency varies by very small amounts (i.e. f_m changes; hence so does $f_c + f_m$ or $f_c - f_m$, but not much because $f_c \geq f_m$). It is easier to see that the signal amplitude changes proportional to changes in baseband voltage (v_m).



(a)



(b)

Figure 3.16 SSB waveform: (a) for an unchanging baseband; (b) with variations in amplitude and frequency of the baseband (after Holds worth and Martin, 1991)

The bandwidth is halved to become

$$B = f_{max} = \text{baseband bandwidth}$$

SSB is used in many applications. One we use nearly every day occurs in our telephone system where signals are LSB modulated onto suppressed sub carriers to enable many to be sent down the same communication channel in FDM.

Another well-known application is in radio communication where SSB is used to conserve bandwidth.

3.6 SSB Circuits

There are two primary methods of generating SSB signals. These are either by the filter method or by the phasing method. The filter method is by far the simplest and most widely used, but we will discuss both types here.

1. **The Filter Method of SSB**

Fig. 3.17 shows a general block diagram of an SSB transmitter using the filter method. The modulating signal usually voice from a microphone, is applied to the audio amplifier whose output is fed to one input of a balanced modulator. A crystal oscillator provides the carrier signal which is also applied to the balanced modulator. The output of the balanced modulator is a DSB signal. An SSB signal is produced by passing the DSB signal through a highly selective band-pass filter. This filter selects either the upper or the lower sideband.

The filter of course, is the critical component in the filter method of SSB generator. Its primary requirement is that it has high selectivity so that it passes only the desired sideband and rejects the other. The filters are usually designed with a bandwidth of approximately 2.5 to 3 kHz, making them only wide enough to pass standard voice frequencies. The sides of the filter response curve are extremely steppy, providing - excellent rejection of the other sideband.

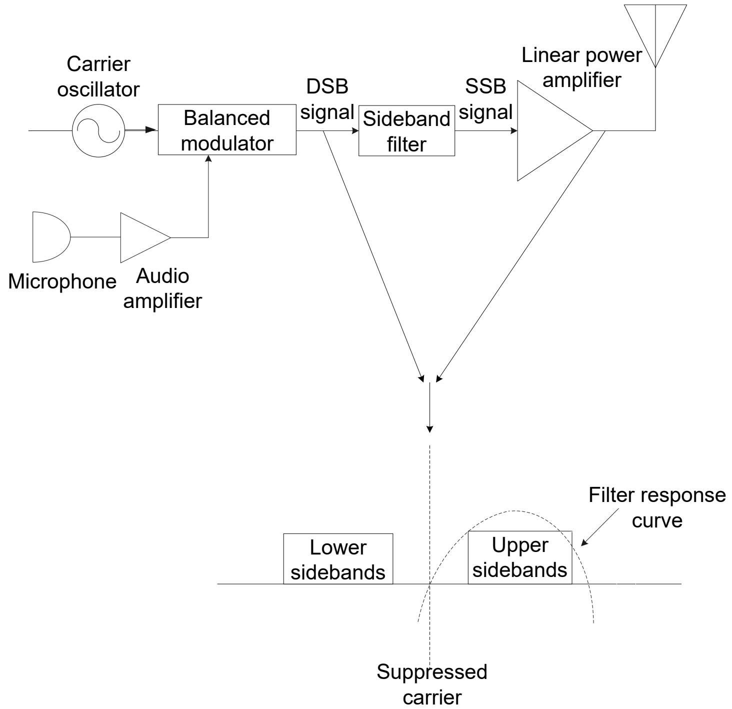


Figure 3.17 An SSB transmitter using the filter method

The filter is a fixed tuned device; the frequencies that it can pass cannot be changed. Therefore, the carrier oscillator frequency must be chosen so that the sidebands fall within the filter bandpass. Usually, the filter is tuned to a frequency in the 455 kHz, 3.35 MHz, or 9 MHz range. Other frequencies are also used, but many commercially available filters are in these frequency ranges.

It is also necessary to select either the upper or the lower sideband. Since the same information is contained in both sidebands, it generally makes no difference which one is selected. However, various conventions in different communications services have chosen either the upper or the lower sideband as a standard. These vary from service to service, and it is necessary to know whether it is an upper or lower sideband to properly receive an SSB signal.

There are two methods of selecting the sideband. Many transmitters simply contain two filters, one that will pass the upper sideband and the other that will pass the lower sideband. A switch is used to select the desired sideband. The other method of selecting the sideband is to provide two carriers oscillator frequencies. Two crystals change the carrier oscillator frequency to force either the upper sideband or the lower sideband to appear in the filter bandpass.

3.7 Vestigial Sideband Modulation (VSB)

This is used for sideband modulating signal: such as TV, where the bandwidth of the modulating signal can extend up to 5.5 MHz. DSB transmission would require 11 MHz bandwidth. This is very excessive in view of transmission bandwidth occupation and of cost.

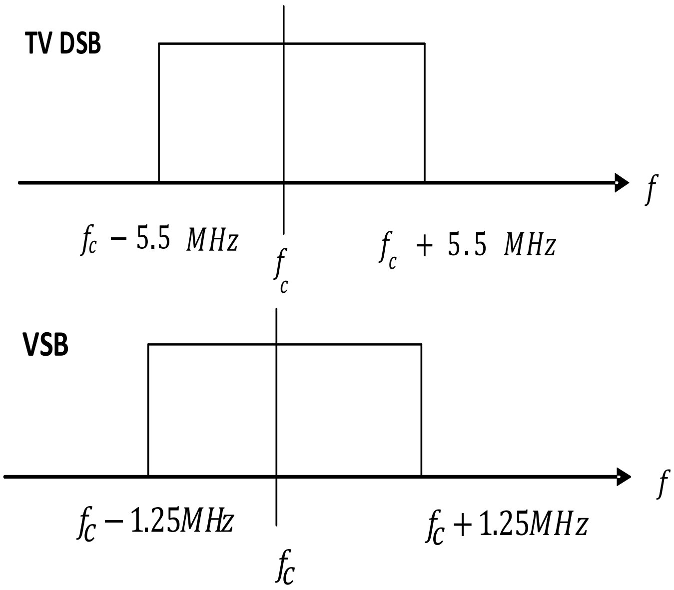


Figure 3.18

DS is a compromise in VSB, one sideband and an important part of the other sideband are transmitted. There is a saving in power and bandwidth compared to DSB. More information is transmitted compared to SSB. VSB have much simpler receiver than SSB.

3.8 Summary

There are three types of amplitude modulation:

1. Full AM
2. Double sideband suppressed carrier (DSBSC)
3. Single sideband (SSB).

For a single sinusoidal baseband a carrier modulated by them is given by the following.

Full AM

$$v_{AM} = E_c(1 + m\cos\omega_m t)\cos\omega_c t$$
$$= E_c\cos\omega_c t + \frac{1}{2}mE_c\cos(\omega_c - \omega_m)t + \frac{1}{2}mE_c\cos(\omega_c + \omega_m)t$$

m is the modulation factor:

$$m = \frac{E_m}{E_c}$$

at the point of modulation $m \leq 1$)

$$m = \frac{A - B}{A + B}$$

on the waveform (see Figure 3.2)

$$m = \frac{2 \text{ x s.f amplitude}}{\text{carrier amplitude}}$$

on the spectrum (see Figure 3.3), and

$$m = \frac{2}{\text{antilog (amplitude ratio in } \frac{dB}{20})}$$

where B is the baseband bandwidth

DSBSC

$$v_{DSBSC} = E_D \cos(\omega_c - \omega_m)t$$

The waveform has a characteristic 180^0 carrier phase reversal at zero crossings of the modulation envelope (Figs. 3.6 and 3.7):

$$B/W = 2f_m = 2\beta$$

SSB

Either

$$v_{SSB} = v_{LSB} = E_s \cos(\omega_c - \omega_m)t$$

or

$$v_{SSB} = v_{USB} = E_s \cos(\omega_c + \omega_m)t$$

$$B/W = f_m = \beta$$

For stationary phasor representation see Fig. 3.10 for the effect of actual basebands see Fig. 3.3. etc.

Two of these amplitude modulation methods are in very common use: full AM for broadcasting, and SSB for telephone and radio communication. But they all have serious deficiencies in dynamic range and in noise immunity. To improve matters we need to go to an entirely different method of modulation FM as described later, but first we will discuss amplitude modulators and demodulators.

3.9 Chapter Review Problems

3.1. Full AM is produced by a signal, $v_m = 3.0\cos(2\pi \times 10^3)t$ volts, modulating a carrier, $v_c = 10.0\cos(2\pi \times 10^6)t$ volts. Calculate:

- (i) m
- (ii) Side frequencies and bandwidth.
- (iii) Ratio of sideband amplitude to carrier amplitude.
- (iv) Maximum and minimum peak-to-peak amplitude of the modulated waveform.
- (v) The percentage of the total power in the side frequencies.

3.2. An AM signal voltage is given by $v = 100\sin(2\pi \times 10^6)t + [20\sin(6250t) + 50\sin(12\,560t)]\sin(2\pi \times 10^6)t$:

- (i) What type of amplitude modulation is being used?
- (ii) Draw its amplitude spectrum
- (iii) Sketch its waveform over one modulation cycle showing quantitative values for important times and amplitudes.
- (iv) Work out the peak and average powers into a 100 Ω load.
- (v) Does *m* have values? If so, what are they?
- (vi) What is the significance (if any) of using sines not cosines?

3.3. Repeat the last problem but with the carrier suppressed (questions (i) to (v) only).

3.4. Discuss the relative merits of the three main types of amplitude modulation. For each give at least one unique advantage relative to the other two, and one disadvantage.

3.5. A carrier, f_1 to f_2, with amplitude proportional to frequency:

- (i) Draw the baseband amplitude spectrum
- (ii) Draw the full AM spectrum showing relative amplitudes for maximum modulation.
- (iii) Give the bandwidths for full AM and DSBSC
- (iv) Derive an expression for the proportion of power in the sidebands for full AM at maximum modulation.

3.6 A radar signal consists of 1 μs pulses of a 10 GHz carrier. The pulses are of constant amplitude and are spaced by 12 μs gaps during which nothing is transmitted.

- **(i)** Sketch the modulation envelope
- **(ii)** Draw the modulation spectrum showing frequencies and relative amplitudes.
- **(iii)** Assign a bandwidth giving reasons
- **(iv)** Repeat for an interval of 99 μs between pulses.

3.7 A signal is band limited to the frequency range 0-5 kHz. It is frequency translated by multiplying it by the signal $v_c = \cos 2\pi f_c\, t$. Find f_c so that the bandwidth of the translated signal is 1% of f_c.

3.8 A DSBSC transmitter delivers 10 W when the modulating signal is d.c. voltage of 10.V. determine the power delivered when the modulating signal is a tone of RMS value 1.6 V.

3.9 Show that the signal $v = \sum[\cos\omega_c\, t.\cos(\omega_i\, t + \theta_i)]$
Is an SSBSC signal. Is it upper or lower sideband? Write an expression for the other sideband. Obtain an expression for the total DSBSC signal.

3.10 Two transmitters, one generating full AM signal and the other an SSB one, have equal mean output power ratings. A single sine wave modulating signal causes the AM signal to have a modulation factor 0.8. if both transmitters are operating at maximum output, compare (in dB) the power contained in the sidebands of the full AM signal with that contained in the SSB one.

3.11 (a) (i) Explain briefly why information signal need to undergo the process of modulation.

(ii) Mention and explain the two main categories of modulation.

(iii) What is frequency Modulation?

(iii) Compare and contrast AM and NBFM. Illustrate with diagrams.

(b) The signal $i_1 = 30cos\ (2\pi f_1 t)$ is used to amplitude modulate the signal
$i_2\ = 40cos\ (2\pi f_2 t)\ , f_1 = 4\ \text{kHz}\ \ \&\ \ f_2 = 4\ \text{MHz}.$

(i) State the names of $i_1 and\ i_2$

(ii) Obtain the modulation index

(iii) Sketch a well labeled waveforms of i_1, i_2 and the resulting amplitude modulation signal.

(iv) Sketch a frequency spectral of i_1, i_2 and the resulting amplitude modulation signal.

(v) Is it right to replace the first signal with $i_1 = 50\ Cos(2\pi f_1\ t)$ if no, explain, state the consequence (if any).

(vi) Suppose the first signal is place with $i_1 =\ 40 + 30cos(2\pi f_1 t) + 20cos(20\pi f_1 t)$, Sketch the frequency spectrum of the resulting amplitude modulated signal.

(c) (i) What are the advantages (if any) of SSB over full AM?

(ii) Draw block diagrams to illustrate the production of DSBSC and SSB by use of balanced modulators (Don't make use of any filter).

(iii) Draw the basic Envelope Detector Circuit.

(University of Ibadan, TEL412- Communication system I 2009/20010 BSc degree Exam).

3.12 (a) (i) What are the advantages (if any) of SSB over full AM?

(ii) Draw block diagrams to illustrate the production of DSBSC and SSB by use of balanced modulators (Don't make use of any filter)

(iii) Draw the basic Envelope Detector Circuit.

(iv) Derive an expression for output of an AM signal.

(b) (i) An AM voltage signal consist of a carrier wave $100\cos 2\pi f_c t$ and a baseband signal of $20\cos 2\pi f_m t + 50\cos 18\pi f_m t$. Determine the expression for the output of the AM signal.

(ii) if f_c = 10 MHz, f_m = 200 kHz, draw the spectrum of the modulated message

(University of Ibadan, TEL412- Communication system I 2009/20010 BSc degree Exam).

3.13 The signal $i_1 = 30\cos(2\pi f_1 t)$ is used to amplitude modulate the signal $i_2 = 40cos\ (2\pi f_2 t)$ $\quad f_1 = 3\text{ kHz}\ \&\ f_2 = 3\text{ MHz}.$

(i) State the names of $i_1\ and\ i_2$

(ii) Obtain the modulation index

(iii) Sketch a well labeled waveform of i_1, i_2 and the resulting amplitude modulation signal.

(iv) Sketch a frequency spectral of i_1, i_2 and the resulting amplitude modulation signal.

(v) Is it right to replace the first signal with $i_1 = 41cos(2\pi f_1\ t)$ if no, explain, state the consequence (if any)

(vi) Suppose the first signal is place with i_1 = 20 + 40cos($2\pi f_1 t) + 10\ cos(20\pi f_1 t)$

Sketch the frequency spectrum of the resulting amplitude modulated signal.

(University of Ibadan, TEL412- Communication system I 2003/2004 BSc degree Exam).

3.14 (a) The signal $v_1 = 37cos(2\pi f_1 t)$ and v_2 = 64cos ($2\pi f_2 t$) are inputs to an amplitude modulator. $f_1 = 15000\text{ Hz}\ \&\ f_2 = 50\text{ MHz}.$

(i) State the names of $v_1\ and\ v_2$

(ii) Obtain the modulation index

(iv) Sketch a frequency spectral of v_1, v_2 and the resulting amplitude modulated signal.

(v) Is it right to replace the first signal with v_1 = 65cos ($2\pi f_1 t$) if no, explain, explain consequence (if any)

(vi) Suppose the first signal is place with $v_1 = 42 + 43cos(2\pi f_1\ t) + 44cos\ (20\pi f_1\ t)$

Sketch the frequency spectrum of the resulting amplitude modulated signal.

(b) With the aid of phase diagrams, briefly compare and contrast NBFM and AM.

(University of Ibadan, TEL412- Communication system I 2005/2006 BSc degree Exam).

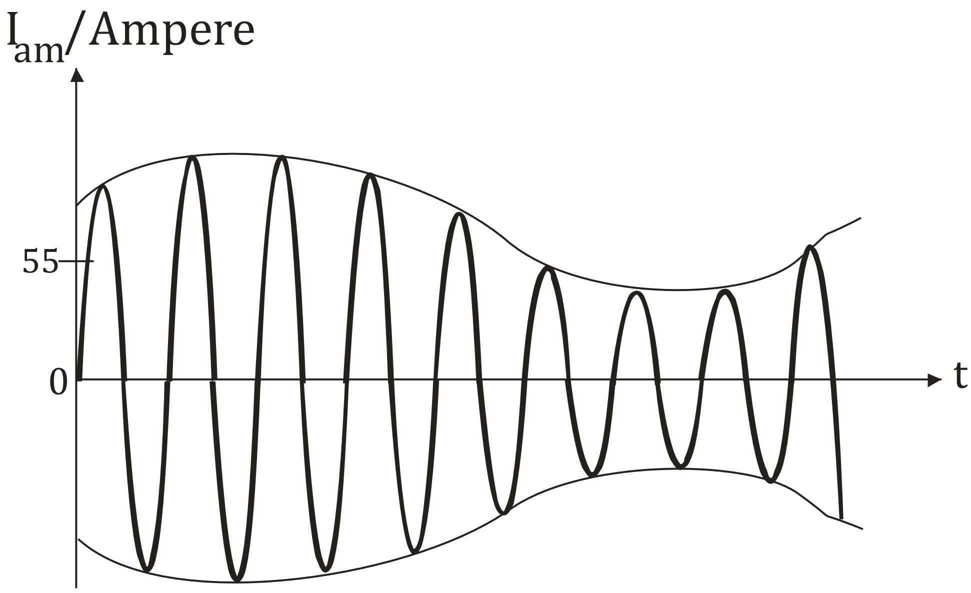


Fig Q5

3.15 Fig. Q5 shows an amplitude-modulated wave. The baseband and carrier are Sinusoidal signals with 20 kHz and 2 MHz frequency respectively. Suppose the modulated index is 0.8, Obtain:

(i) I_m and I_c Hence sketch the frequency spectrum.

(ii) Ratio of power in the side bands to that in the carrier.

(iii) The value of I_c to change modulation index to 50% without change in I_m.

(University of Ibadan, TEL412- Communication system I 2010/2011 BSc degree Exam).

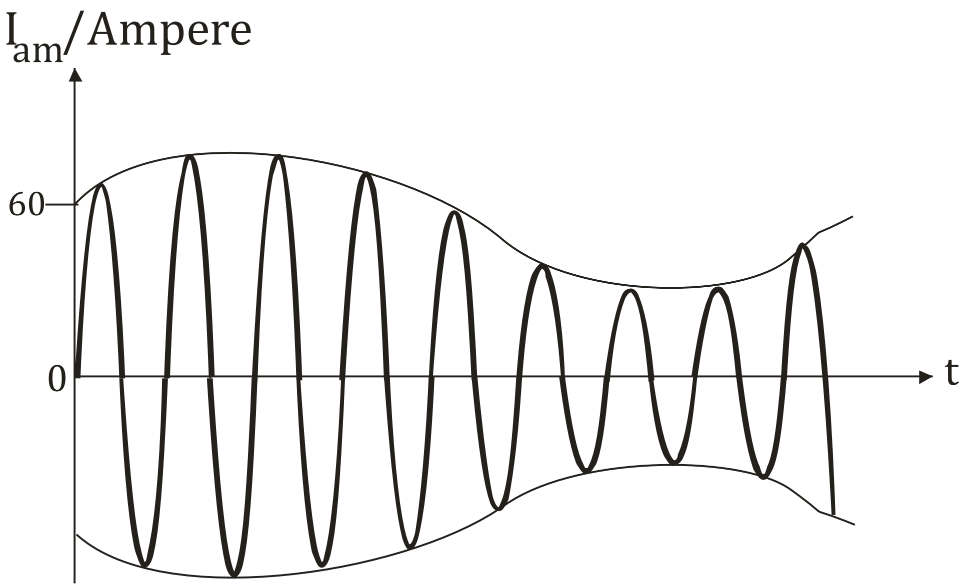


Fig Q6

3.16 Fig.Q6 shows an amplitude-modulated wave. The baseband and carrier are Sinusoidal signals with 40 kHz and 4 MHz frequency respectively. Suppose the modulated index is 0.6, Obtain:

(i) I_m and I_c Hence sketch the frequency spectrum.

(ii) Ratio of power in the side bands to that in the carrier.

(iii) The value of I_c to change modulation index to 60% without change in I_m.

(University of Ibadan, TEL412- Communication system I 2001/2002 BSc degree Exam).

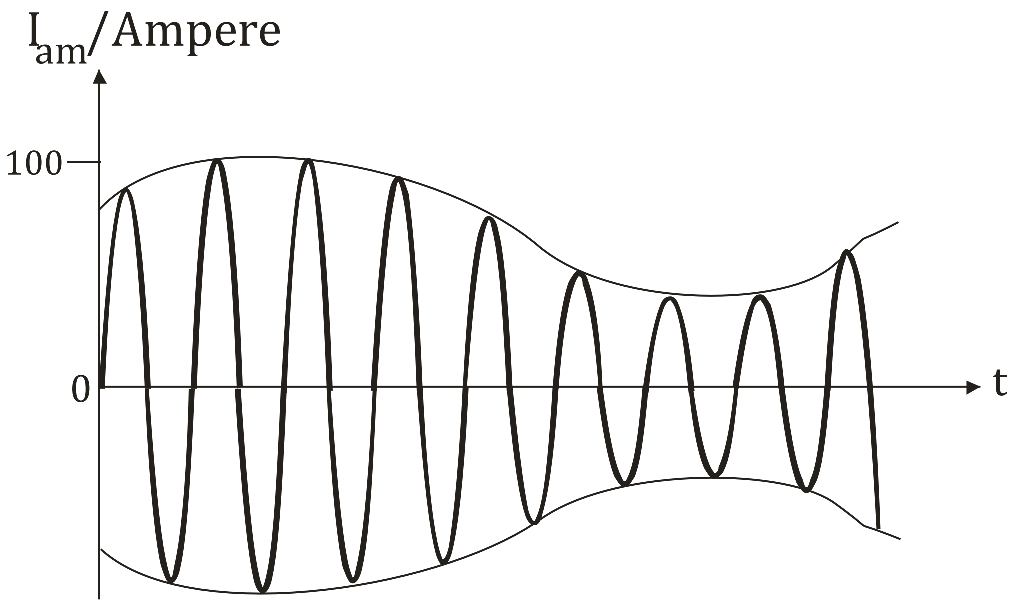


Fig Q7

3.17 Fig. Q7 shows an amplitude-modulated wave. The baseband and carrier are Sinusoidal signals with 40 kHz and 4 MHz frequency respectively. Suppose the modulated index is 0.6, Obtain:

(i) I_m and I_c Hence sketch the frequency spectrum.

(ii) Ratio of power in the side bands to that in the carrier.

(iii) The value of I_m to change modulation index to 60% without change in I_c.

(University of Ibadan, TEL412- Communication system I 2004/2005 BSc degree Exam).

CHAPTER 4

FREQUENCY MODULATION THEORY

4.0 Introduction

In Chapter 3 (Section 3.6) we highlight the general expression for a carrier waveform

$$v_c = E_c \cos \theta_c = E_c \cos (\omega_c t + \emptyset_c)$$

ANGLE MOD AM FM PM

leads to the three types of analogue modulation.

Changes to in the E_c lead to amplitude modulation which has already been dealt with. The other two types occur when θ_c is altered by a baseband and come under the general heading of angle modulation. Changing ω_c produces frequency modulation whereas changes in $\emptyset_c$ cause phase modulation. These last two often seem very similar in action (after all, frequency is merely rate of change of phase) and have similar analytical treatments, although their effect is quite different in practice (and in application). This difference is seen clearly in Fig. 3.1 for a square wave baseband but is much harder to see for a continuously varying baseband. Because they are actually very different, FM is dealt with here and PM is treated separately in Chapter 5.

4.1 Frequency Modulation Principles

In FM, the carrier amplitude remains constant, while the carrier frequency is changed by the modulating signal. As the amplitude of the information signal varies, the carrier frequency shifts in proportion. As the modulating signal amplitude increases, the carrier frequency increases. If the amplitude of the modulating signal decreases, the carrier frequency decreases. The reverse relationship can also be implemented. A decreasing modulating signal will increase the carrier frequency above its center value, whereas an increasing modulating signal will decrease the carrier frequency below its center value. As the modulating signal amplitude varies, the carrier frequency varies above and below its normal center frequency with no modulation. The amount of change in carrier frequency produced by the modulating signal is known as the frequency deviation. Maximum frequency deviation occurs at the maximum amplitude of the modulating signal.

The frequency of the modulating signal determines how many times per second the carrier frequency deviates above and below its nominal center frequency. If the modulating signal is a 100 Hz sine wave, then the carrier frequency will shift above and below the center frequency 100 times per second. This is called the frequency deviation rate.

An FM signal is illustrated in Fig. 4.1(c). Normally, the carrier as shown in Fig. 4.1(a) is a sine wave, but it is shown as a triangular wave here to simplify the illustration. With no modulating signal applied, the carrier frequency is a constant-amplitude sine wave at its normal constant center frequency.

The modulating information signal Fig. 4.1(b) is a low-frequency sine wave. As the sine wave goes positive, the frequency of the carrier increases proportionately. The highest frequency occurs at the peak amplitude of the modulating signal. As the modulating signal amplitude decreases, the carrier frequency decreases. When the modulating signal is at zero amplitude, the carrier will be at its center frequency point.

When the modulating signal goes negative, the carrier frequency will decrease. The carrier frequency will continue to decrease until the peak of the negative half cycle of the modulating sine wave is reached. Then, as the modulating signal increases toward zero, the frequency will again increase. Note in Fig. 4.1 (c) how the carrier sine waves seem to be first "compressed" and then "stretched" by the modulating signal.

Assume a carrier frequency of 50 MHz. If the peak amplitude of the modulating signal causes a maximum frequency shift of 200 kHz,

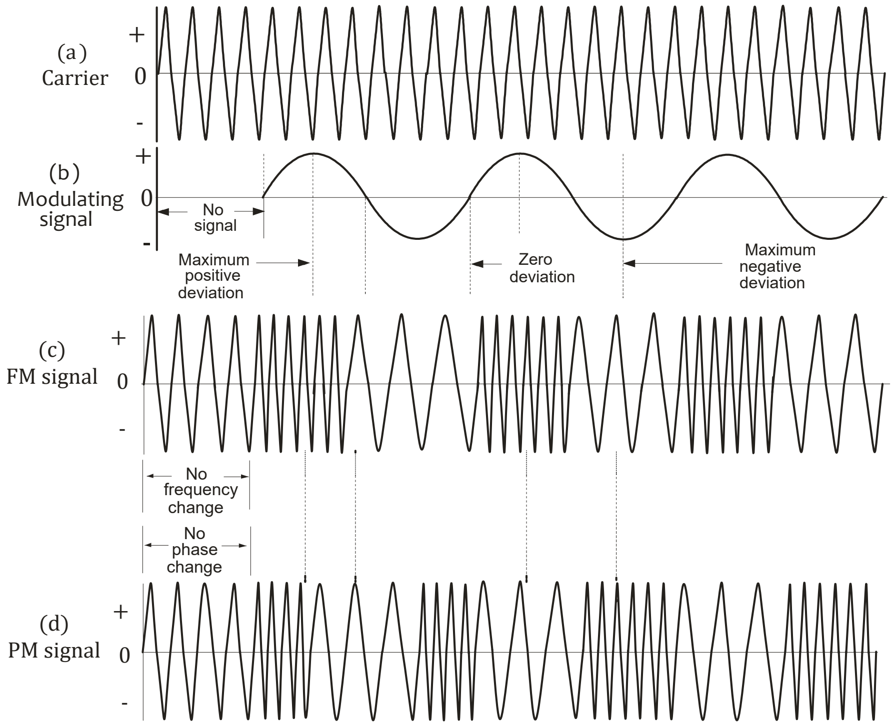


Figure 4.1. Frequency modulation and phase modulation signals the carrier is drawn as a triangular wave for simplicity, but in practice it is a sine wave.

4.2 How FM and PM Differ

It is important to point out that it is the dynamic nature of the modulating signal that causes the frequency variation at the output of the phase shifter. In other words, FM is produced only as long as the phase shift is being varied. One way to understand this better is to assume a modulating signal like that shown in Fig. 4.2(a). It is a triangular wave whose positive and negative peaks have been clipped off at a fixed amplitude. During time t_0, the signal is zero so the carrier is at its center frequency.

Applying this modulating signal to a frequency modulator will produce the signal shown in Fig. 4.2(b). During the time that the waveform is rising at t_1, the frequency is increasing. During the time that the positive amplitude is constant at t_2, the FM output frequency is constant. During the time the amplitude decreases and goes negative at t_3, the frequency will decrease. Then, during the constant-amplitude negative alternation at t_4, the frequency remains constant at a lower frequency. During t_5, the frequency increases.

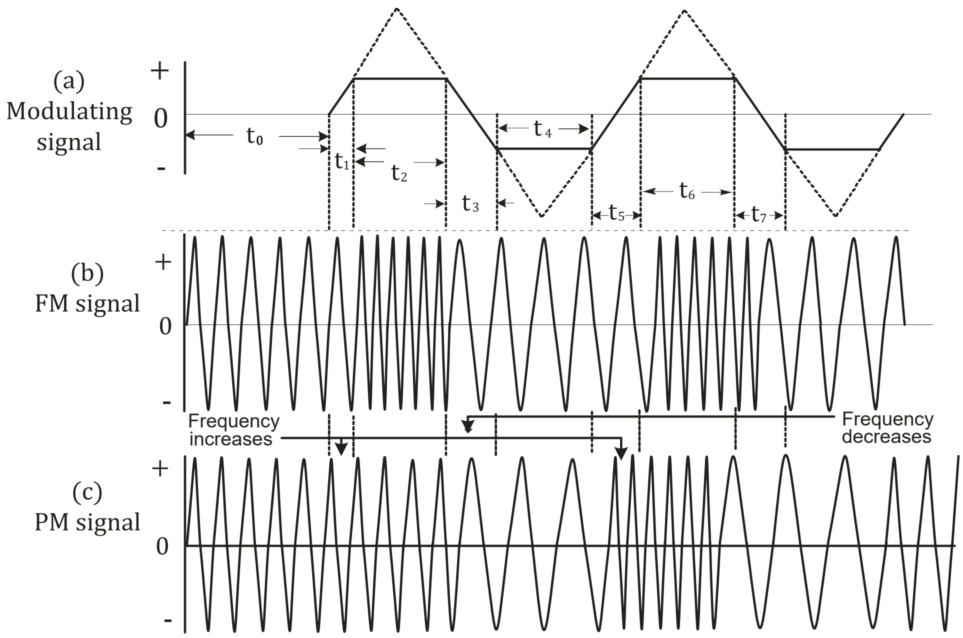


Figure 4.2. A frequency shift occurs in PM only when the modulating signal amplitude varies.

When the modulating signal is applied to a phase shifter, the output frequency will change only during the time that the amplitude of the modulating signal is varying. Refer to the PM signal in Fig. 4.2(c). During increases or decreases in amplitude at t_1, t_3, and t_5, a varying frequency will be produced. However, during the constant-amplitude positive and negative peaks no frequency change takes place. The output of the phase shifter will simply be the carrier frequency which has been shifted in phase. This clearly illustrates that frequency variations take place only if the modulating signal amplitude is varying.

As it turns out, the maximum frequency deviation produced by a phase modulator occurs during the time that the modulating signal is changing at its most rapid rate. For a sine wave modulating signal, the rate of change of the modulating signal is greatest when the modulating wave changes from plus to minus or from minus to plus. The maximum rate of change of modulating voltage occurs exactly at the zero crossing points as depicted in Fig. 4.2(c). In contrast, note that in an FM wave the maximum deviation occurs at the peak positive and negative amplitude of the modulating voltage. So, although a phase modulator does indeed produce FM, maximum deviation occurs at different points of a modulating signal. Of course, this is irrelevant since both the FM and the PM waves contain exactly the same information and, when demodulated, will reproduce the original modulating signal.

In FM, maximum deviation occurs at the peak positive and negative amplitudes of the modulating signal. In PM, the maximum amount of leading or lagging phase shift occurs at the zero crossings of the modulating signal. Recall that we said that the frequency deviation at the output of the phase shifter depends upon the rate of change of the modulating signal. The faster the modulating signal voltage varies, the greater the frequency deviation produced. For the reason of this, the frequency deviation produced in PM increases with the frequency of the modulating signal. The higher the modulating signals frequency, the shorter its period and the faster the voltage changes. Higher modulating voltages produce greater phase shift which in turn produces greater frequency deviation. However, higher modulating frequencies produce a faster rate of change of the modulating voltage, and therefore also, produce greater frequency deviation. In PM then, the carrier frequency deviation is proportional to both the modulating frequency and the amplitude.

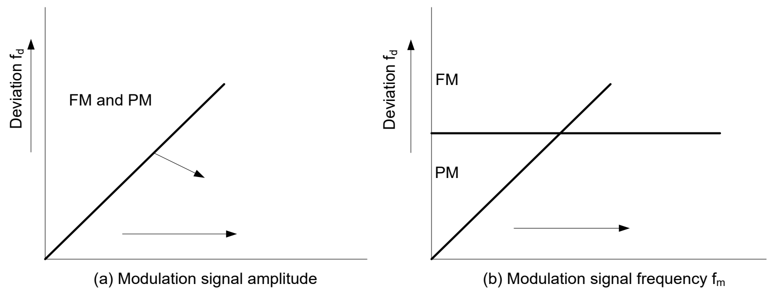


Figure 4.3. Relationship between frequency deviation, modulating signal amplitude and frequency for FM and PM

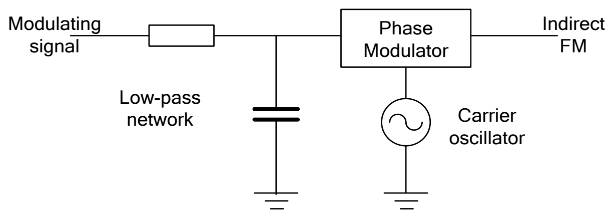


Figure 4.4. Production of indirect FM using low pass filter

In FM, frequency deviation is proportional only to the amplitude of the modulating signal regardless of its frequency. A low-pass filter compensates for higher phase shift and frequency deviation at the higher modulating frequencies to produce indirect FM. The relationship between carrier deviation and modulating signal characteristics is summarized with illustration of Fig. 4.3.

4.3 Common FM Applications

1. FM radio broadcasting
2. TV sound broadcasting
3. Two-way mobile radio
4. Police, fire, public service
5. Marine
6. Amateur radio
7. Family radio
8. Cellular telephone (analog phones)
9. Digital data transmission

4.4 Sidebands and the Modulation Index

Every modulation process commonly produces sidebands. As we saw in AM, when a constant-frequency sine wave modulates a carrier, two side frequencies are produced. The side frequencies are the sum and difference of the carrier and the modulating frequency. In FM and PM, sum and difference sideband frequencies are produced. In addition, a theoretically infinite number of pairs of upper and lower sidebands are also generated. As a result, the spectrum of an FM/PM signal is usually wider than an equivalent AM signal. A special narrowband FM signal whose bandwidth is only slightly wider than that of an AM signal can also be generated.

Fig. 4.5 shows an example of the spectrum of a typical FM signal produced by modulating a carrier with a single-frequency sine wave. Note that the sidebands are spaced from the carrier f_c and are spaced from one another by a frequency equal to the modulating frequency f_m. If the modulating frequency is 500 Hz, the first pair of sidebands are above and below the carrier by 500 Hz. The second pair of sidebands are above and below the carrier by 2 X 500 Hz = 1000 Hz, or 1 kHz, and so on. Note also that the amplitudes of the sidebands vary. If each sideband is assumed to be a sine wave with a frequency and amplitude as indicated in Fig. 4.5 and all these sine waves were added together, then the FM signal producing them would be created.

As the amplitude of the modulating signal varies, of course, the frequency deviation will change. The number of sidebands produced, their amplitude, and their spacing depend upon the frequency deviation and modulating frequency. Keep in mind that an FM signal has constant amplitude. If an FM signal is a summation of the sideband frequencies, then we can see that the sideband amplitudes must vary with frequency deviation and modulating frequency if their sum is to produce a fixed-amplitude FM signal.

Although the FM process produces an infinite number of upper and lower sidebands, only those with the largest amplitudes are significant in carrying the information. Typically any sideband with an amplitude less than 1 percent of the unmodulated carrier is considered insignificant. As a result, this markedly narrows the bandwidth of an FM signal.

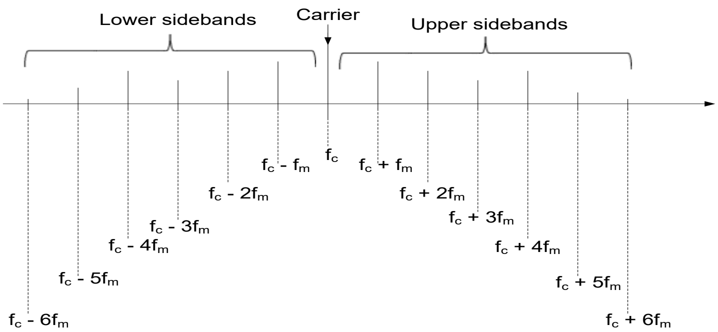


Figure 4.5. Frequency spectrum of an FM signal

4.5 Percentage of Modulation

In AM, the amount or degree of modulation is usually given as a. percentage of modulation. The percentage of modulation is the ratio of the amplitude of the modulating signal to the amplitude of the carrier. When the two factors are equal, the ratio is 1 and we say that 100 percent modulation occurs. If the modulating signal amplitude becomes greater than the carrier amplitude, then over-modulation and distortion occur.

Such conditions do not exist with FM or PM. Since the carrier amplitude remains constant during modulating with FM and PM, the percentage-of-modulation indicator used in AM has no meaning. Further, increasing the amplitude or the frequency of the modulating signal will not cause over-modulation or distortion. Increasing the modulating signal amplitude simply increases the frequency deviation. This, in turn, increases the modulation index, which simply produces more significant sidebands and a wider bandwidth. For practical reasons of spectrum conservation and receiver performance, there is usually some limit put on the upper frequency deviation and the upper modulating frequency. As indicated earlier, the ratio of the maximum frequency deviation permitted to the maximum modulating frequency is referred to as the deviation ratio.

The audio in TV broadcast is transmitted by FM. The maximum deviation permitted is 25 kHz, and the maximum modulating frequency is 15 kHz. This produces a deviation ratio of

$$\mathrm{d} = \frac{25}{15} = 1.6666$$

In standard two-way mobile radio communications using FM, the maximum permitted deviation is usually 5 kHz. The upper modulating frequency is usually limited to 2.5 kHz, which is high enough for intelligible voice transmission. This produces a deviation ratio of

$$\mathrm{d} = \frac{5}{2.5} = 2$$

The maximum deviation permitted can be used in a ratio with the actual carrier deviation to produce a percentage of modulation for FM. Remember, in commercial FM broadcasting the maximum allowed deviation is 75 kHz. If the modulating signal is producing only a maximum deviation of 60 kHz, then the FM percentage of modulation is

$$\text{FM percent modulation} = \frac{\text{actual carrier deviation}}{\text{maximum carrier deviation}} \text{x}100 = \frac{60}{75} \text{x}100 = 80\%$$

When maximum deviations are specified, it is important that the percentage of modulation be held to less than 100 percent. The reason for this is that FM stations operate in assigned frequency channels. Every FM station channel is adjacent to other channels containing other stations. If the deviation is allowed to exceed the maximum, a greater number of pairs of sidebands will be produced and the signal bandwidth may be excessive. This can cause undesirable adjacent channel interference.

4.6 Frequency Modulation vs. Amplitude Modulation

In general, FM is considered to be superior to AM. Although both modulation types are suitable for transmitting information from one place to another, both are capable of equivalent fidelity and intelligibility. FM typically offers some significant benefits over AM, as follows:

A) FM offers better noise immunity
B) It rejects interfering signals because of the capture effect.
C) It provides better transmitter efficiency.

Its disadvantage lies in the fact that it uses an excessive amount of spectrum space. However, let's consider each of these points in more detail.

a) Noise Immunity

The primary benefit of FM over AM is its superior noise immunity. Noise is interference to a signal generated by lightning, motors, automotive ignition systems, and any power line switching that produces transients. Such noise is typically narrow spikes of voltage with very broad frequency content. They add to a signal and interference with it. If the noise signals are strong enough, they can completely obliterate the information signal.

Noise is essentially amplitude variations. An FM signal, on the other hand, has constant carrier amplitude. Because of this, FM receiver contain limiter circuits that deliberately restrict the amplitude of the receiver signal.

Any amplitude variations occurring on the FM signal are effectively clipped off. This does not hurl the information content of the FM signal, since it is contained solely within the frequency variations of the carrier. Because of the clipping action of the limiter circuits, noise is almost completely eliminated.

b) Capture Effect

Another major benefit of FM is that interfering signals on the same frequency will be effectively rejected. Because of the limiters built into FM receivers, a peculiar effect takes place when two or more FM signals occur simultaneously on the same frequency. If the signal of one is more than twice the amplitude of the other, the stronger signal will "capture" the channel and will totally eliminate the weaker, interfering signal. This is known as the **capture effect** in FM. When two AM signals occupy the same frequency, both signals will generally be heard regardless of their relative signal strengths. When one AM signal is significantly stronger than the other, naturally the stronger signal will be intelligible. However, although the weaker signal will be unintelligible, it will still be heard in the background. When the signal strengths of the AM signals are nearly the same, they will interfere with one another making both of them nearly unintelligible. In FM, the capture effect allows the stronger signal to dominate while the weaker signal is eliminated.

However, when the strengths of the two FM signals begin to be nearly the same, the capture effect may cause the signals to alternate in their domination of the frequency. At some time one signal will be stronger than the other, and it will capture the channel. At other times, the signal strength will reverse and the other signal will capture the channel. You may have experienced this effect yourself when listening to the FM radio in your car while driving on the highway. You may be listening to a strong station on a particular frequency, but as you drive, you move away from that station. At some point, you may begin to pick up the signal from another station on the same frequency. When the two signals are approximately the same amplitude, you will hear one station dominate and then the other as the signal amplitudes vary during your driving. However, at some point, the stronger signal will eventually dominate. In any case, once the strong signal dominates, the weaker is not heard at all on the channel.

c) Higher transmitter efficiency

Since only assigned transmitter (carrier) signal frequency is modulated, it means that only a fraction of available energy of audio power is required to produce 100% modulation as compared to high power needed in AM transmission.

4.7 Pre-emphasis and De-emphasis

Despite the fact that FM has superior noise rejection qualities, noise still interferes with an FM signal. This is particularly true for the high-frequency components in the modulating signal. Since noise is primarily sharp spikes of energy, it contains a considerable number of harmonics and other high-frequency components. These high frequencies can at times be larger in amplitude than the high-frequency content of the modulating signal. This causes a form of frequency distortion that can make the signal unintelligible.

Most of the content of a modulating signal, particularly voice, is at lower frequencies. In voice communications systems, the bandwidth of the modulating signal is deliberately limited to a maximum of approximately 3 kHz. The voice is still intelligible despite the bandwidth limitations. After all, telephones cut off at 3 kHz and give good voice quality. However, music would be severely distorted by such a narrow bandwidth because it contains high-frequency components necessary to high fidelity. Typically, however, these high-frequency components are of a lower amplitude. For example, musical instruments typically generate their signals at low frequencies but contain many lower level harmonics that give them their unique sound. If their unique sound is to be preserved, then the high-frequency components must be passed. This is the reason for such a wide bandwidth in high-fidelity sound systems. Since these high-frequency components are at a very low level, noise can obliterate them.

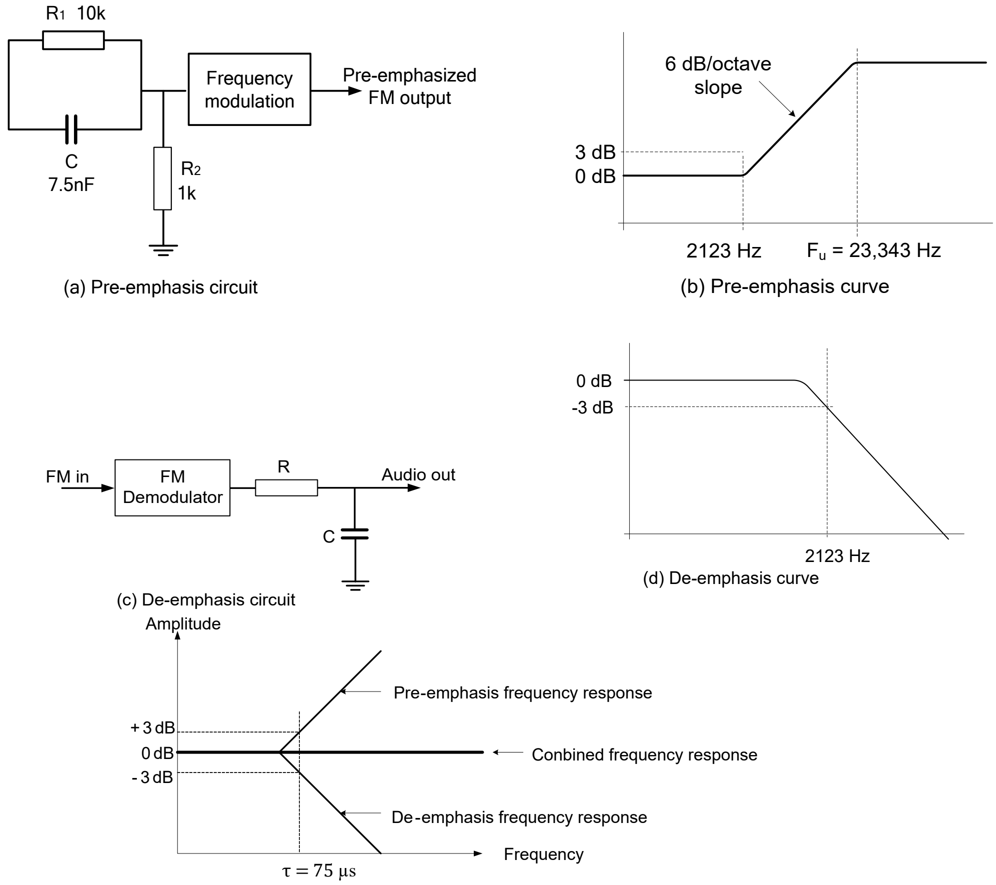


Figure 4.6 Pre-Emphasis and De-Emphasis.

To overcome this problem, most FM systems use a technique known as pre-emphasis, which helps offset high-frequency noise interference. At the transmitter, the modulating signal is passed through a simple network which amplifies the high-frequency components more than the low-frequency components. The simplest form of such a circuit is the high-pass filler of the type shown in Fig. 4.6(a). Specifications dictate a time constant t of 75 microseconds (µs) where τ = R_1C. Any combination of resistor and capacitor (or resistor and inductor) giving this time constant will be satisfactory. Such a circuit has a lower break frequency f_1 of 2123 Hz. This means that frequencies higher than 2123 Hz; will be linearly enhanced. The output amplitude increases with frequency at a rate of 6 dB per octave. The pre-emphasis curve is shown in Fig. 4.6(b). This pre-emphasis circuit increases the energy content of the higher-frequency signals so that they tend to become stronger than the high-frequency noise components. This improves the signal-to-noise ratio and increases intelligibility and fidelity.

The pre-emphasis circuit also has an upper break frequency f_u, where the signal enhancement flattens out. See Fig. 4.6(b).This upper break frequency is computed with the expression

$$f_u = \frac{R_1 + R_2}{2\pi R_1 C}$$

It is usually set at some high value beyond the audio range. An f_u of greater than 20 kHz is typical.

To return the frequency response to its normal level, a de-emphasis circuit is used at the receiver. This is a simple low-pass filler with a time constant of 75 µs. See Fig. 4.6(c). It features a cut-off 2123 Hz and causes signals above this frequency to be attenuated at the rate of 6 dB per octave. The response curve is shown in Fig. 4.6(d). As a result, the pre-emphasis at the transmitter is exactly offset by the de-emphasis circuit the receiver, providing a normal frequency response. The combined effect of pre-emphasis and De-emphasis is to increase the high-frequency components during transmission so that they will be stronger and not masked by noise.

4.8 Transmission Efficiency

The third advantage of FM over AM is in trans-lining efficiency. Recall that AM can be pro-need by both low-level and high-level modulation techniques. The most effective is high-level modulation, in which a class C amplifier is used as the final RF power stage and is modulated by a high-power modulation amplifier. The AM transmitter must produce both very high RF and modulating signal power. In addition, at very high power levels large modulation amplifiers are impractical. Under such conditions, low-level modulation must be used.

The AM signal is generated at a lower level and then amplified with linear amplifiers to produce the final RF signal. Because linear amplifiers operate class A or class B, they are far less efficient than class C amplifiers. However, linear amplifiers must be used if the AM information is to be preserved.

An FM signal has a constant amplitude, and, therefore, it is not necessary to use linear amplifiers to increase its power level. In fact, FM signals are always generated at a lower level and then amplified by a series of class C amplifiers to increase their power. The result is greater use of available power because class C amplifiers are far more efficient.

Disadvantage of FM

Perhaps the greatest disadvantage of FM is that it simply uses too much spectrum space. The bandwidth of an FM signal is considerably wider than that of an AM signal transmitting similar information. Although the modulation index can be kept low to minimize the bandwidth used, still the bandwidth is typically larger than that of an AM signal. Further, reducing the modulation index also reduces the noise immunity of the FM signal. In commercial two-way FM radio systems, the maximum allowed deviation is 5 kHz, with a maximum modulating frequency of 3 kHz. This produces a deviation ratio of 5/3 - 1.67. This is usually referred to as narrowband FM (NBFM).

Since FM occupies so much bandwidth, it typically has been used only at the very high frequencies. In fact, it is seldom used in communications below frequencies of 30 MHz. Most FM communications work is done at the VHP, UHF, and microwave frequencies. It is only in these portions of the spectrum where adequate bandwidth is available for FM signals and where line-of-sight transmission is prevalent. This means that the range of communication is more limited.

Binary Signals

FM is also used to transmit digital data. When the modulating signal is binary, a binary 1 input produces one earner frequency, and a binary 0 input produces another carrier frequency. This modulation technique is known as frequency-shift keying (FSK).

Fig. 4.7 shows an FSK signal, A binary 0 input produces a 1070 Hz carrier, and a binary 1 produces a 1270 Hz carrier.

At one time, FSK was used in computer modems that transmit data through the telephone system. Today, FSK is used primarily to transmit binary data by radio.

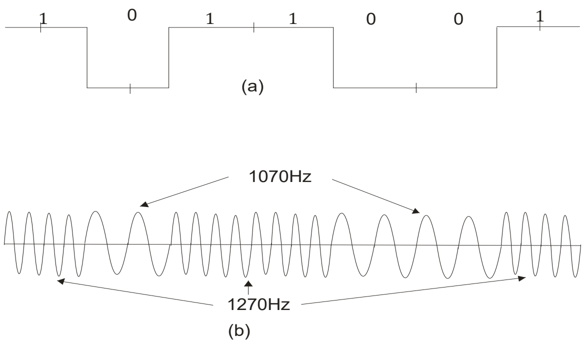


Figure 4.7 Frequency-shift keying, (a) Binary signal, (b) FSK signal.

4.9 Frequency Modulators

The basic concept of FM is to vary the carrier frequency in accordance with the modulating signal. The carrier is generated by either an LC or a crystal oscillator circuit.

The object then isto find a way to change the frequency of oscillation. In an LC oscillator, the carrier frequency is fixed by the values of the inductance or capacitance in a tuned circuit. The carrier frequency, therefore, may be changed by varying either this inductance or the capacitance.The idea is to find a circuit or component that converts a modulating voltage into a corresponding change in capacitance or inductance.
When the carrier is generated by a crystal oscillator, the frequency is fixed by the crystal.

However, keep in mind that the equivalent circuit of a crystal is an LCR circuit with both series and parallel resonant points. By connecting an external capacitor to the crystal, minor variations in operating frequency can be obtained. Again, the objective is to find a circuit or component whose capacitance will change in response to the modulating signal.

4.10 Voltage-Variable Capacitor

The component most frequently used in this application is a varactor or voltage-variable capacitor (VVC). Also known as a variable capacitance diode, or varicap, this component is basically a semiconductor junction diode that is operated in a reverse-bias mode.

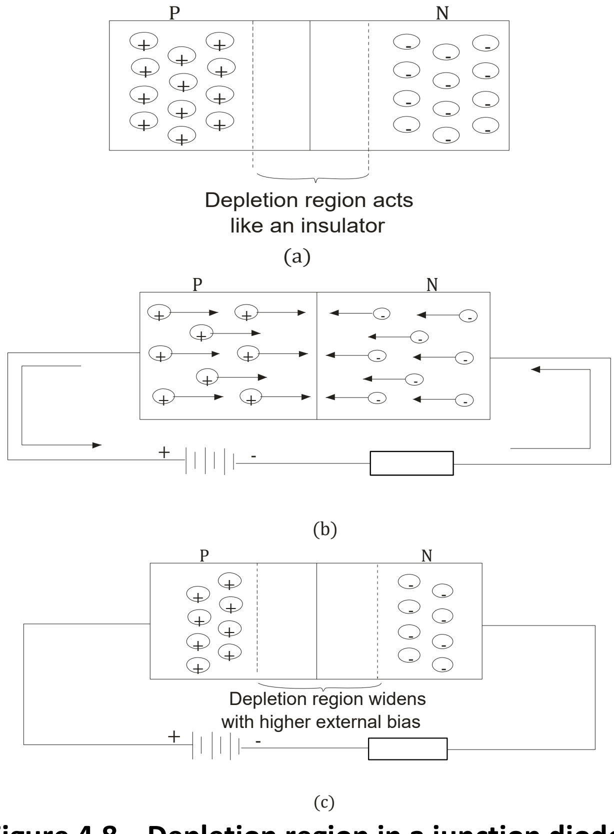


Figure 4.8 Depletion region in a junction diode.

Refer to Fig. 4.8. When a junction diode is formed, P- and N-type semiconductors are joined to form a junction. Some electrons in the N-type material drift over into the P-type material and neutralize the holes there. Thus a thin region where there are no free carriers, holes, or electrons is formed. This is called the depletion region It acts like a thin insulator that prevents current from flowing through the device. See Fig. 4.8 (a).

If you apply a forward bias to the diode, it will conduct. The external potential forces the holes and electrons toward the junction, where they combine and cause a continuous current inside the diode as well as externally.
The depiction layer simply disappears. See Fig. 4.8 (b).

If an external reverse bias is applied to the diode, as in Fig. 4.8(c), no current will flow. The bias actually increases the width of the depletion layer. The width of this depletion mode depends upon the amount of the reverse bias. The higher the reverse bias, the wider the depletion layer and the less chance for current flow.

A reverse-biased junction diode appears to be a small capacitor. The P- and N-type materials act as the two plates of the capacitor, and the depletion region acts as the dielectric. The width of the depletion layer determines the width of the dielectric and, therefore, the amount of capacitance. If the reverse bias is high, the depletion region will be wide and the dielectric will cause the plates of the capacitor to be widely spaced, producing a low value of capacitance. Decreasing the amount of reverse bias narrows the depletion region and, therefore, the plates of the capacitor will be effectively closer together and produce a higher capacitance.

All junction diodes exhibit this characteristic of variable capacitance as the reverse bias is changed. However, varactors or VVCs have been designed to optimize this particular characteristic. Such diodes are made so that the capacitance variations are as wide and linear as possible.

4.11 Varactor Modulator

Fig. 4.9 shows the basic concept of a varactor frequency modulator. The L_1 and C_1 represent the tuned circuit of the carrier oscillator. Varactor diode D_1 is connected in series with capacitor C_2 across the tuned circuit. The value of C_2 is made very large at the operating frequency so that its reactance is very low. As a result, when C_2 is connected in series with the lower capacitance of D_1. The effect is as if D_1 were connected directly across the tuned circle total effective circuit capacitance then capacitance of D_1in parallel with C_1. This fixes the center carrier frequency

The capacitance of D_1, of course, is controlled by two factors: a fixed dc bias and the modulating signal. In Fig. 4.9, the bias on D_1 at the voltage divider which is made up of and $R_1 + R_2$ Usually either R_1 or R_2; is made variable so that the center carrier frequency can be adjusted over a narrow range. The modulating signal is applied through C_3 and the RFC. The C_3 is a blocking capacitor that keeps the dc bias out of the modulating signal circuit. The RFC is a radio frequency choke whose reactance is high at the carrier frequency to prevent the carrier signal from getting into modulating signal circuits.

The modulating signal derived from the mi-lone is amplified and applied to the modulator. As the modulating signal varies, it adds to subtracts from the fixed-bias voltage.

Thus the effective voltage applied to D_1 causes its capacitance to vary. This, in turn, produces a deviation of the carrier frequency as desired. A positive-going signal at point A adds to the reverse bias, decreasing the capacitance and easing the carrier frequency.

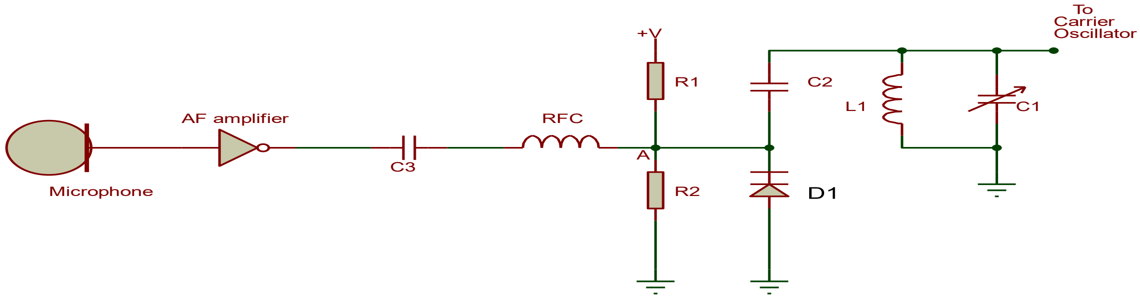


Fig 4.9 Frequency modulation with a VVC.

A negative-going signal at A subtracts from the bias, increasing the capacitance and decreasing the carrier frequency.

The main problem with the circuit in Fig. 4.9 is that most LC oscillators are simply not stable enough to provide a carrier signal. Despite the quality of the components and the excellence of design, the LC oscillator frequency will vary because of temperature changes, circuit voltage variations, and other factors. Such instabilities cannot be tolerated in most modern electronic Communication systems. As a result, crystal oscillators are normally used to set the carrier frequency. Not only do crystal oscillators provide a highly accurate carrier frequency, but their frequency stability is superior over a wide temperature range.

4.12 Frequency-Modulating a Crystal Oscillator

It is possible to vary the frequency of a crystal oscillator by changing the value of capacitance in series or in parallel with the crystal; Fig.4.10 shows a typical crystal oscillator. When a small value of capacitance is connected in series with the crystal, the crystal frequency can be "pulled" slightly from its natural resonant frequency. By making the series capacitor a varactor diode, frequency modulation of the crystal oscillator can be achieved. The modulating signal is applied to the varactor diode D_1 which changes the oscillator frequency.

The important thing to note about an FM crystal oscillator is that only a very small frequency deviation is possible. Greater deviations can be achieved with LC oscillators. Rarely can the frequency of a crystal oscillator be changed more than several hundred hertz from the nominal crystal value. The resulting deviation may be less than the total deviation

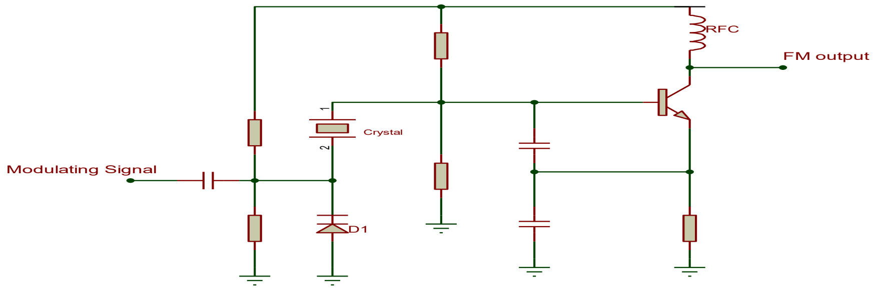


Figure 4.10 Frequency Modulation of A Crystal Oscillator With A VVC

desired. To achieve a total frequency shift of 75 kHz as in commercial FM broadcasting, other techniques must be used. In two-way (narrowband) FM communications systems, the narrower deviations are acceptable.

Although a deviation of only several hundred cycles may be possible at the crystal oscillator frequency, the total deviation can be increased by using frequency multiplier circuits after the carrier oscillator. When the FM signal is applied to & frequency multiplier, both the frequency of operation and the amount of

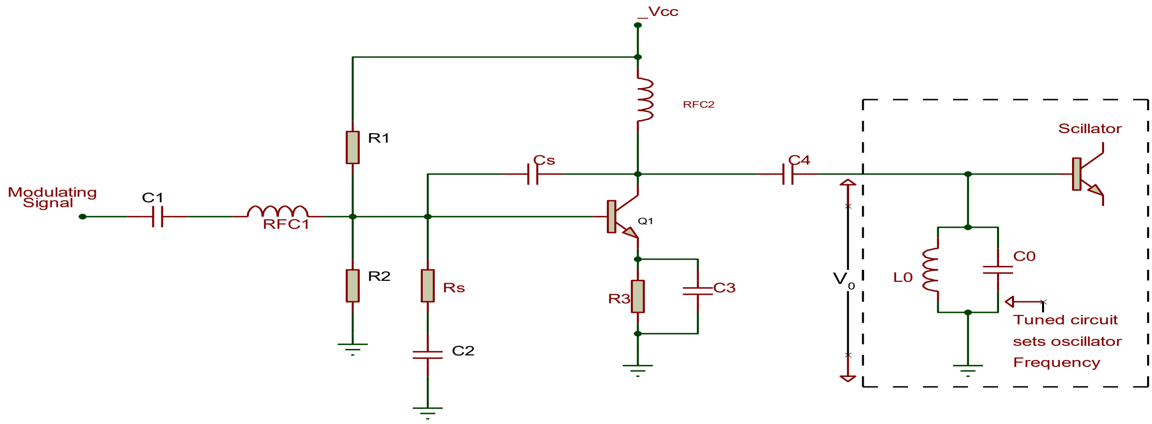


Figure 4.11 A reactance modulator.

deviation is increase. Typical frequency multipliers may increase the base (oscillator) frequency by 24 to 32 times.

For example, assume that the desired output frequency from an FM transmitter is 168 MHz and that the carrier is generated by a 7 MHz crystal oscillator. This is followed by frequency multiplier circuits that increase the frequency by a factor of 24 (7 MH x 24 = 168 MHz)

Assume further that the desired maximum frequency deviation is 5 kHz. Frequency modulation of the crystal oscillator may produce a maximum deviation of only 200 Hz. When multiplied by the factor of 24 in the frequency multiplier circuits, this deviation, of course, is increased to $200 \text{ X } 24 = 4800 \text{ Hz}$, or 4.8 kHz. Frequency multiplier circuits will he discussed in more detail in the chapter on transmitters.

4.13 Reactance Modulator

Another way to produce direct frequency modulation is to use a reactance modulator. This circuit uses a transistor amplifier to act like either a variable capacitor or an inductor. When the circuit is connected across the tuned circuit of an oscillator, the oscillator frequency can be varied by applying the modulating signal to the amplifier.

A reactance modulator is illustrated in Fig. 4.11. It is basically a standard common-emitter class a amplifier. Resistors R_1 and R_2 form a voltage divider to bias the transistor into the linear region. R_3 is an emitter bias resistor which is bypassed with capacitor C_3. Instead of a collector resistor, a radio frequency choke (RFC_2) is used to provide a high impedance load at the operating frequency.

Now, note that the collector of the transistor is connected to the top of the tuned circuit in the oscillator. Capacitor C_4 has very low impedance at the oscillator frequency; its main purpose is to keep the direct current from the collector of Q_t from being shorted to ground through the oscillator coil L_o. As you can see, the reactance modulator circuit is connected directly across the parallel limed circuit that sets the oscillator frequency.

The oscillator signal from the tuned circuit V_o is connected back to an RC phase-shift circuit made up C_s and R_S. Capacitor C_2 in series with R_s, has a very low impedance at the operating frequency, so it does not affect the phase shift. However, it does prevent R_s, from disturbing the dc bias on Q_1. The value of C_s, is chosen so that its reactance at the oscillator frequency is about 10 or more times the value of R_s. If the reactance is much greater than the resistance, the circuit will appear predominantly capacitive; therefore the current through the capacitor and R_S will lead the applied voltage by about 90°. This means that the voltage across R_s, that is applied to the base of Q_1 leads the voltage from the oscillator.

Since the collector current is in phase with the base current, which in turn is in phase with the base voltage, the collector current in Q_1 leads the oscillator voltage V_o by 90°. Of course, any circuit whose current leads its applied voltage by 90° looks capacitive to the source voltage. This means that the reactance modulator looks like a capacitor to the oscillator-tuned circuit.

The modulating signal is applied to the modulator circuit through C_1 and RFC_1. The RFC helps keep the RF signal from the oscillator out of the audio circuits from which the modulating signal usually comes. The audio modulating signal will vary the base voltage and current of Q_1 according to the intelligence to be transmitted. The collector current will also vary in proportion. As the collector current amplitude varies, the phase-shift angle changes with respect to the oscillator voltage, which is interpreted by the oscillator as a change in the capacitance. So as the modulating signal changes, the effective capacitance of the circuits varies and the oscillator frequency is varied accordingly. An increase in capacitance lowers the frequency, whereas a lower capacitance increases the frequency. The circuit produces direct frequency modulations.

If you reverse the positions of R_s and C_s in the circuit of Fig. 4.11, the current in the phase shifter will still lead the oscillator voltage by W. However, it is voltage from across the capacitor that is now applied to the base of the transistor. This voltage lags the oscillator voltage by 90". With this arrangement, the reactance modulator acts like an inductor. The equivalent inductance changes as the modulating signal is applied. Again, the oscillator frequency is varied in proportion to the amplitude of the intelligence signal.

The reactance modulator is one of the best FM circuits because it can produce frequency deviation over a wide frequency range. It is also highly linear; that is, distortion is minimal. The circuit can also be implemented with a field-effect transistor (FET) in place of the NPN bipolar shown in Fig. 4.11.

4.14 Voltage-Controlled Oscillator

Oscillators whose frequencies are controlled by an external input voltage are generally referred to as voltage-controlled oscillators (VCOs). A voltage-controlled crystal oscillator is generally referred to as a VXO. Although VCOs are used primarily for FM, there are other applications in which some form of voltage-to-frequency conversion is required.

In high-frequency communication circuits. VCOs are ordinarily implemented with discrete-component transistor and varactor diode circuits. However, there are many different types of lower-frequency VCOs in common use. These include 1C VCOs using AT multi-vibrator-type oscillators whose frequency can be controlled over a wide range by an ac or dc input voltage. These VCOs typically have an operating range of less than 1 Hz to approximately 1 MHz.

The output is either a square or triangular wave rather than a sine wave. A typical 1C VCO is shown in Fig. 4.12(a). This is a general block diagram of the popular NE566. External resistor R] at pin 6 sets the value of current produced by the internal current sources. The current sources linearly charge and discharge external capacitor C_1 at pin 7. An external voltage V_c applied at pin 5 may also be used to vary the amount of current produced by the current sources.

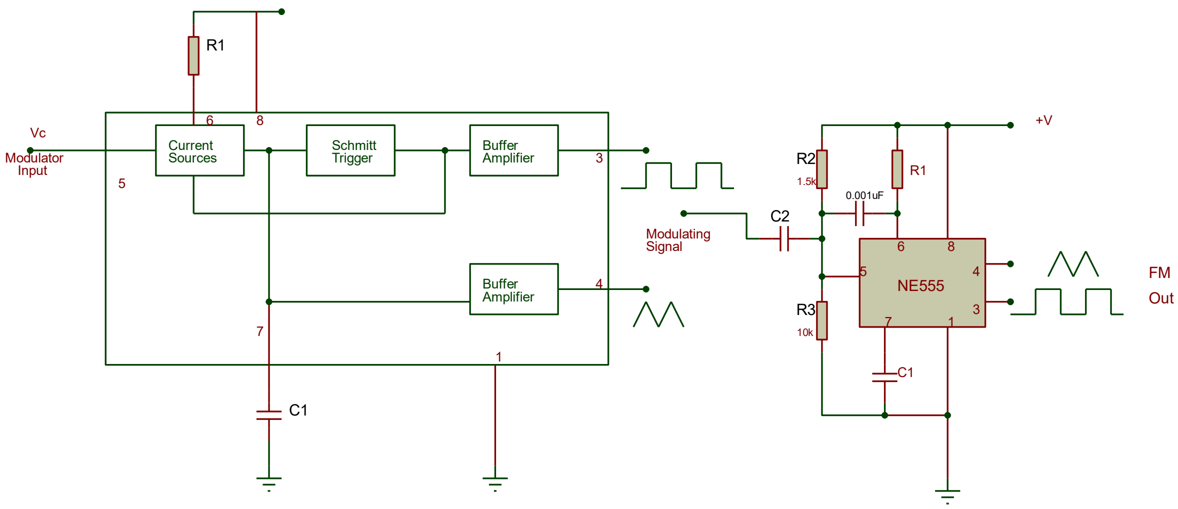


Figure 4.12 Frequency Modulation with an IC VCO

The Schmitt trigger circuit is a level detector which switches when the capacitor charges or discharges to a specific voltage level. The Schmitt trigger controls the current source by switching the current sources between charging and discharging. A linear sawtooth of voltage is developed across the capacitor by the current source. This is buffered and made available at pin 4. The Schmitt trigger output is a square wave at the same frequency available a pin 3. If a sine wave output is desired, usually the triangle wave is filtered with a tuned circuit resonant to the desired carrier frequency.

A complete FM circuit is shown in Fig. 4.12(b). The current sources are biased with a voltage divider made up of R_2 and R_3. The modulating signal is applied through C_2 to the voltage divider at pin 5. The 0.001 µF capacitor between pins 5 and 6 is used to prevent unwanted spurious oscillations.

The center carrier frequency of the circuit is set by the values of R_1 and C_1. Carrier frequencies up to 1 MHz may be used with this 1C. The outputs may be filtered or used to drive other circuits such as a frequency multiplier if higher frequencies and deviations are necessary. The modulating signal can vary the carrier frequency over a nearly 10:1 range, making very large deviations possible. The deviation is linear with respect to the input amplitude over the entire range.

4.15 General basis for FM

The basic principle of frequency modulation is that a baseband voltage (v_m) changes the carrier frequency **linearly** by a **small** amount (i.e. keeping $\delta f_c \ll f_c$). In this way frequency changes carry the information which can then be recovered by the receiver.

Thus the basic equation is

$$\delta\omega_c \ll \omega_c$$

If there are any only one frequency component, i.e.

$$v_m(t) = E_m \cos\omega_m t$$

$\delta\omega_c \propto v_m$	**Basis**
Or	**for**
$\delta\omega_c = K v_m$	**FM**

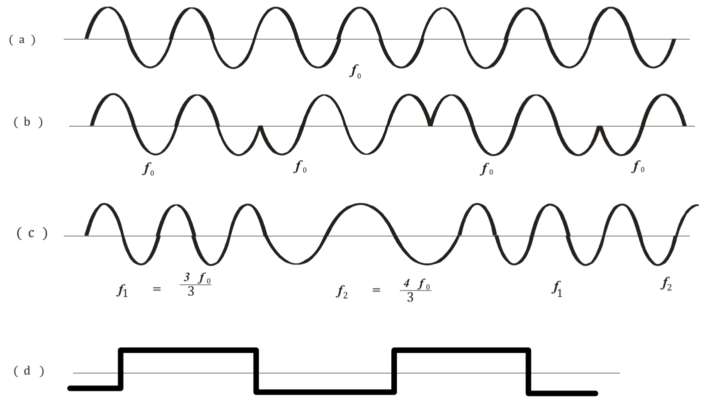


Figure 4.13. Frequency and phase modulation by a square wave baseband: (a) unmodulated carrier waveform, ω_c and $\emptyset_c$ unvarying; (b) modulated by step changes of phase, $\nabla\emptyset_c$ = 180º; (c) modulated by step changes of frequency, f_1= 3f$_c$/4, f_2 = 4f$_c$/3; (d) the square wave baseband.

K is the **modulation sensitivity** having units of rad $s^{-1}\,V^{-1}$, although more likely to be expressed in practice as kHz mV^{-1}. K is often written K to distinguish it from, K_p used for phase modulation.

$\delta\omega$ is the amount by which the baseband has caused the instantaneous carrier frequency; ω_i, of the carrier to **deviate** from the unmodulated frequency; ω_c. That is

$$\omega_i = \omega_c + \delta\omega$$

K = modulation sensitivity (rad/s/vot)

$$\delta\omega_c = KE_m \cos \omega_m t$$

$$= \Delta\omega_c \cos \omega_m t$$

Sine $\Delta\omega_c = KE_m$ (maxim deviation) which occurs at $\cos \omega_m t = \omega_i$

$\omega_i =$ instantaneous

$$\uparrow \omega_i = \ \omega_c t\, \delta\omega_c$$

At first sight you might think you could write simply that

v_{FM} = $E_c\cos \omega_i t$ **WRONG**

But this would not be correct because of the previous history of ω_i which would introduce an unknown additional phase term. Instead you have to write that

$$v_{FM} = E_c \cos \theta_i$$

Where θ_i is the instantaneous phase angle of the carrier, and then go back to the basic idea that frequency is the rate of change of phase so that $\omega_i = d\theta_i/dt$.

The correct expression for θ_i can noz be found by integrating ω_i from time zero to the present instant, t. That is

$$d\theta_i = \int \omega_i dt$$

therefore,

$$\int_0^{\theta_i} d\theta_i = \int_0^t \omega_i \, dt$$

$$\theta_i = \int_0^t (\omega_c + \delta\omega) \, dt$$

$$= \int_0^t (\omega_c + Kv_m) \, dt$$

$$= \int_0^t \omega_c \, dt + K\int v_m \, dt$$

$$\delta\omega_c = KE_m \cos \omega_m t$$

$$= \Delta\omega_c \cos \omega_m t$$

Thus

$$\theta_i = \omega_c t + K\int_0^t v_m dt$$

$$v_{FM} = E_c \cos(\omega_c t + K\int_0^t v_m dt)$$

$$v_{FM} = E_c \cos(\omega_c t + K\int_0^t v_m dt)$$

General expression for FM waveform

This is a completely general expression for any baseband signal frequency modulated onto a carrier.

4.15.1 Frequency modulation by a single sinusoidal baseband

Obviously the expression above does not tell you much. In particular it gives no idea of the spectrum produced in the carrier by the act of frequency modulating it, to see this it is necessary, as for AM, to restrict the analysis to a baseband consisting of a single frequency. This is not as restricting as it see,s because any real signal is made up of a number of such single frequency signals as was shown by Fourier analysis in Chapter 7.

Thus it is both valid and useful to use

$$v_m = E_m \cos \omega_m t$$

Which gives

$$\delta\omega_c = KE_m \cos \omega_m t$$

$$= \Delta\omega_m \cos \omega_m t$$

$$2\pi\Delta f_c = kE_m$$

$$\Delta f_c = \frac{KE_m}{2\pi}$$

Where $\Delta\omega$ and Δf are the **maximum** deviations which occur (when $\cos \omega_m t = \pm 1$). Both quantities are actually referred to merely as the **deviation.** They are important parameters of any f.m. system (e.g. $\pm$ 75 kHz for VHF).

$$\Delta\omega_c = KE_m \quad \text{……DEVIATIONS}$$

$$\Delta f_c = \frac{KE_m}{2\pi}$$

Putting v_m into the general expression for v_{FM} at the end of the last section gives

$$v_{FM} = E_c \cos\left(\omega_c t + K\int_0^t E_m \cos \omega_m t\, dt\right)$$
$$= E_c \cos\left(\omega_c t + KE_m\int_0^t \cos \omega_c t\, dt\right)$$
$$= E_c \cos\left(\omega_c t + \Delta\omega \int_0^t \cos \omega_c t\, dt\right)$$
$$= E_c \cos\left(\omega_c t + \frac{\Delta\omega}{\omega m} \sin \omega_m t\right)$$

Now write the quantity $\Delta\omega/\omega_m$ as a single constant, β, giving

$$v_{FM} = E_c \cos(\omega_c t + \beta \sin \omega_m t)$$

$$\beta = \frac{\Delta\omega_c}{\omega_m} = \frac{\Delta f_c}{f_m} \qquad \rightarrow \qquad (\textbf{modulation index})$$

Modulation index in **f.m** can be greater than 1. By Bessel when $\beta \geq 0.2$ rad, v_{FM} can be reduced to

$$v_{FM} = E_c e^{(j\omega_c t\, \beta \sin \omega_m t)}$$
$$v_{FM)} = E_c[J_o\, \beta \cos \omega_c t + J_1\, \beta[\cos(\omega_c + \omega_m)t - \cos(\omega_c - \omega_m)t] + J_2\, \beta[\cos(\omega_c + 2\omega_m)t + \cos(\omega_c - 2\omega_m)t] + J_3\, \beta[\cos(\omega_c + 3\omega_m)t - \cos(\omega_c - 3\omega_m)t + \cdots]$$

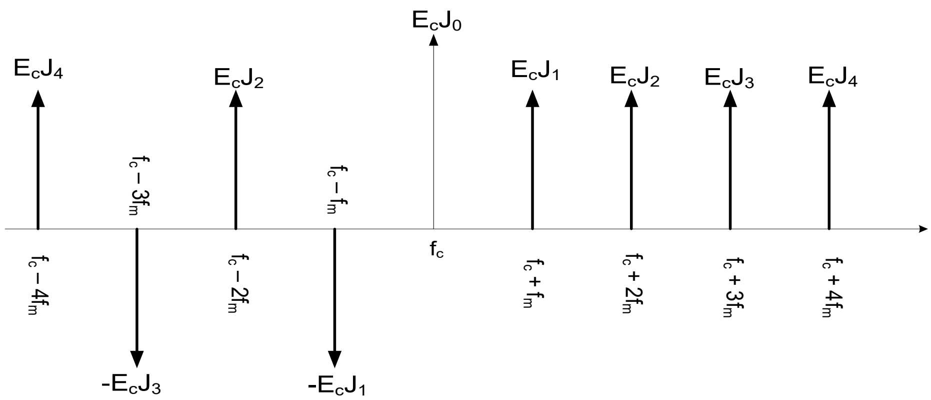


Figure 4.17 Frequency spectrum of the resulting general expression

Example 4.1

A 9 W, 120 MHz higher frequency voltage signal is frequency modulated with a 20 kHz frequency signal such that the peak frequency deviation is 11 kHz. The resistance level is 3 Ω, determine

(i) Amplitude of the high frequency voltage signal.

(ii) The modulation index.
(iii) The Carson's bandwidth.
(iv) The modulation sensitivity if the amplitude of the lower frequency signal is 5.658 V.
(v) State the amplitudes of all components in the resulting frequency modulated signal. (Quote the frequency and the amplitude of each frequency component). Sketch the frequency spectrum of the frequency modulation.
(vi) Write down the resulting frequency modulated equation.

(TEL 412, University of Ibadan 2004/2005 Session Degree Exam).

Solution

(i) $\mathrm{E_c} = ?, \mathrm{P} = 9\ \mathrm{W}, \mathrm{f_c} = 120\ \mathrm{MHz},\ \mathrm{f_m} = 20\ \mathrm{kHz},\ \Delta\mathrm{f_c} = 11\ \mathrm{kHz}, \mathrm{R} = 3\ \Omega$

From Power(P) = $\frac{1}{2}\frac{E_c{}^2}{\mathrm{R}}$, $E_c{}^2 = P \times 2 \times R$

$$\mathrm{E_c}^2 = 9 \times 2\mathrm{R} = 9 \times 2 \times 3 = 54$$

$$\mathrm{E_c} = \sqrt{54} = \mathbf{7.3485\ V}$$

$$\text{(ii) Modulation index } (\beta) = \frac{\Delta\mathrm{f_c}}{\mathrm{f_m}} = \frac{11 \times 10^3}{20 \times 10^3} = 0.55 \approx \mathbf{0.6}$$

$$\text{(iii) Carsons Bandwidth } = 2(1+\beta)\mathrm{f_m} = 2 \times 1.55 \times 20 \times 10^3 = \mathbf{62\ kHz}$$

$$\text{(iv) Modulation sensitivity (K)} = \frac{\Delta\omega_c}{\mathrm{E_m}} = \frac{2\pi \times 11 \times 10^3}{5.658}$$

$$= \mathbf{12.22 \times 10^3\ rad/s\ V}$$

Now using $\beta = 0.6$, refer to the Bessel table for the value of J_n's we obtain the table 4.1 for frequency and amplitude.

Table 4.1

Frequency $\mathbf{f_c}$	**Amplitude**
120.00 MHz	$\mathrm{E_cJ_0} = 7.3485 \times 0.9120 = 6.7018\mathrm{V}$
120.02 MHz	$\mathrm{E_cJ_1} = 7.3485 \times 0.2867 = 2.1068\mathrm{V}$
119.98 MHz	$-\mathrm{E_cJ_1} = -7.3485 \times 0.2867 = -2.1068\mathrm{V}$
120.04 MHz	$\mathrm{E_cJ_2} = 7.3485 \times 0.0437 = 0.3211\mathrm{V}$
119.96 MHz	$\mathrm{E_cJ_2} = 7.3485 \times 0.0437 = 0.3211\mathrm{V}$
120.06 MHz	$\mathrm{E_cJ_3} = 7.3485 \times 0.0044 = 0.0323\mathrm{V}$
119.94 MHz	$-\mathrm{E_cJ_3} = -7.3485 \times 0.0044 = -0.0323\mathrm{V}$
120.08 MHz	$\mathrm{E_cJ_4} = 7.3485 \times 0.0003 = 0.0022\mathrm{V}$
119.92 MHz	$\mathrm{E_cJ_4} = 7.3485 \times 0.0003 = 0.0022\mathrm{V}$

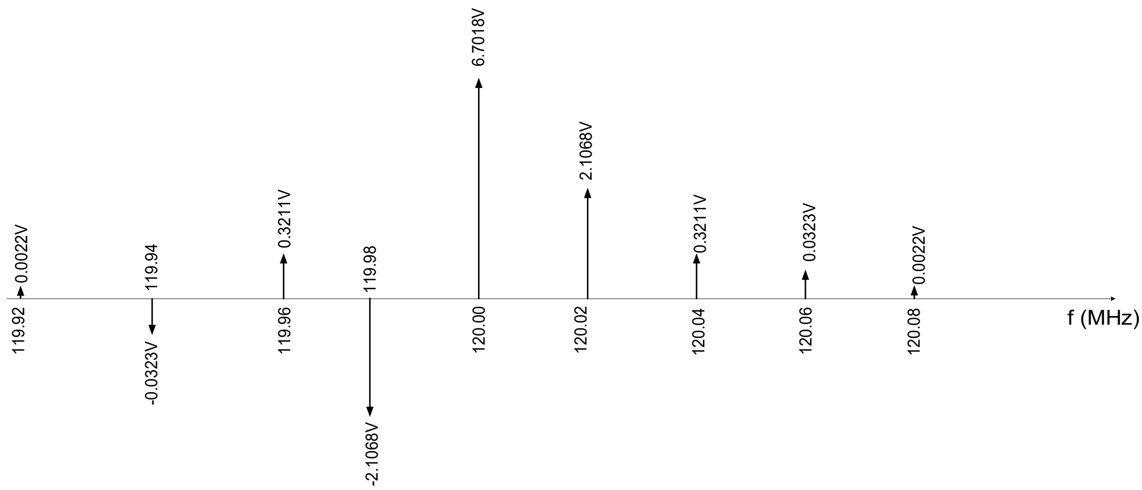


Figure 4.18 Frequency spectrum of the resulting voltage signal

$$v_{FM} = E_cJ_0\beta\cos\omega_ct + E_cJ_1\beta\cos(\omega_c + \omega_m)t - E_cJ_1\beta\cos(\omega_c - \omega_m)t + E_cJ_2\beta\cos(\omega_c + 2\omega_m)t + E_cJ_2\beta\cos(\omega_c - 2\omega_m)t + E_cJ_3\beta\cos(\omega_c + 3\omega_m)t - E_cJ_3\beta\cos(\omega_c - 3\omega_m)t + E_cJ_4\beta\cos(\omega_c + 4\omega_m)t + E_cJ_4\beta\cos(\omega_c - 4\omega_m)t.$$

$$v_{FM} = 6.7018\cos(240\pi)t + 2.1068\cos(240.4\pi)t - 2.1068\cos(239.96\pi)t +0.3211\cos(240.8\pi)t + 0.3211\cos(239.92\pi)t + 0.0323\cos(240.12\pi)t - 0.0323\cos(239.88\pi)t + 0.0022\cos(240.16\pi)t + 0.0022\cos(239.84\pi)t\ V$$

FM waveform has an infinite number of sidebands. The bandwidth of V_{Fm} depends on β. The spectrum contains the carrier components and side frequencies at humour of the modulating frequency.

The amplitudes of the various spectra components are given by a mathematical function known as Bessel function of the first kind denoted by: $J_n(m_f)$ being the modulation and n is the order of the side frequency. In mathematical notation; mf is termed the argument and n is termed the order. The Bessel functions are available in both graphical and tabular forms

$J_0(\beta) \rightarrow$ Amplitude of the carrier components
$J_1(\beta) \rightarrow$ Amplitude of the Ist order Frequency components $= (f_c \pm f_m)$
$(f_c \pm 2f_m)$
Note: odd order $\rightarrow$ we have $(f_c - f_m)$
Even order $\rightarrow$ we have $(f_c + f_m)$

The amplitude of the V_{fm} is constant. Thus the possible components add up to give a constant amplitude FM wave.

The **modulation index** β is another important quantity in any f.m. communication system. Here it is given by

$$\beta = \frac{\Delta\omega}{\omega_m} = \frac{\Delta f}{f_m}$$

Modulation Index

β values obviously change with f_m, so β will vary over a band of signal frequencies and can be very large. For example, 300 Hz audio with 75 kHz deviation gives a value of β of 150 whereas 15 kHz audio gives only 5. This is considered more in Section 4.16.

Now the general expression for v_{FM} can be analyse one stage further to give

$$v_{FM} = E_c \cos \omega_c t \cos (\beta \sin \omega_m t) - E_c \sin \omega_c t \sin (\beta \sin \omega_m t)$$

It is this expression which can be analysed to give the spectral components of an f.m. signal. There are two ways of doing this, one requiring simplifying assumptions and the other not.

4.16 Narrow-Band Frequency Modulation (NBFM)

Here a simplifying assumption is made that β is small – so small in fact that it makes $\beta \sin \omega_m t$ also small, allowing the following further assumptions to be made:

$$\text{Cos } (\beta \sin \omega_m t) = 1$$

$$\text{Sin } (\beta \sin \omega_m t) = \beta \sin \omega_m t$$

The usual criterion for these assumptions to be valid is that $\beta \leq 0.2$ rad. [check that cos (0.2) = 0.980 and sin(0.2) = 0.199 – good enough?]

Now the expression for v_{FM} can be written in a much simplified form as

$$v_{FM} = E_c \cos \omega_c t \,.\, 1 - E_c \cos \omega_c t(\beta \sin \omega_m t)$$

$$= E_c \cos \omega_c t - \beta E_c \cos \omega_c t \sin \omega_m t$$

Giving

$$v_{NBFM} = E_c \cos \omega_c t - \tfrac{1}{2}\beta E_c \cos(\omega_c + \omega_m)t + \tfrac{1}{2}\beta E_c \cos(\omega_c + \omega_m)t$$

This is a very simple expression, quite easy to understand. It consists of a carrier and one sideband pair similar to full AM, the only differences being that β is used instead of m and the lower sideband is negative. The two spectra are compared in Fig. 4.19. The value of β can be measured from this spectrum just as for m in amplitude modulation.

The fact that the lower sideband is negative is crucial and makes the modulated waveform quite different as can be seen in Fig. 4.20.

The frequency shift in NBFM at $\beta = 0.2$ is so small (βf_m compared with f_c) that it cannot be illustrated visibly, unlike the envelope variations for m = 0.2. The effect of this negative lower sideband can be seen much more graphically by considering of this negative lower sideband can be much graphically by considering the stationary phasor diagrams in Fig. 4.21.

Here you can see very clearly how the change in sign of the lower sideband causes changes in phase angle for NBFM but in amplitude only for full AM. It also Shows how FM produces a phase modulation type of effect ($\Delta\varnothing_{FM} = \beta = \Delta f/f_m$). Note also that there is a small amount of amplitude variation in NBFM (about 3% maximum) but this does not matter as an FM demodulator will ignore it.

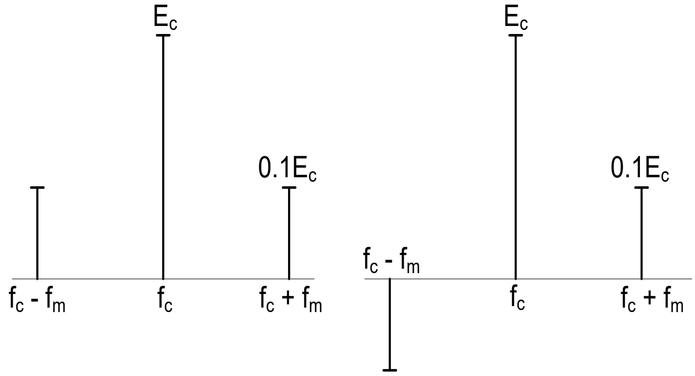


Figure 4.19. NBFM (right) and full AM (left) spectra compared (for β = m = 0.2)

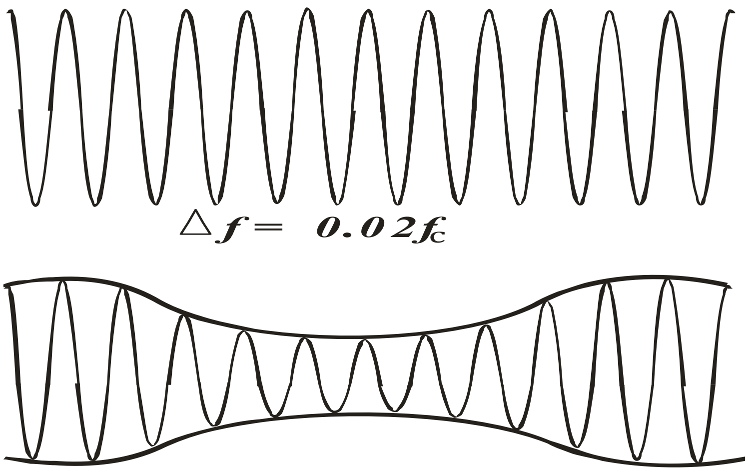


Figure 4.20. NBFM (upper) and full AM (lower) waveform compared (for β=m= 0.2)

From Fig. 4.20. you can see that the **bandwidth** of an NBFM signal is the same as that of full AM

$\beta = 2f_m$ …….NBFM BANDWIDTH

Hence, the name **narrow** band – compared with WBFM.

NBFM is not just a curiosity but forms a vital part of many f.m. systems. This is partly because of its narrow bandwidth, but also because the deviation is so small that it is easy to make it linear.

The bandwidth depends on the value of the modulation index, β. For small values of β, only the first term (carrier) and the first order term $(f_c \pm f_m)$ are significant and the bandwidth will be: $2\omega_m$. As β increases, more terms become significant and the bandwidth increases significantly.

The condition where β is small enough to make all the terms except the first two $(J_0\beta)$ and $(J_1\beta)$ negligible is the condition for narrow band FM (NBFM). NBFM is an example of linear modulation: A value of say $\beta < 0.2$ is taken to be sufficient to satisfy this condition. A value of $\beta = 0.5$ can also be read.

Although AM and NBFM signals have similarities, they are produced by different methods of modulation. The similarities and differences are portrayed by considering their phasor representation.

$$v_{am} = E_c \cos\omega_c t + \frac{m}{2} E_c \cos(\omega_c + \omega_m)t + \frac{m}{2} E_c \cos(\omega_c - \omega_m)t.$$

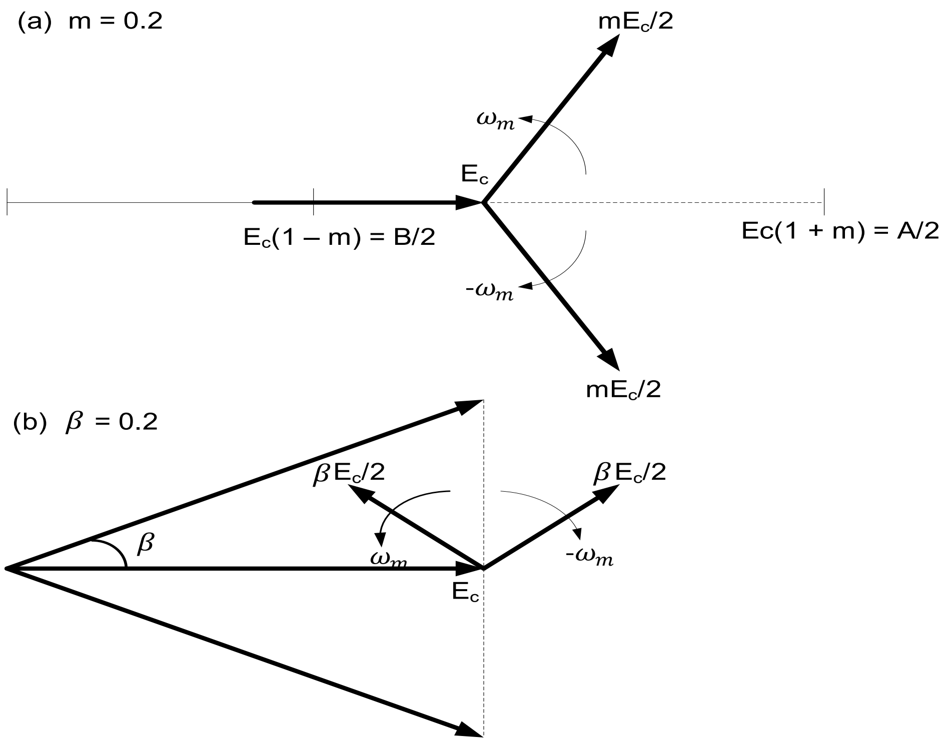


Figure 4.21. Stationary phasor diagram for (a) full AM and (b) NBFM

4.17 Wide Band Frequency Modulation (WBFM)

For values of β larger than 0.2, the general expression at the end of section 4.15 cannot be simplified any further but must be used as it. Luckily the awkward parts can be expanded by using the following relationships (which can are merely stated here and not derived):

$\cos(\beta \sin \omega_m t) = J_0(\beta) + 2J_2(\beta)\cos 2\omega_m t + 2J_4(\beta)\cos 4\omega_m t + \cdots$

$\sin(\beta \sin \omega_m t) = 2J_1(\beta)\sin \omega_m t + 2J_3(\beta)\sin 3\omega_m t + 2J_5(\beta)\sin 5\omega_m t + \cdots$

Where $J_n(\beta)$ etc. Are Bessel functions of β of order n.

To use these expressions it is not necessary to understand Bessel functions in full but a certain familiarity with them is useful. They come from solving a particular type of differential equation which often arises in actual physical situations. Bessel function values all have the form of pseudo-periodic functions of β, "decaying" as β increases, and of "period" increasing with n. They start at zero for $\beta = 0$ except for J_0 which starts at one. All this is shown in Fig. 4.22.

Also, to use Bessel functions quantitatively it is necessary to able to look up their values in tables such as those given in Tables 4.2 and 4.3.

Now let us see the effect of these relationships on the general expression for a

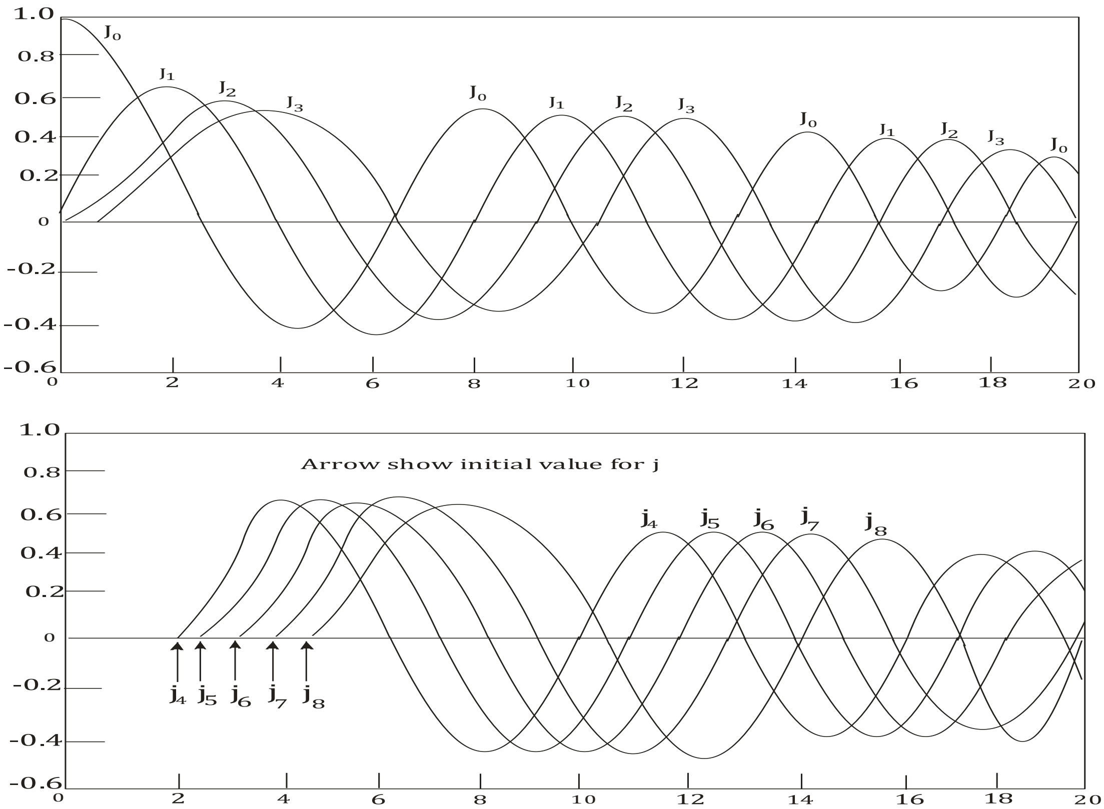


Figure 4.23. How Bessel functions vary with β and n (after Brown and Glazier, 1964)

Carrier modulated by a sinusoidal baseband in WBFM

$v_{WBFM} = E_c \cos \omega_c t \cos(\beta \sin \omega_m t) - E_c \sin \omega_c \sin(\beta \sin \omega_m t)$

$= E_c \cos \omega_c t[J_0(\beta) \cos 2\omega_m t + \cdots] - E_c \sin \omega_c t\, [2J_1(\beta) \sin \omega\, 3\omega_m t + \cdots]$

$= J_0(\beta)\, E_c \cos \omega_c t + 2J_2(\beta)\, E_c \cos \omega_m t + \cdots - 2J_1(\beta)\, E_c \sin \omega_m t - 2J_3(\beta)\, E_c \sin 3\omega_m t - \cdots$

$= J_0(\beta)\, E_c \cos \omega_c t + J_2(\beta)\, E_c \cos(\omega_c - 2\omega_m)t + J_2(\beta)\, E_c \cos(\omega_c + 2\omega_m) + \cdots - J_1(\beta)\, E_c \cos(\omega_c - \omega_m)t + J_1(\beta)\, E_c \cos(\omega_c + \omega_m)t - J_3(\beta)\, E_c \cos(\omega_c - 3\omega_m)t + J_3(\beta) E_c \cos(\omega_c + 3\omega_m)t - \cdots$

Table 4.2. Bessel function values, β = 0-2.5 (0.2 steps) and n = 0-8 (after Betts, 1970)

N	$J_n(0.2)$	$J_n(0.4)$	$J_n(0.6)$	$J_n(0.8)$	$J_n(1.0)$	$J_n(1.25)$	$J_n(1.5)$	$J_n(1.75)$	$J_n(2.0)$	$J_n(2.5)$
0	+0.9900	+0.9604	+0.9120	+0.8463	+0.7652	+0.6459	+0.5118	+0.3690	+0.2239	-0.0484
1	+0.0995	+0.1960	+0.2867	+0.3688	+0.4401	+0.5106	+0.5579	+0.5802	+0.5767	+0.4971
2	+0.0050	+0.0197	+0.0437	+0.0758	+0.1149	+0.1711	+0.2321	+0.2940	+0.3528	+0.4461
3	+0.0002	+0.0013	+0.0044	+0.0102	+0.0196	+0.0369	+0.0610	+0.0918	+0.1289	+0.2166
4			+0.0003	+0.0010	+0.0025	+0.0059	+0.0118	+0.0209	+0.0340	+0.0738
5					+0.0002	+0.0007	+0.0018	+0.0038	+0.0070	+0.0195
6							+0.0002	+0.0006	+0.0012	+0.0042
7								+0.0001	+0.0002	+0.0008
8										+0.0001

Table 4.3. Bessel function values, β = 0-20 (1.0 steps) and n=0-25 (after Betts, 1970)

N	$J_n(1)$	$J_n(2)$	$J_n(3)$	$J_n(4)$	$J_n(5)$	$J_n(6)$	$J_n(7)$	$J_n(8)$	$J_n(9)$	$J_n(10)$
0	+0.7652	+0.2239	-0.2601	-0.3971	-0.1776	+0.1506	+0.3001	+0.1717	-0.0903	-0.2459
1	+0.4400	+0.5767	+0.3391	-0.0660	+0.3276	-0.2767	-0.0047	+0.2346	+0.2453	+0.0435
2	+0.1149	+0.3528	+0.4861	+0.3641	+0.0466	-0.2429	-0.3014	-0.1130	+0.1448	+0.2546
3	+0.0196	+0.1289	+0.3091	+0.4302	+0.3648	+0.1148	-0.1676	-0.2911	-0.1809	+0.0584
4	+0.0025	+0.0340	+0.1320	+0.2811	+0.3912	+0.3576	+0.1578	-0.1054	-0.2655	-0.2196
5	+0.0002	+0.0070	+0.0430	+0.1321	+0.2611	+0.3621	+0.3479	+0.1858	-0.0550	-0.2341
6		+0.0012	+0.0114	+0.0491	+0.1310	+0.2458	+0.3392	+0.3376	+0.2043	-0.0145
7		+0.0002	+0.0025	+0.0152	+0.0534	+0.1296	+0.2336	+0.3206	+0.3275	+0.2167
8			+0.0005	+0.0040	+0.0184	+0.0565	+0.1280	+0.2235	+0.3051	+0.3179
9				+0.0009	+0.0055	+0.0212	+0.0589	+0.1263	+0.2149	+0.2919
10				+0.0002	+0.0015	+0.0070	+0.0235	+0.0608	+0.1247	+0.2075
11					+0.0004	+0.0020	+0.0083	+0.0256	+0.0622	+0.1231
12						+0.0005	+0.0027	+0.0096	+0.0274	+0.0634
13						+0.0001	+0.0008	+0.0033	+0.0108	+0.0290
14							+0.0002	+0.0010	+0.0039	+0.0120
15								+0.0003	+0.0013	+0.0045
16									+0.0004	+0.0016
17									+0.0001	+0.0005
18										+0.0002
19										
20										
21										
22										
23										
24										
25										

Table 4.3 (continued)

N	$J_n(11)$	$J_n(12)$	$J_n(13)$	$J_n(14)$	$J_n(15)$	$J_n(16)$	$J_n(17)$	$J_n(18)$	$J_n(19)$	$J_n(20)$
0	-0.1712	+0.0477	+0.2069	+0.1711	-0.0142	-0.1749	-0.1699	-0.0134	+0.1466	+0.1670
1	-0.1768	-0.2234	-0.0703	+0.1334	+0.2051	+0.0904	-0.0977	-0.1880	-0.1057	+0.0668
2	+0.1390	-0.0849	-0.2177	-0.1520	+0.0416	+0.1862	+0.1584	-0.0075	-0.1578	-0.1603
3	+0.2273	+0.1951	+0.0033	-0.1768	-0.1940	-0.0438	+0.1349	+0.1863	+0.0725	-0.0989
4	-0.0150	+0.1825	+0.2193	+0.0762	-0.1192	-0.2026	-0.1107	+0.0696	+0.1806	+0.1307
5	-0.2383	-0.0735	+0.1316	+0.2204	+0.1305	-0.0575	-0.1870	-0.1554	+0.0036	+0.1512
6	-0.2016	-0.2437	-0.1180	+0.0812	+0.2061	+0.1667	+0.0007	-0.1560	-0.1788	-0.0551
7	+0.0184	-0.1703	-0.2406	-0.1508	+0.0345	+0.1825	+0.1875	+0.0514	-0.1165	-0.1842
8	+0.2250	+0.0451	-0.1410	-0.2320	-0.1740	-0.0070	+0.1537	+0.1959	+0.0929	-0.0739
9	+0.3089	+0.2304	+0.0670	-0.1143	-0.2200	-0.1895	-0.0429	+0.1228	+0.1947	+0.1251
10	+0.2804	+0.3005	+0.2338	+0.0850	-0.0901	-0.2062	-0.1991	-0.0732	+0.0916	+0.1865
11	+0.2010	+0.2704	+0.2927	+0.2357	-0.1000	-0.0682	-0.1914	-0.2041	-0.0984	+0.0614
12	+0.1216	+0.1953	+0.2615	+0.2855	+0.2367	+0.1124	-0.0486	-0.1762	-0.2055	-0.1190
13	+0.0643	+0.1201	+0.1901	+0.2536	+0.2787	+0.2368	+0.1228	-0.0309	-0.1612	-0.2041
14	+0.0304	+0.0650	+0.1188	+0.1855	+0.2464	+0.2724	+0.2364	+0.1316	-0.0151	-0.1464
15	+0.0130	+0.0316	+0.0656	+0.1174	+0.1813	+0.2399	+0.2666	+0.2356	+0.1389	-0.0008
16	+0.0051	+0.0140	+0.0327	+0.0661	+0.1162	+0.1775	+0.2340	+0.2611	+0.2345	+0.1452
17	+0.0019	+0.0057	+0.0149	+0.0337	+0.0665	+0.1150	+0.1739	+0.2286	+0.2559	+0.2331
18	+0.0006	+0.0022	+0.0063	+0.0158	+0.0346	+0.0668	+0.1138	+0.1706	+0.2235	+0.2511
19	+0.0002	+0.0008	+0.0025	+0.0068	+0.0166	+0.0354	+0.0671	+0.1127	+0.1676	+0.2189
20		+0.0003	+0.0009	+0.0028	+0.0074	+0.0173	+0.0362	+0.0673	+0.1116	+0.1647
21			+0.0003	+0.0010	+0.0031	+0.0079	+0.0180	+0.0369	+0.0675	+0.1106
22			+0.0001	+0.0004	+0.0012	+0.0034	+0.0084	+0.0187	+0.0375	+0.0676
23				+0.0001	+0.0004	+0.0013	+0.0037	+0.0089	+0.0193	+0.0381
24						+0.0005	+0.0015	+0.0039	+0.0093	+0.0199
25						+0.0002	+0.0006	+0.0017	+0.0042	+0.0098

$$v_{WBFM} = E_c[\cdots - J_3(\beta)\cos(\omega_c - 3\omega_m)t + J_2(\beta)\cos(\omega_c - 2\omega_m)t - J_1(\beta)\cos(\omega_c - \omega_m) + J_0(\beta)\cos\omega_c t + J_1(\beta)\cos(\omega_c + \omega_m)t + J_2(\beta)\cos(\omega_c + 2\omega_m) + J_3(\beta)\cos(\omega_c + 3\omega_m)t + \cdots]$$

Note that now, unlike NBFM, there are many more sideband pairs (up to $\beta + 1$ of them). It also appears that there is a pattern about their signs – all positive except the odd lower ones – but this neat pattern is distortrd in reality by the signs of the Bessel functions which are just as likely to be negative as positive as shown in Fig. 4.23.

Of much more significance are the magnitudes of the spectral lines which are the same for both upper and lower components of each sideband pair producing characteristic patterns of amplitude symmetrical about the carrier position. The carrier itself is, otherwise, of no special importance and may be very small or even zero in value. All this is shown in Fig. 4.24 for a range of β and Δf.

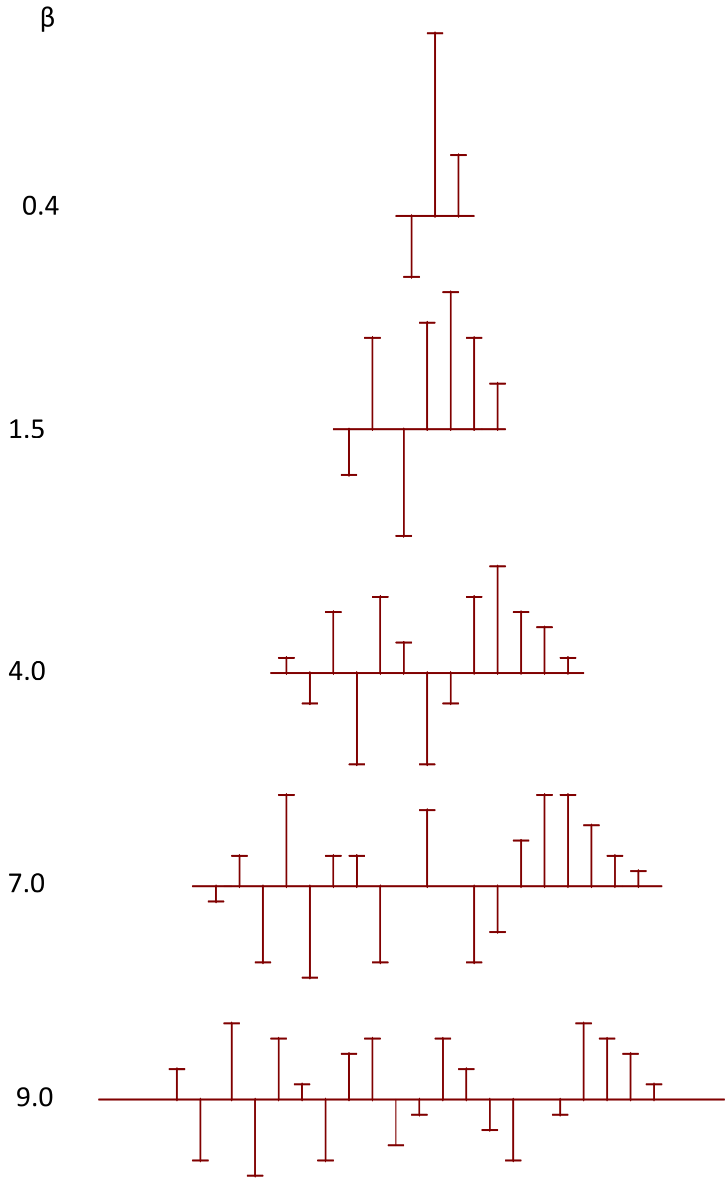


Figure 4.24. Total spectra for small β values.

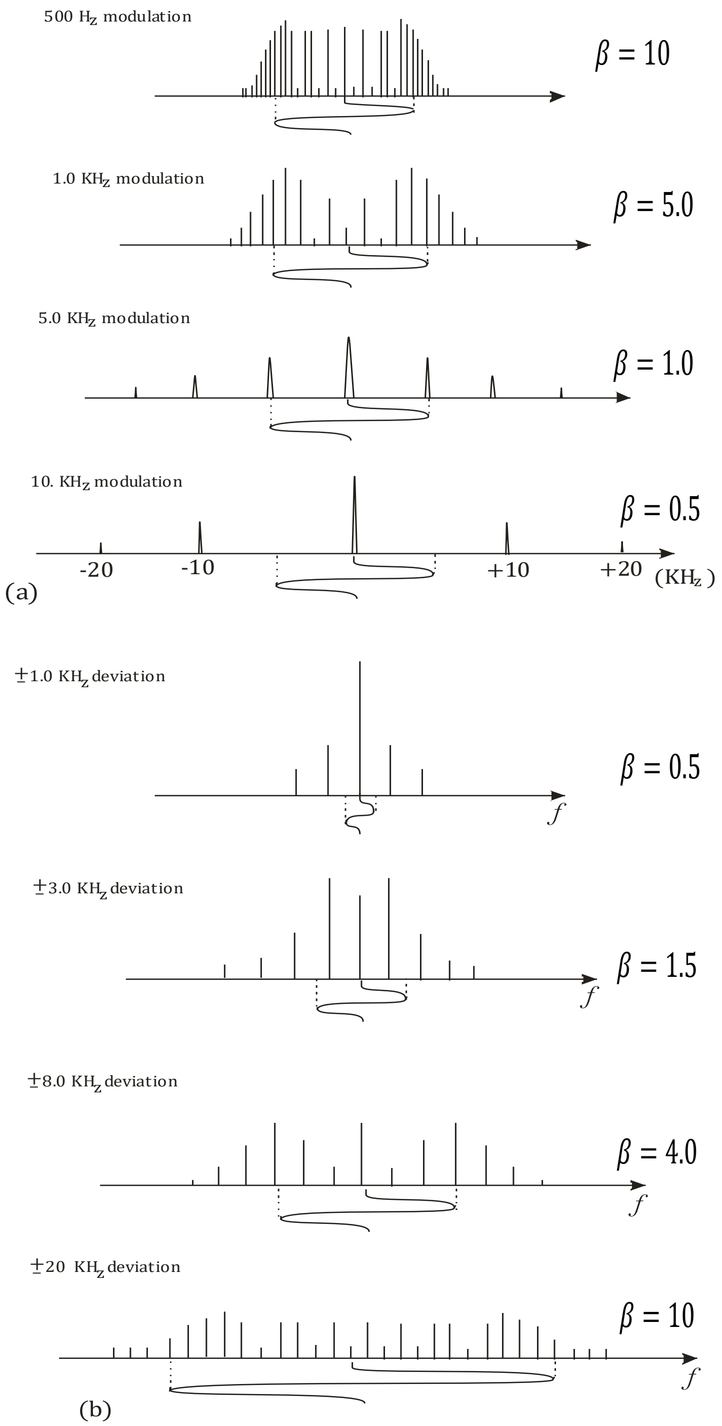


Figure 4.25 Frequency-modulated amplitude spectra for various β values: (a) Δf constant (± 5 kHz), P_m and β varying; (b) f_m constant (1 kHz), Δf (and β) varying (after Hardy, 1986)

Note how the line amplitudes fall off rapidly at frequencies beyond its deviation. This is related to bandwidth which is discussed later (section 4.17).

The general rule is that the lines are strongest for deviations at which the baseband waveform lingers for the longest time. This is near the peaks for a sinusoidal and produces the characteristics "eared" envelope shape as in Fig. 4.24. Other waveforms will have their spectra at different places and, for example, there will be an even greater concentration at the extremes for a square wave. An example is given in Fig. 4.25.

Another way in which the Bessel function values are useful is to help identify the β value for certain spectra. This is because their zero values give zero amplitudes for the spectra lines for which they are a multiplier. Some examples are given in Table 4.4.

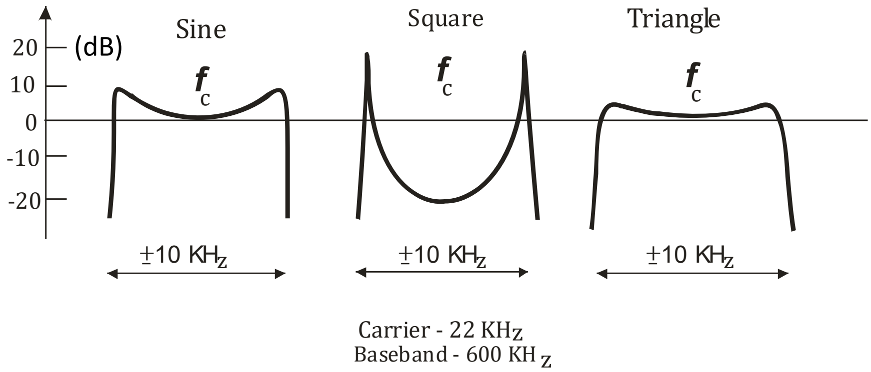


Figure 4.26. Spectra concentration near deviation "pauses"

Table 4.4. Spectra line zeros

At	$\beta = 2.4$	The carrier is zero
At	$\beta = 3.8$	The first pair are zero
At	$\beta = 5.2$	The second pair are zero
At	$\beta = 5.5$	The carrier is again zero
At	$\beta = 6.4$	the third pair are zero
At	$\beta = 7.0$	The first pair are again zero
At	$\beta = 7.6$	The fourth pair are zero

And so on (e.g. the carrier is again zero at 8.7, 11.9, 15.0 and 18.0)

It is very illustrative to set up a simple FM system (e.g. using the VCO input to a function generator) and vary v_m, checking β values from the zeros above and also by measuring bandwidth as below. This forms an excellent teaching experiment in the laboratory.

As mentioned earlier, the FM has infinite no. of sidebands as illustrated in Figure ... (plots of sideband magnitudes for several values of β).

Magnitude of the spectra components of the higher order sidebands become negligible for all practical purposes; giving rise to a finite bandwidth.

How many sidebands are important to the FM transmission of a signal? Depends on the intended application and the fidelity requirements. A sideband is significant if its magnitude is equal to or exceeds 1% of the unmodulated carrier i.e. if $J_m(\beta) \geq 0.01$ then the order is significant. The no of significant sidebands sidebands for different values of β can be found from a plot a table of Bessel function.

The bandwidth for FM signals can be appear given as $\simeq 2(\Delta\omega_c + \omega_m)$

$$\frac{\beta}{\omega} \simeq 2\omega_m(H\beta)$$

$$\beta = \frac{\Delta\omega_c}{\omega_c}$$

This is known as Carson's rule

As you can see from the spectra in Fig. 4.24, there is no clear cut limit on bandwidth for WBFM as there is for NBFM and full AM. On the other hand, spectral line magnitude does fall off very rapidly at deviations greater than $\pm\Delta f$ is a good approximation known as the **nominal bandwidth,** especially at larger β values, but, for general use, more quantitative definitions are required.

The most common of these is the **Carson bandwidth** defined as that bandwidth which will let trough at least 98% of the signal power. The usefulness of this definition is that it always requires $\beta + 1$ sideband pairs at any value of β. Since the spectral lines are f_m apart, this means that the bandwidth is B where

$$B = 2(\beta + 1)f_m \qquad \textbf{Carson Bandwidth}$$

Which can also be written $B = 2(\Delta f + f_m)$

Note that this includes one more sideband pair than the nominal bandwidth but that that extra width is negligible if β is large.

Another definition is the **1% bandwidth** defined as that bandwidth lying inside the sideband pair beyond which all spectral amplitudes are smaller than 1% of the largest line amplitude. This is usually wider than the Carson bandwidth.

For $\beta = 1$ the Carson bandwidth gives the following values:

giving $P = \frac{1}{2}E_c{}^2$. Also,

$$v_{FM} = J_0(5)\cos\omega_c t - J_1(1)\cos(\omega_c - \omega_m)t$$
$$+ J_2(1)\cos(\omega_c - 2\omega_m)t + (1)\cos(\omega_c + 2\omega_m)t$$
$$= 0.765E_c - 0.440E_c + 0.115RE_c + 0.115E_c$$

Therefore

$$P = 0.5[0.765^2 + 2(0.440)^2 + 2(0.115)^2]E_c{}^2$$
$$= [0.4994]E_c{}^2$$

This is 99.93% of the total power – well within the Carson definition. The third sideband pair with $J_3(1)$ of 0.0196 add very little – calculate them to check.

4.18 Modulation by a band of frequencies

So far we have only considered by a single sinudoid signal. What happens when, as is usually the case, the signal is a band of frequencies – a true **baseband?** An example of such a band is a single telephone channel of 0-4 kHz nominal (300-3400 Hz usable). Obviously the first step is to consider the band as the sum of a whole lot frequencies, and then see what they have in common and where they differ.

The main **common factor** is Δf which is the **same** for all f_m values because it depends on signal amplitude only, not frequency. On the other hand,($\beta = \Delta f/f_m$). The baseband **waveform** itself will still be periodic but non-sinusoidal and also varying continuously in shape as E_m and f_m change rapidly. Thus, this is not particularly useful except to give an instantaneous idea of what the signal might look like. It does give a very qualitative picture of the range of wavelengths involved. The **spectrum** will still consist of separate lines. There will be many more of them, closer together, and they will be continuously varying in position as the spectral content of v_m changes. Instantaneously there will still be symmetry in the amplitude spectrum but this is of little use.

The **bandwidth** will depend partially onf_m which must therefore be large enough to allow for the largest f_m value. For example, a telephone channel requires

$$B = 2(\Delta f + 3400)\ \text{Hz}$$

4.19 Similarities: AM & NBFM

some frequency components i.e. $\omega_c, \omega_c + \omega_m, \omega_c - \omega_m$. Hence, same bandwidth $2\omega_m$. Linear modulation.

Differences

Modulation is added is quadrature with the carrier in NBFM, where as modulation is added in place to the carrier in AM. NBFM gives rise to phase variation τ very little amplitude change whereas the AM gives amplitude change where as the AM gives amplitude variation τ .... phase deviation.

Note: NBFM is used in telemetry and mobile communication.

Commercial FM Transmission

Carrier frequencies are spaced at 200 kHz intervals in the range 88-108 MHz are assigned for commercial FM transmissions. Maximum deviation in this case (150 kHz being used) is 75 kHz above or 75 kHz below the centre. The 200 kHz available to each station in compassion with 10 kHz for AM allows the transmission of high fidelity programme material τ room to square (i.e. leaving some frequency range unused). WBFM is used to fill the band.

Suppose the FM is 15 kHz

$$\Delta f_c = 5\ \text{kHz}$$
$$f_m = 15\ \text{kHz}$$

$$\beta = \frac{\Delta\omega_c}{\omega_m} = \frac{\Delta f_c}{f_m} = \frac{75}{15} = 5$$
$$\beta = 2(\Delta f_c + f_m)$$
$$\beta = 2(15 + 75)$$
$$\boldsymbol{\beta = 180\ \mathbf{kHz}}$$

Average power in FM

$$V_{fm}(t) = E_c \cos(\omega_c t + \beta \sin \omega_m t)$$

$\text{Power} \left(\frac{E_c}{\sqrt{2}}\right)^2 = \frac{{E_c}^2}{2}$

Total average power is a constant regardless of the modulation index.

4.20 Summary of expressions

Basis of Fm:

$$\delta\omega = K v_m$$

General expression for any baseband:

$$v_{FM} = E_c \cos(\omega_c t + K \int_0^t v_m dt)$$

General expression for a single sinusoidal baseband:

$$v_{FM} = E_c \cos(\omega_c t + \beta \sin \omega_m t)$$

NBFM:

$$v_{NBFM} = E_c \cos\omega_c t - \frac{1}{2}\beta E_c \cos(\omega_c + \omega_m)t$$

WBFM:

$$\begin{aligned} v_{WBFM} &= E_c \cos \omega_c t \cos(\beta \sin \omega_m t) - E_c \sin \omega_c t \sin(\beta \sin \omega_m t) \\ &= \cdots - J_3(\beta)\cos(\omega_c - 3\omega_m)t + J_2(\beta)\cos(\omega_c - 2\omega_m)t \\ &\quad - J_1(\beta)\cos(\omega_c - \omega_m)t + J_0(\beta)\cos\omega_c + J_1(\beta)\cos(\omega_c + \omega_m)t \\ &\quad + J_2(\beta)\cos(\omega_c + 2\omega_m)t + J_3(\beta)\cos(\omega_c + 3\omega_m)t + \cdots \end{aligned}$$

Definitions

Modulation sensitivity (K) $\delta\omega = K V_m$

Deviation ($\Delta\omega$) $\Delta\omega = K E_m$

Modulation index (β) $= \Delta\omega/\omega_m = \Delta f/f_m$

Carson bandwidth (β) $= 2(\Delta f + f_m) = 2(\beta + 1)f_m$

Nominal bandwidth $\beta = 2\Delta f$

4.21 Conclusion

Frequency modulation is capable of both high linearity and large dynamic range. Also, it is much less susceptible to amplitude noise corruption than amplitude modulation provided the signal power is large enough. By using higher carrier frequencies, larger baseband bandwidths can be used enabling quality stereo sound broadcasting.

True NBFM (Δf of the order of Hertz) is used in broadcasting to provide highly linear modulation which is then converted to WBFM for transmission. Low Δf values (e.g. ± 2.5 kHz) are also used in communications applications such as mobile radio, CB, and amateur radio. These give noise advantages whilst keeping within small channel bandwidths.

4.22 Chapter Review Problems

4.1. For f_m = 3 kHz, E_m = 3.0 V, $\beta = 4.0, E_c = 2.0$ V, and $f_c = 200$ kHz calculate:

(i)The frequency deviation (Δf)

(ii) The frequency modulation sensitivity (K)

(iii) The maximum and minimum instantaneous frequencies (f_m)

(iv) The Carson bandwidth (β)

(v) The f.m. side frequencies and amplitudes.

4.2. Sketch the spectrum for 4.1(v) above, showing the bandwidth.

4.3. Repeat problems 4.1. and 4.2. for $\beta = 0.004$ and $\beta = 40$ [you will need plenty of time].

4.4. For $\Delta f = 75$ kHz, f_c = 100 MHz, and an audio baseband extending from 300 Hz to 3400 Hz calculate:

(i) The r.f. bandwidth

(ii) The maximum and minimum modulation indices

(iii) The highest baseband frequency which could be sent without using NBFM

(iv) The modulation index for the stereo pilot tone at 19 kHz.

4.5. A square wave baseband at 2.5 kHz repetition frequency and $\pm$ 6 V magnitude is frequency modulated on to a carrier at (kHz with modulation sensitivity of 2500 rad V^{-1}:

(i) What is the modulation index?

(ii) What is the frequency deviation?

(iii) What is the bandwidth?

(iv) Sketch the modulated waveform

(v) Draw the instantaneous amplitude spectrum at the two extremes of high voltage.

(vi) Draw the f.m. spectrum and compare with that in (v).

4.6. Draw the phasor diagram for $\beta = 0.2$ showing especially:

Maximum and minimum phase deviations (compare with calculated values).

The phasor position when $f_i = f_c$; at min f_i; at max f_i; and v_m is 45° from v_c.

Estimate how much AM occurs (calculate equivalent m).

4.7. Explain why FM is used for high-quality sound broadcasting. Also explain why it is not used for medium-wave broadcasting.

4.8. Find the value of the modulation index (β) for:

(i) $v_{FM}= E_c\cos\,(\omega_c t + \int K \cos \omega\; tdt)$.

(ii) $v_{FM}= 1.0 \cos(10^5 t + 10 \sin 10^3 t)$:

(iii) $v_{FM} = 5.0 \cos\,(2\pi x 10^4 t)$.

(iv) $1.0 \cos 100t$ frequency modulated on to 1.0 cos 10^5t with $K = 10^4$.

(v) As (iv) but phase modulated with β_p = 1 (See chapter 5)

4.9. Find the frequency separation for the first pair of sidebands for each of the waveform in Problem 4.8.

4.10. Work out the relative sideband amplitudes for f.m. waveforms for which β = 0.1, 2.0, 20, and 100. Go up to the $(n + 1)$th pair.

4.11. Sketch the phasor diagrams for both full AM and NBFM $(m = \beta = 0.2)$ with sidebands (LSB and USB) at the following relationships to the carrier:

(i) LSB in phase with the carrier

(ii) LSB in anti-phase with the carrier

(iii) LSB leads carrier by $\frac{\pi}{2}$

(iv) LSB lags carrier by $\frac{\pi}{2}$.

4.12. For the situations in Problem 4.11 work out:

(i) Amplitude modulation depth

(ii) Amount of AM in the NBFM signal

(iii) Maximum phase deviation in the NBFM signal

(iv) The greatest rate of change in NBFM if f_m = 1 kHz.

(v) The instantaneous frequency at the points of maximum phase deviation for f_c = 100 kHz.

4.13. Find the percentage power transmitted within the Carson bandwidth for $\beta = 0.5, 2.0, 10\; and\; 50$. Then calculate the percentage power lost if the outermost sideband pair is not sent.

4.14. For VHF FM radio find the highest modulating signal frequency for which Δf meets Carson's rule.

4.15. A carrier signal whose voltage is given by

$$v_c= 10 \cos\,(2\pi \times 10^6)t$$

Is frequency modulated by a single frequency tone whose voltage is

$$v_m = 2\cos(2\pi \times 10^4)t$$

Using a modulation index of 5.

Write down the expression for the modulated signal showing clearly how frequency change is related to modulation amplitude. From this obtain an expression showing the amplitudes and frequencies of all components of the spectrum of the modulated signal which fall within the Carson bandwidth.

Explain the criteria for deciding on the bandwidth you use and verify if it is correct in this case:

4.16 The circuit shown in Fig. 4.26 has a frequency-modulated input signal whose waveform is given by the expression

$$v_c = E_c \cos\left(\omega_c t + \int_0^t K v_m dt\right)$$

Where v_m is the waveform of the original baseband, K is the modulation sensitivity, and E_c is the carrier amplitude.

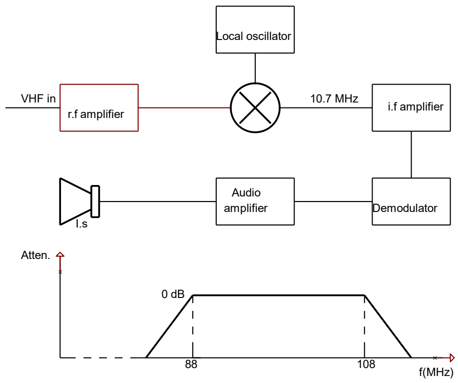


Figure 4.26. Circuit for Problem 4.16

From this expression obtain a relationship giving the modulation index in terms of frequency deviation for a single sinusoidal signal.

Given that: the centre frequency of the carrier is 90 MHz
The centre frequency of the i.f. signal is 10.7 MHz
The frequency deviation is $\pm$ 75 kHz
The baseband is from 30 Hz to 15 kHz

Calculate: (i) the extreme values modulation index
(ii) the maximum Carson bandwidth
(iii) The local oscillator frequency
(iv) the probable practical; i.f. bandwidth.

Derive and additional expression you use. [Note: this question requires some understanding of the organization and operation of a radio receiver].

4.17. Define the modulation of a frequency-modulated signal. A carrier has a waveform given by $v_c = 5.0 \cos(2\pi\ x\ 10^8)t$. It is frequency modulated by a baseband signal whose waveform is given by v_m = 1.0 cos ω_mt, where $\omega_m = 2\pi f_m$ and f_m is in the range 200 Hz to 20 kHz. The frequency deviation is kept constant at 200 kHz. For f_m at the lower limit of its range calculate the modulation index, the modulation sensitivity and the bandwidth. Any expression you use must be derived or explained, whichever is most appropriate.

Now calculate the same three quantities for the situation where f_m is at the upper limit of the baseband range. In addition sketch the amplitude spectrum of the modulated signal showing the signs and relative magnitude of each component.

4.18. Give the basic definition of frequency modulation and show that this leads to a phase change proportional to the integral pf the baseband signal voltage.

A signal sinusoidal baseband at 3 kHz is frequency modulated on to a carrier at 300 kHz with a frequency deviation of 450 Hz. Derive an expression for the waveform of the modulated carrier and sketch its spectrum relative magnitudes.

Using a stationary phasor diagram, show how this modulation produces phase changes in the carrier. Compare this with a full AM signal at low modulating factor.

4.19. An FM mono broadcasting station is transmitting music band limited to 15 kHz. It uses a Carson bandwidth of 150 kHz. For a tone at the upper limit of the baseband calculate the modulation index. Give the full Carson spectrum of the broadcast signal modulated by that tone for a carrier at 95 MHz exactly, showing the relative amplitude of each spectral component, including sign. Derive all the theory you need to obtain the spectrum.
State the likely minimum value of modulation index for a stereo broadcast using the same baseband and broadcast bandwidth.

4.20 A 18 W, 120 MHz higher frequency signal is frequency modulation with 40 KHz lower frequency signal such that the peak frequency deviation is 22 kHz. The resistance level is 4 Ω.

Determine:

(i) the amplitude of the higher frequency signal.

(ii) The modulation index.

(iii) The Carson's bandwidth.

(iv) The modulation sensitivity if the amplitude of the lower frequency signal is 5.68 V.

(v) State the amplitude of all components in the resulting frequency modulated signal. (Quote the frequency and the amplitude of each frequency component). Sketch the frequency spectrum of the frequency modulated signal.

(University of Ibadan, TEL412- Communication system I 2004/2005 BSc degree Exam).

4.21 (a) What are the advantages (if any) of SSB over full AM?

(i) Draw block diagrams to illustrate the production of DSBSC and SSB by use of balanced modulators (Don't make use of any filter).
(ii) Draw the basic Envelope Detector Circuit.

4.22 A 18 W, 150 MHz higher frequency current signal is frequency modulated with a 44 kHz lower frequency signal such that peak frequency deviation is 16 kHz.
The resistance level is 6 Ω. Determine:
(i) the amplitude of higher frequency current signal.
(ii) the modulation index.
(iii) the Carsons bandwidth.
(iv) the modulation sensitivity if the amplitude of the lower frequency signal is 6.2655 mA.
(v) State the amplitudes of all components in the resulting frequency modulated signal. (Quote the frequency and the amplitude of each frequency component). Sketch the frequency spectrum of the frequency of the frequency modulated signal.
(University of Ibadan, TEL412- Communication system I 2003/2004 BSc degree Exam).

4.23 (a) A 10 MHz carrier is frequency modulated by 500 kHz sinusoidal signal such that the peak frequency deviation is 50 kHz. Determine the approximate bandwidth of FM. What type is FM signal?
(b) A 18 W, 240 MHz higher frequency current signal is frequency modulated with a 25 kHz lower frequency signal such that peak frequency deviation is 15 kHz. The resistance level is 3 Ω.
Determine:
(i) the amplitude of higher frequency voltage signal.
(ii) the modulation index.
(iii) the Carsons bandwidth.
(iv) the modulation sensitivity if the amplitude of the lower frequency signal is 5.5658 V.
(v) State the amplitudes of all components in the resulting frequency modulated signal. (Quote the frequency and the amplitude of each frequency component). Sketch the frequency spectrum of the phase modulated signal. (Calculate power in cash component and sum for all components in the phase modulated signal).
(University of Ibadan, TEL412- Communication system I 2000/2001 BSc degree Exam).

CHAPTER 5

PHASE MODULATION

5.0 Introduction

Another way to produce angle modulation is to vary the amount of phase shift of a constant-frequency carrier in accordance with a modulating signal as shown in Fig. 4.2. The resulting output is a PM signal. Imagine a modulator circuit whose basic function is to produce a phase shift.

Remember that a phase shift refers to a time separation between two sine waves of 'the same frequency. Assume that we can build a phase shifter that causes the amount of phase shift to vary with the amplitude of the modulating signal. The greater the amplitude of the modulating signals, the greater the phase shifts. Assume further that positive alternations of the modulating signal produce a lagging phase shift and negative signals produce a leading phase shift.

If a constant-amplitude-constant-frequency carrier sine wave is applied to the phase shifter, the output of the phase shifter will be a PM wave. As the modulating signal goes positive, the amount of phase lag increases with the amplitude of the modulating signal. This means that the carrier output is delayed. That delay increases with the amplitude of the modulating signal. The result at the output is as if the constant-frequency carrier signal had been stretched out or its frequency lowered.

When the modulating signal goes negative, the phase shift becomes leading. This causes the carrier sine wave to be effectively speeded up or compressed. The result is as if the carrier frequency had been increased.

Phase modulation produces frequency modulation. Since the amount of phase shift is varying, the effect is as if the carrier frequency is changed. Since PM produces FM, PM is often referred to as *indirect FM.*

5.1 Phase Modulators

Most modern FM transmitters use some form PM to produce indirect FM. The reason for using PM instead of direct FM is that the car-oscillator can be optimized for frequency accuracy and stability. Crystal oscillators or crystal-controlled frequency synthesizers can be used to set the carrier frequency accurately maintain solid stability.

The output of the carrier oscillator is fed to a phase modulator where the phase shift is made vary in accordance with the modulating signal. Since phase variations produce frequency variations, indirect FM is the result.

5.2 Basic Phase-Shift Circuits

The simplest phase shifters are *RC* networks like those shown in Figs. 5.1(a) and (b). Depending upon the values of *R* and *C,* the output of the phase shifter can be set to any phase angle between 0° and 90°. In Fig. 5.1(a) the output leads the input by

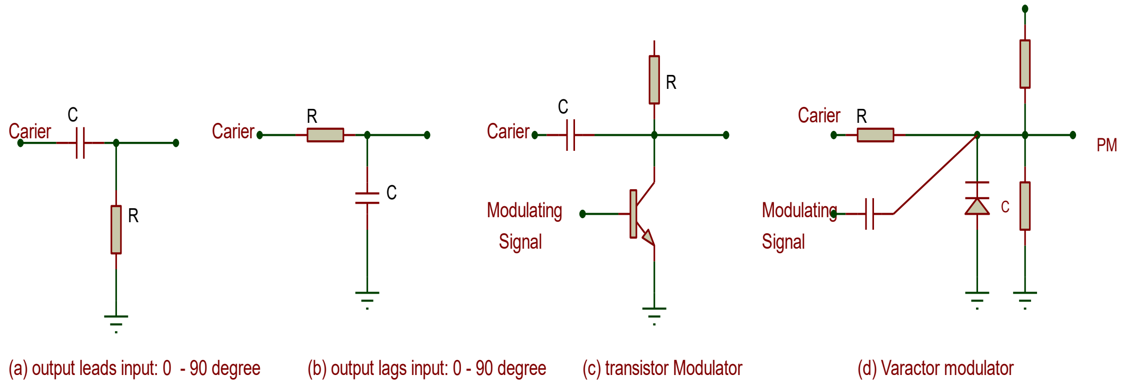


Figure 5.1 Simple PM circuits.

some angle between 0° and 90°. For example, when X_c equals *R,* the phase shift is 45°.

A low-pass version of the same *RC* filter can also be used, as shown in Fig. 5.1(b). Here the output is taken from across the capacitor, so it lags the input voltage by some angle between 0° and 90°.

One of these simple phase-shift circuits can be used as a phase modulator if the resistance or capacitance can be made to vary with the modulating signal. One way to do this is to substitute a transistor for the resistor in the circuit in Fig. 5.1(a). The resulting phase-shift circuit is shown in Fig. 5.1(c). The transistor simply acts as a variable resistor that varies in response to the modulating signal. If the modulating signal increases, the transistor base current and collector current increase. Therefore, the effective transistor resistance decreases. Lowering the resistance increases the amount of phase shift. This causes a corresponding frequency increase. If the amplitude of the modulated signal decreases, the base and collector currents decrease and the effective transistor resistance increases. This decreases the amount of phase shift, and, as a result, the amount of frequency shift decreases. An FET can be substituted for the bipolar transistor in Fig. 5.1(c) with comparable results.

Fig. 5.1(d) shows how a varactor can be used to implement a simple low-pass phase-shift modulator. Here the modulating signal causes the capacitance of the varactor to change. If the modulating signal amplitude increases, it adds to the varactor bias from R_1 and R_2, thereby causing the capacitance to decrease. This causes the reactance to increase; thus, the circuit produces less phase shift. A decreasing modulating signal subtracts from the reverse bias on the varactor diode, thereby increasing the capacitance or decreasing the capacitive reactance. This increases the amount of phase shift.

5.3 Practical Phase Modulators

A common phase modulator is shown in Fig. 5.2. It uses a phase shifter made up of a capacitor and the variable resistance of a field-effect transistor Q_1. The carrier signal from a crystal oscillator or a phase-locked loop frequency synthesizer is applied directly to the output through C_1 and C_2. The carrier signal is also applied to the gate of the FET through C_1. The series capacitance of C_1 and C_2 and the FET source to drain resistance produce a leading phase shift of current in the FET and a

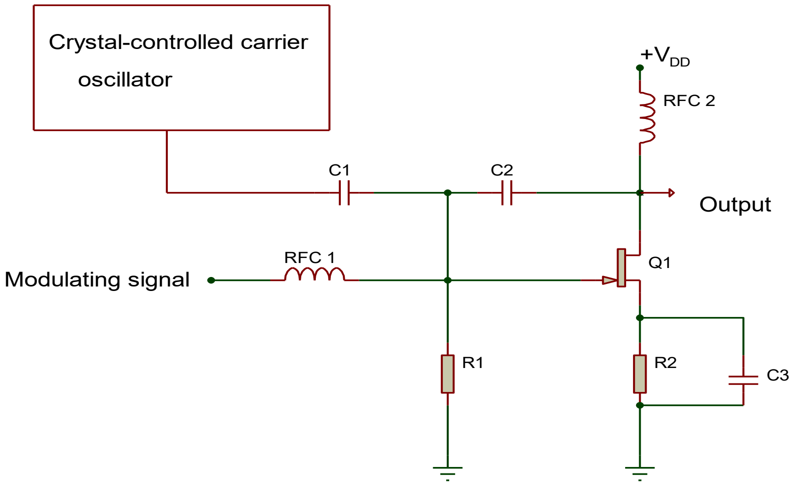


Figure 5.2 An Improved Phase Modulator

leading voltage at the output. The carrier signal applied to the gate of the FET also varies the FET current. C_1 and R_1 produce a leading phase shift less than 90°. The leading voltage across RI also controls the current in Q_1. With two signals controlling the FET current, the result is a phasor sum of the two currents.

The modulating signal is applied to the gate of the FET. RFC_1 keeps the carrier RF out of the audio circuits. The audio signal now also controls the FET current. This changes the amplitude relationships of the other two controlling inputs, thereby producing a phase shift that is directly proportional to the amplitude of the modulating signal. The carrier output at the FET drain varies in phase and amplitude. The signal is then usually passed on to a class C amplifier or frequency multiplier which removes the amplitude variations but preserves the phase and frequency variations.

The problem with such simple phase modulators is that they are capable of producing only a small amount of phase shift because of the narrow range of linearity of the transistor or varactor. The total amount of phase shift is essentially limited to ±20°. Such a limited amount of phase shift produces, in turn, only a limited frequency shift.

5.4 Using a Tuned Circuit for Phase Modulation

An improved method of PM is to use a parallel tuned circuit to produce the phase shift. At resonance, the parallel resonant circuit will have a very high resistance. Off resonance, the circuit will act inductively or capacitively and as a result, will produce a phase shift between its current and applied voltage.

Fig. 5.3 shows the basic impedance Z response curve of a parallel resonant circuit. Also, shown is the phase variation *A*. At the resonant frequency f_r the inductive and capacitive reactance's are equal, and therefore, their effect cancel one another. The result is extremely high impedance at f_r. The circuit acts resistive at this point, and, therefore, the phase angle between the current and the applied voltage is ZERO.

At frequencies below resonance X_1 decreases and X_c increases. This causes the circuit in act like an inductor. Therefore, the current will lag the applied voltage. Above resonance, X_l increases while, X_c decreases. This causes the circuit to act capacitively, and the current leads the applied voltage. If the *Q* of the resonant circuit is relatively high, the phase shift will be quite pronounced, as shown in Fig. 5.3. The same effect is achieved if f is constant and either *L* or *C* is varied. For a relatively small change in *L* or *C* a significant phase shift can be produced. The idea then is

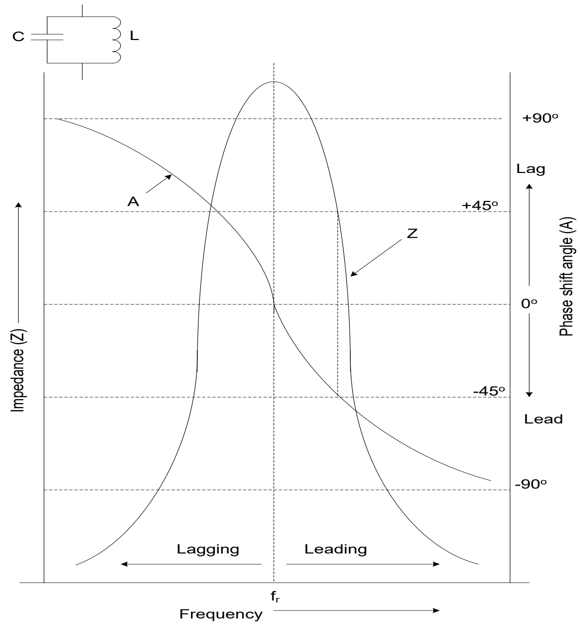


Figure 5.3 Impedance and Phase Shift Versus Frequency of *a* Parallel Resonant Circuit.

to cause the inductance in capacitance to vary with the modulating voltage and thus produce a phase shift.

A variety of circuits have been developed based on this technique, bin one illustrating its operation is shown in Fig. 5.4. The parallel limed circuit made up of *L* C_1 and C_2 is part of the output circuit of an RF amplifier driven by the carrier oscillator. Capacitor C_1 is large so that its reactance at the carrier frequency is low. Therefore, the resonant frequency is controlled by C_2. A varactor diode D_1 is connected in parallel with C_2 in the timed circuit and, therefore, will provide a capacitance change with the modulating signal. The voltage divider made up of R_1 and R_2 Sets the reverse bias on D_1. The value of C_3 is very large and simply acts as a dc blocking capacitor, preventing bias from being applied to the tuned circuit. Its value is very large, so it is essentially an ac short at the carrier frequency. Therefore, it is the capacitance of D_1 and C_2 that controls the resonant frequency.

The modulating signal is first passed through a low-pass network R_3-C_5 that provides the amplitude compensation necessary to produce FM. The modulating signal appears across potentiometer R_4. In this way, the desired amount of modulating signal can be tapped off and applied to the phase-shift circuit. The potentiometer acts as a deviation control. The higher the modulating voltage, the greater the frequency deviation. The modulating signal is applied to the varactor diode through capacitor C_4. The RFC has a high impedance at the carrier frequency to minimize the loading of the tuned circuit which reduces *Q*. With zero modulating voltage, the value of the capacitance of D_1 along with capacitor C_2 and the inductor *L* set the resonant frequency of the tuned circuit. The PM output across *L* is inductively coupled to the output.

When the modulating signal goes negative, it subtracts from the reverse bias of D_1. This increases the capacitance of the circuit and lowers the reactance, making the circuit appear capacitive. Thus a leading phase shift is produced. The parallel *LC* circuit looks like a capacitor to the output resistance of the RF amplifier, so the output lags the input. A positive-going modulating voltage will decrease the capacitance, and thus the tuned circuit will become inductive and produce a lagging phase shift. The *LC* circuit looks like an inductor to

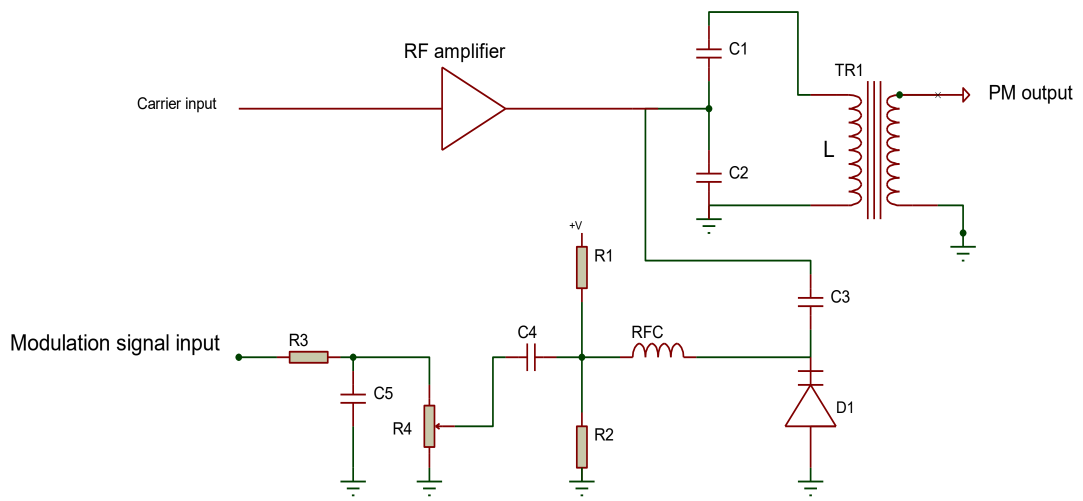


Figure 5.4 One form of phase modulator

output resistance of the RF amplifier, so the output leads the input. The result at the output is a relatively wide phase shift which, in turn, produces excellent linear frequency deviation.

Although phase modulators are relatively easy to implement, they have two main disadvantages. First, the amount of phase shift they produce and the resulting frequency deviation is relatively low. For that reason, the carrier is usually generated at a lower frequency and frequency multipliers are used to increase the carrier frequency and the amount of frequency deviation. Second, all the phase-shift circuits described above produce amplitude variations as well as phase changes. In the simple phase-shift circuits of Figure 5.1, the phase shifters are all voltage dividers. When the value of one of the components is changed, the phase shifts but the output amplitude changes as well. This is also true of the tuned-circuit phase shifter. As a result, some means must be used to remove the amplitude variations.

Both these problems are solved by feeding the output of the phase modulator to class C amplifiers used as frequency multipliers. The class C amplifiers eliminate any amplitude variations, while at the same time they increase the carrier frequency and the deviation to the desired final values. Class C frequency multipliers will be discussed in the next chapter.

Phase modulation is the third and last of the analogue modulation methods. It is also an angle modulation like FM, and is very similar to it in lots of ways. It is obtained by altering the phase constant, $\emptyset_c$, of the carrier in proportion to a baseband voltage shown, as before, on the carrier equation as

$$v_c = E_c\cos(\omega_c t + \underbrace{\emptyset_c}_{\text{PM}})$$

Modulation changes $\emptyset_c$ by $\delta\emptyset$ where

$$\delta\emptyset = K_p v_m \qquad \text{Definition of phase modulation}$$

The analysis is very similar to that for FM and so will not be given in full detail here.

5.5 General Analysis

Assume $\emptyset_c = 0$ when $v_m = 0$. then

$$v_{PM} = E_c\cos\left(\omega_c t + K_p v_m\right)$$

When v_m is a sinusoidal written as

$$v_m = E_m\cos\omega_m t$$

The modulated signal becomes

$$v_{PM} = E_c\cos\left(\omega_c t + K_p E_m\cos\omega_m t\right)$$

Which is more usually written as

$$v_{PM} = E_c\cos(\omega_c t + \beta_p\cos\omega_m t)$$

Where β_p is the **phase modulation index.**

Then, by analogy to FM, there will be a maximum **phase deviation.** $\Delta\emptyset$, where

$$\Delta\emptyset = \beta_p = K_p E_m$$

But, because there is a 2π ambiguity, that is

$$\Delta\emptyset + 2\pi = \Delta\emptyset$$

Useful phase changes must lie between $+\pi$ and $-\ \pi$.

Expansion of the equation in β_p above gives

$$v_{PM} = E_c[\cos\omega_c t\cos(\beta_p\cos\omega_c t) - sin\omega_c t\sin(\beta_p\cos\omega_m t)]$$

Which is very similar to the corresponding equation for FM and leads to narrow-band and wideband modulations = NBPM and WBPM.

Example 5.1

The carrier signal $v_c = 2\cos(4\pi \times 10^5)t$ is phase modulated by the information bearing signal $v_m = 3\cos(6\pi \times 10^3)t$ the modulation sensitivity is $\frac{4}{3}\ rad/\mathrm{V}$.

i. Find the phase deviation.
ii. Find the bandwidth of the phase-modulated signal .
iii. For each frequency component within the bandwidth of the phase-modulated signal. write the amplitude and the phase as well as the powers.
iv. Sketch the frequency and phase spectrum of the phase modulation.
v. What is the power in the carrier.

University of Ibadan, TEL412- Communication system I 2000/2001 B.Sc. degree Exam.

Solution

(i) $\text{Phase deviation} = \Delta\emptyset = B_p = K_p E_m$

Given that $v_c = 2\cos(4\pi \times 10^5)t$, $v_m = 3\cos(6\pi \times 10^3)t$ and $K = 4/3$

$$\Delta\emptyset = B_p = \frac{4}{3} \times 3 = \mathbf{4}$$

(ii) $Bandwidth = 2(\beta + 1)f_m$ = $2(4+1) \times 3 \times 10^3$

= $\mathbf{30\ kHz}$

(iii)

Frequency at $\beta = 4$	Amplitude at $\beta = 4$	Phase radian	Power (W)
J_0 200	-0.7942	0	0.3154
$J_1 \begin{cases} 203 \\ 197 \end{cases}$	$\begin{cases} +0.132 \\ +0.132 \end{cases}$	$-\pi/2$ $-\pi/2$	0.0174
$J_2 \begin{cases} 206 \\ 194 \end{cases}$	$\begin{cases} -0.7282 \\ -0.7282 \end{cases}$	0 0	0.5303
$J_3 \begin{cases} 209 \\ 191 \end{cases}$	$\begin{cases} +0.8604 \\ +0.8604 \end{cases}$	$-\pi/2$ $-\pi/2$	0.7403
$J_4 \begin{cases} 212 \\ 182 \end{cases}$	$\begin{cases} +0.5622 \\ +0.5622 \end{cases}$	0 0	0.3161
$J_5 \begin{cases} 215 \\ 179 \end{cases}$	$\begin{cases} -0.2642 \\ -0.2642 \end{cases}$	$-\pi/2$ $-\pi/2$	0.0698
$J_6 \begin{cases} 218 \\ 176 \end{cases}$	$\begin{cases} -0.0982 \\ -0.0982 \end{cases}$	0 0	0.0096
$J_7 \begin{cases} 221 \\ 173 \end{cases}$	$\begin{cases} +0.0304 \\ +0.0304 \end{cases}$	$-\pi/2$ $-\pi/2$	0.0009
$J_8 \begin{cases} 224 \\ 170 \end{cases}$	$\begin{cases} +0.0080 \\ +0.0080 \end{cases}$	0 0	0.0001
$J_9 \begin{cases} 227 \\ 167 \end{cases}$	$\begin{cases} -0.0018 \\ -0.0018 \end{cases}$	$-\pi/2$ $-\pi/2$	0.0000
$J_{10} \begin{cases} 230 \\ 164 \end{cases}$	$\begin{cases} -0.0004 \\ -0.0004 \end{cases}$	0 0	0.0000
Total Phase Power			$\mathbf{1.9998}$

(iv)

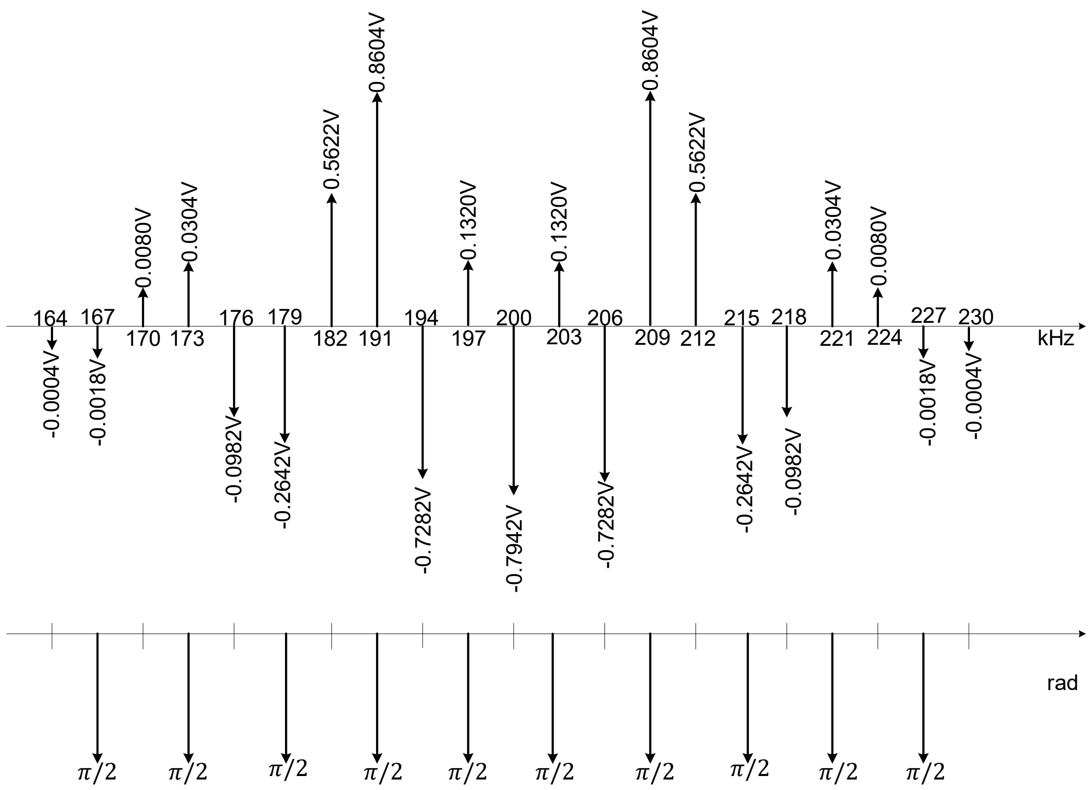


Figure 5.5 Frequency and phase spectrum of the modulated signal

(v) Power in the carrier theoretically $= \frac{V_c{}^2}{2R} = \frac{2^2}{2} = \mathbf{2\ W}$

But the power in the phase modulated wave using Bessel table is $\mathbf{1.9998\ W.}$

This shows that the sum of power in each amplitude component is equal to the carrier power.

5.6 Narrow-Band Phase Modulation

The general expression for NBPM is

$$v_{NBPM} = E_c \cos \omega_c t - \beta_p E_c \sin \omega_c t \cos \omega_m t$$
$$= E_c \cos \omega_c t - \frac{1}{2}\beta_p E_c[\sin(\omega_c - \omega_m)\, t + \sin(\omega_c + \omega_m)\, t]$$

Which is superficially similar to the expression for NBFM. However, further analysis shows the difference because

$$v_{NBPM} = E_c \cos \omega_c t - \frac{1}{2}\beta_P E_c \{\cos [\frac{\pi}{2} - (\omega_c - \omega_m)t] + \cos[\frac{\pi}{2} - (\omega_c + \omega_m)t]\}$$

$$= E_c \cos \omega_c t - \frac{1}{2}\beta_P E_c \{\cos [(\omega_c - \omega_m)t - \frac{\pi}{2}] + cos[(\omega_c + \omega_m)t - \frac{\pi}{2}]\}$$

$$= E_c \cos \omega_c t + \frac{1}{2}\beta_P E_c \{\cos [(\omega_c - \omega_m)t - \frac{\pi}{2} + \pi] + \cos[(\omega_c + \omega_m t) - \frac{\pi}{2}]\}$$

$$= E_c \cos \omega_c t + \frac{1}{2}\beta_P E_c \{\cos\left[(\omega_c - \omega_m)t + \frac{\pi}{2}\right] + \cos\left[(\omega_c + \omega_m)t + \frac{\pi}{2}\right]\}$$

This is, of course, just a carrier and a single sideband pair whose amplitude spectrum is exactly the same as for AM and NBFM with bandwidth $2f_m$.

Obviously the phase spectrum is different and this is where the uniqueness of

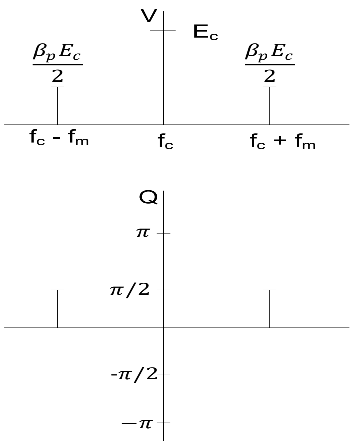


Figure 5.6 Amplitude and phase spectra for NBPM

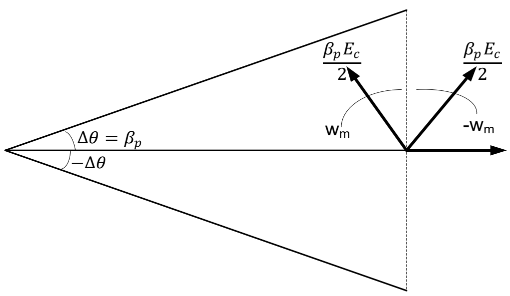


Figure 5.7 NBPM in stationary phasors

5.7 Wide Band Phase Modulation

The general expression for PM above is

$$v_{PM} = E_c[\cos \omega_c t cos(\beta_p \cos \omega_m t) - \sin \omega_c t \sin(\beta_p \cos \omega_m t)]$$

As for FM, Bessel function expressions are needed to expand this further. These can easily be obtained from those for FM (Section 5.6) by writing $\cos\omega_m t = \sin(\frac{\pi}{2} - \omega_m t)$etc., and are

$$cos(\beta_p \cos\omega_m t) = J_0(\beta_p) - 2J_2(\beta_p)\cos 2\omega_m t + 2J_4(\beta_p) - \cdots$$
$$sin(\beta_p \cos\omega_m t) = 2J_1(\beta_p)\cos\omega_m t - 2J_3(\beta_p \cos 3\omega_m\, t) + \cdots$$

Leading to the full expression for WBPM showing the signs of all side frequencies;

$$\begin{aligned}
&v_{WBPM} = \cdots + J_4(\beta_P)\cos(\omega_c - 4\omega_m)t \\
&+J_3(\beta_P)\cos\left[(\omega_c - 3\omega_m)t - \frac{\pi}{2}\right] - J_2(\beta_P)\cos(\omega_c - 2\omega_m)t \\
&-J_1(\beta_P)\cos\left[(\omega_c - \omega_m)t - \frac{\pi}{2}\right] + J_0(\beta_P)\cos\omega_c\, t \\
&-J_1(\beta_P)\cos\left[(\omega_c + \omega_m)t - \frac{\pi}{2}\right] - J_2(\beta_P)\cos(\omega_c + 2\omega_m)t \\
&+J_3(\beta_P)\cos\left[(\omega_c + 3\omega_m)t - \frac{\pi}{2}\right] + J_4(\beta_P)\cos(\omega_c + 4\omega_m)t - \cdots
\end{aligned}$$

The amplitude are of course exactly the see the same as for WBFM (for the same β value) with the same bandwidths and again the difference lies only in the phases which are not the same.

5.8 Comparison of PM and FM

These two methods are closely related and, in some respects, very similar. They both have the effect of changing carrier phase as a function of baseband voltage. This can be seen clearly from their general expressions

$$v_{FM} = E_c\cos(\omega_c t + \beta\int v_m dt)$$
$$v_{PM} = E_c\cos(\omega_c t + \beta_p v_m)$$

Thus it is not surprising that the fuller expressions for NBPM and WBPM are very similar to those for NBFM and WBFM. These show that for both narrow and wide modulations the amplitude spectra are the same, the difference between PM and FM lying in the phases of the harmonic sidebands.

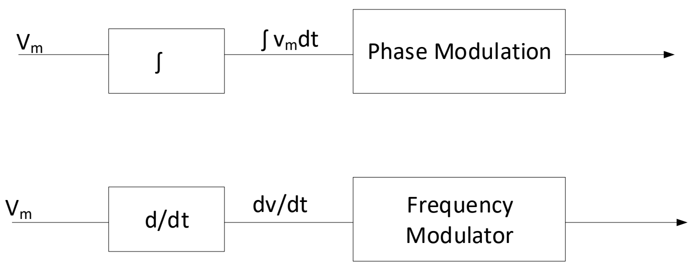


Figure 5.8 FM and PM obtained from each other

The similarity gives simple ways of obtaining one modulation from the other by integrating or differentiating the baseband first. Fig. 5.8 illustrates this.

The relationships can easily be shown analytically: if

$$v_m = E_m \cos \omega_m t$$

Then

$$\int v_m dt = \frac{E_m}{\omega_m} \sin \omega_m t$$

And

$$v_{PM} = E_c[cos\omega_c t \cos\left(\frac{\beta_p}{\omega_m} sin\omega_m t\right) - sin\, \omega_c t \sin\left(\frac{\beta_P}{\omega_m} sin\omega_m t\right)]$$

$$= E_c[\cos \omega_c t \cos (\beta \sin \omega_c t) - \sin \omega_c t \sin (\beta \sin \omega_m t)]$$

$$v_{PM} = (\text{writing } \frac{\beta_P}{\omega_m} = \beta)$$

With a similar analysis for differentiation.

As we shall see, this method is a convenient way to get NBFM using an NBPM modulator because the circuitry is simpler and does not need tuned circuits.

For simple baseband waveform the difference between FM and PM can be seen clearly (see Fig. 4.16) but it is not all easy to see with a continuously varying baseband. Try sketching them for a sinusoidal baseband. This does bring home the great similarity between the two methods.

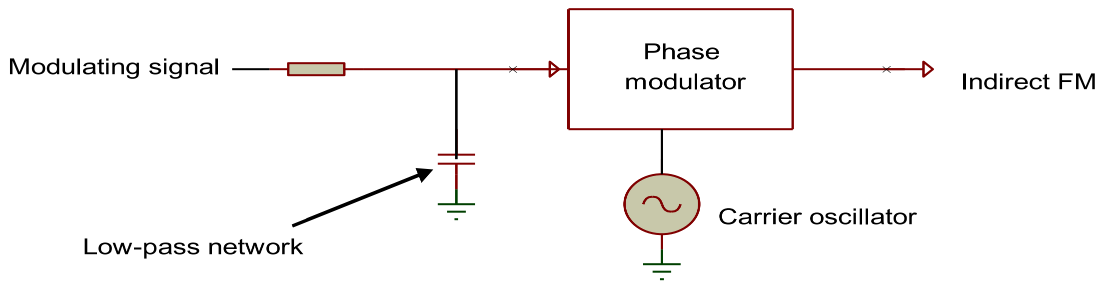


Figure 5.9 A low-pass filter compensates for higher phase shift and frequency deviation at the higher modulating frequencies to produce indirect FM.

5.9 Converting PM to FM

To make PM compatible with FM, we must compensate for the deviation produced by the frequency changes in the modulating signal. This is illustrated in Fig. 5.9. This low-pass filter causes the higher modulating frequencies to be attenuated in amplitude. Although the higher modulating frequencies will produce a greater rate of change and thus a greater frequency deviation, this is offset by the lower amplitude of the modulating signal, which will produce less phase shift and less frequency deviation. This network compensates for the excess frequency deviation caused by higher modulating frequencies. The result is an output that is the same as an FM signal. The FM produced by a phase modulator is called indirect FM.

Although both FM and PM are widely used in communications systems, most angle modulation is PM. The reason for this is that acrystal oscillator with higher frequency accuracy and stability can be used to produce the carrier. In FM, crystal oscillators cannot, in general, be frequency-modulated over a very wide range. However, the crystal oscillator can drive a phase modulator that can produce the desired FM. Further, most phase modulators are simpler to implement than frequency modulators. You will see this where practical FM and PM circuits are discussed in Chapter 4.

5.10 Use of PM

There are two common uses:

1. For NBFM generation in such applications as CB and mobile radio.
2. As WBPM for digital signals in BPSK as described in the next chapter

The first needs simpler circuitry (i.e. no tuned circuits) than the direct f.m. method for applications where only narrow-baseband modulation is required. The second can also be done simply- using a multiplier- and has very wide applications.

5.11 Generation and Demodulation of PM

Narrow-baseband modulation can be obtained from a simple RC low-pass filter as in Fig. 5.8. with C consisting of a fixed capacitor in series with a varactor (providing $\delta C \propto v_m$).

The unmodulated carrier goes in and a phase-modulated one comes out. The analysis is as follows:

$$\frac{V_{NBPM}}{V_c} = \frac{\frac{1}{j\omega C}}{R + \frac{1}{j\omega C}} \qquad c = c_O + c_r$$

$$= \frac{1}{1 + j\omega CR}$$

$$= \frac{1}{(1 + \omega^2 C^2 R^2)^{\frac{1}{2}}} < \theta \ (\tan\theta = \omega CR)$$

If ωCR is kept small (< 0.2) then $\theta \simeq -\omega CR$. So if C changes by δC, then θ changes by $\delta\theta$ where

$$\theta + \delta\theta = \omega(C + \delta C)R$$

Therefore

$$\delta\theta = -\omega_c RKv_m$$

That is

$$\delta\theta \propto v_m$$

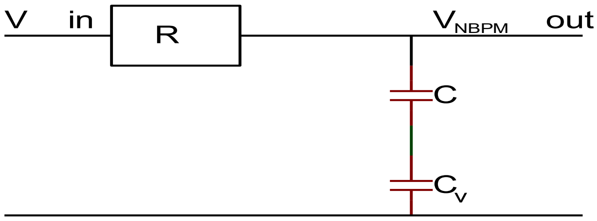


Figure 5.4 RC NBPM modulator

And thus linear phase modulation has been achieved. Of course the assumptions are that the fractional change in $\delta\theta$ must be small ($< 11^o$) so that only NBPM can be achieved. But, because the modulation is linear, it can be used for analogue basebands and in particular, with a pre-integrator, to give NBFM as already mentioned above.

By contrast WBPM is really only used for discrete basebands with fixed voltage levels between which it is switched. The simplest of these is the binary PSK method described in the next chapter and its detailed description will be left to then. There the phase is either 0^o(representing a one) or 180^0 (representing a zero).

Demodulation of continuous PM is rarely needed as it is not used as a mode of transmission. And for the discretion of the demodulation of binary PSK see the next chapter.

5.13 Summary

Phase modulation occurs when θ_c changes linearly proportional to v_m, that is

$$v_{PM} = E_c\cos\left(\omega_c t + K_P v_m\right)$$

Where $K_P$is the **phase modulation sensitivity**.

For single sinusoidal basebands

$$v_{PM} = E_c\cos\left(\omega_c t + \beta_P \cos\omega_m t\right)$$

Where β_P is the **phase modulation index** and $\beta_P = K_p E_m$.

NBPM occurs when $\beta_p \leq 0.2$ and gives one sideband pair, both negative, as

$$v_{NBPM} = E_c \cos \omega_c t - \frac{1}{2}\beta_p E_c \sin(\omega_c - \omega_m)\, t - \frac{1}{2}\beta_p E_c \sin(\omega_c + \omega_m)\, t$$

Its **stationary phasor** diagram shows clearly how it differs from FM and AM.

WBPM gives $\beta + 1$ sideband pairs with the same amplitude spectrum as WBFM bit with sign and phase differences (of $\frac{\pi}{2}$). The full expression for WBPM involves Bessel function and is given in **Section 4.17**.

PM and FM are very similar and one can be obtained from the other by integrating or differentiating the baseband before modulating.

PM is a convenient way of getting NBFM (e.g. for mobile radio). It is also widely used for the modulation by frequency translation of digital signals where its multilevel capability os of great importance.

5.14 Conclusion

This is a good lead-in as any to the next chapter which describes a very important class of usage for these analogue modulation methods-the simple, but widely used, shift keying methods. Read on.

5.15 Chapter Review Problems

5.1. For f_m =3 kHz, E_m =3.0 V, β =4.0, E_c =2.0 V *and* f_c =200 kHz, calculate

(i) The phase deviation ($\Delta\emptyset$).

(ii) The phase modulation sensitivity and maximum instantaneous phases.

(iii) The bandwidth.

(iv) The amplitudes and signs of the carrier and all the side frequencies within the bandwidth.

[compare with Problem **4.1**.]

5.2. A square wave baseband at 2.5 kHz repetition frequency and $\pm$ 6 V magnitude is frequency modulated on to a carrier at 25 kHz with modulation sensitivity of 2500 rad V^{-1}.

(i) What is the modulation index?

(ii) What is the phase deviation?

(iii) What is the bandwidth?

(iv) Sketch the modulated waveform

(v) Draw the instantaneous amplitude spectrum at the extremities of deviation,

(vi)Draw the p.m. spectrum showing amplitude and sign.

[compare with Problem **4.5**.]

5.5. Discuss the relative advantages of using FM and PM for

(i) Speech basebands
(ii) Digital basebands
(iii) Keying modulations
(iv) Linearity of modulation and demodulation
(v) Obtaining one from the other
(vi) Wideband modulating.

5.6. Sketch the statinary phasor diagram for NBPM at $\beta = 0.1$. attempt also a sketch at $\beta = 1.0$.

5.7 (a) Explain briefly why you need modulate your information bearing signal.

(b) The carrier signal v_e = 3cos (6πt x 10^5) is phase modulated by the information bearing signal v_e=2cos(8πt x 10^3). The modulation sensitivity is 2.50 rad/V.

(i) Find the phase deviation

(ii) Find the Bandwidth of the phase-modulated signal

(iii) For each frequency component within the bandwidth of the phase-modulated signal. Write the amplitude and the phase.

(University of Ibadan, TEL412- Communication system I 2003/2004 BSc degree Exam.

5.9 A 9 W, 120 MHz higher frequency signal is phase modulation with 20 KHz lower frequency signal such that the peak frequency deviation is 11 KHz. The resistance level is 3 Ω. Determine:

(i) the amplitude of the higher frequency signal.

(ii) The modulation index.

(iii) The Carson's bandwidth.

(iv) The modulation sensitivity if the amplitude of the lower frequency signal is 5.68 V.

(v) State the amplitude of all components in the resulting phase modulated signal. (Quote the frequency and the amplitude of each frequency component). Sketch the frequency and phase spectrum of the phase modulated signal.

(University of Ibadan, TEL412- Communication system I 2004/2005 BSc degree Exam).

5.10. The carrier signal $V_c = 4\cos(4\pi \times 10^6)t$ is phase modulated by the information bearing signal $V_m = 6\cos(12\pi \times 10^3)t$ the modulation sensitivity is $\frac{2}{3}$ rad/V.

Find the phase deviation

Find the bandwidth of the phase-modulated signal .

For each frequency component within the bandwidth of the phase-modulated signal. write the amplitude and the phase as well as the powers:

Sketch the frequency and phase spectrum of the phase modulation.

What is the power in the carrier.

(University of Ibadan, TEL412- Communication system I 2004/2005 BSc degree Exam).

CHAPTER 6

SIGNALS AND SYSTEMS

6.0 Introduction to Signal Processing

Signal processing is the science of analyzing and interpreting the signals produced by physical quantities. In engineering, signal carrying information are high energy microwave pulses in high energy signals necessary for performing machine tool operations e. g Telephone, radio signals or the pulse that dictate the operation of a digital computer. Signal may be the cause of an event or the consequence of an action. In fact, the characteristics of a signal may be out of a broad range of shapes, amplitudes, time duration and perhaps other physical properties. This signal may be expressed in analytical form call an expression. In other cases, the signal may be given in a graphical form. A signal is an electric quantity such as voltage, current or field strength whose modulation represent codes of information about the source from which it comes.

Often, an electrical engineer is concerned with the design, analysis and synthesis of systems. Signals carry information energy from one point of a system to another. In other words, the engineer is concerned with analysis, detection and processing of signals. The engineer will have to consider signals of many different amplitudes, shapes and time duration that can vary significantly, Examples: Radar systems, high echoes (weak and hidden noise) signal, waveshape and recurrent rate signal in patient's heart cardiogram (heart signal).

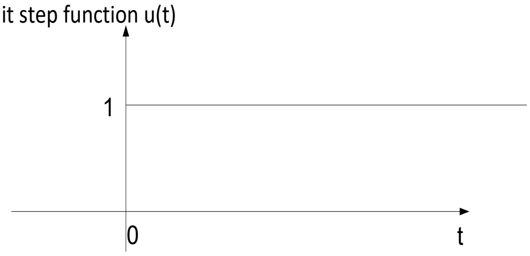
In electrical engineering, signals are encountered as variations of current and voltage versus time. Signals can be converted from one form to the other by means of a transducer e.g. Pressure, conductivity, light intensity, speed, acceleration can be converted to voltage or current variations (i.e. electrical signal). It is important for the engineer to understand the different classes of signals and the various mathematical means for their description. Only with a good understanding of the signals will Engineers be able to carry out proper analysis of these signals bearing in mind that Analog system process continuous or analog signals while Digital systems process digital signals.

6.1 Catalog of signals and their mathematical description

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. Next, we provide a brief description of the mathematical models of signals that are frequently used to characterize many of the physical channels that we encounter in practice.

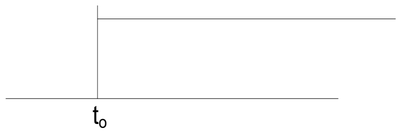
6.1.1 Non periodic continuous signals

(a) Unit step function u(t)

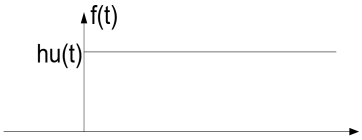


$u(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$

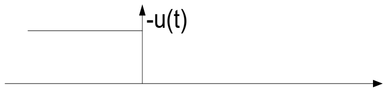
Shifted unit step functions



$u(t - to) = \begin{cases} 1 & t \geq t_0 \\ 0 & t < t_0 \end{cases}$

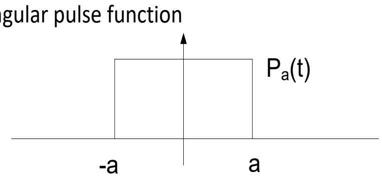


f(t) = hu(t)= $\begin{cases} h & t \geq o \\ o & t < o \end{cases}$

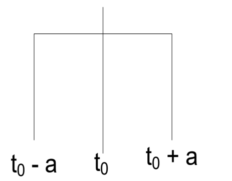


f(t) = -u(t)= $\begin{cases} 1 & t \leq o \\ o & t > o \end{cases}$

(b) Rectangular pulse function



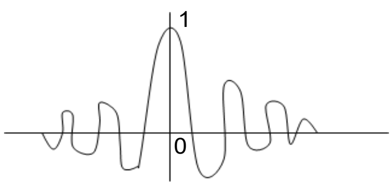
f(t) = $P_a(t)$= u(t + a) – u(t – a) $= \begin{cases} 0 & |t| \geq a \\ 1 & |t| \leq a \end{cases}$



f(t) = $P_a(t)$= u(t - t_0+ a) – u(t - t_0– a)

$= \begin{cases} 1 & |t - t_0| \\ 0 & |t - t_0| \end{cases}$

(c) Sinc function



$\text{Sinc (t)} = \frac{\sin t}{t}$

$-\delta < t < \delta$

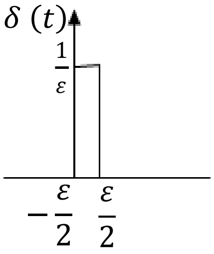
6.1.2 Non periodic discrete signals

The non periodic discrete signals are those signals that doesn't repeat themselves. These signals are:

(a) Delta function

The delta function $\delta(t)$ is also called an impulse or Dirac delta function. It occupies a central place in signal analysis. Many physical quantities such as point sources, point charges, voltage or current sources acting for very short times can be modeled as delta functions.

Theoretically



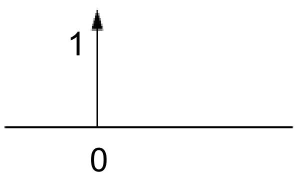
$\delta\ (t - t_0) = \ o$

$t \neq t_0$

f(t) = $\int_{t_1}^{t_2} \delta(t - t_0)dt = 1$

$\int_0^{\infty} \delta(t)dt = 1$ Area under the delta function is unity

Practically



$\delta\ (t) = \ o\ ,\ t = 0$

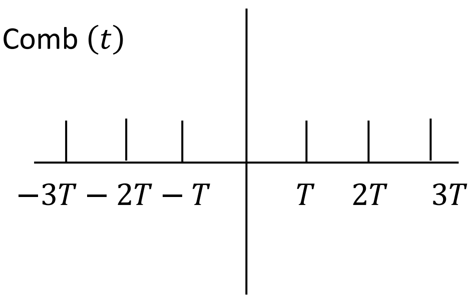
f(t) = $\int_0^{\varepsilon} \delta(t)dt = 1$ 0< t < ε

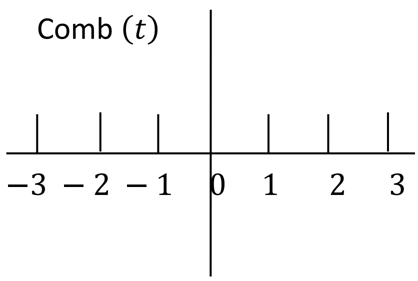
(b) Impulse function can be used to extract a function $f(t)$ from the specific value that exists at the time where $f(t)$ is a continuous function at t_0.

(c) Comb function

This is any array of delta functions that are spared T units apart and that extend from $-\infty$ to $+\infty$

Comb $(t_1) = \sum_{-\infty}^{\infty} \delta(t - kT)\ k = 0, t_1, t_2.$

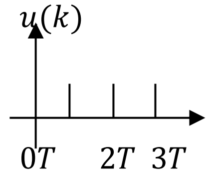




(d) Discrete step function $\boldsymbol{u(k)}$

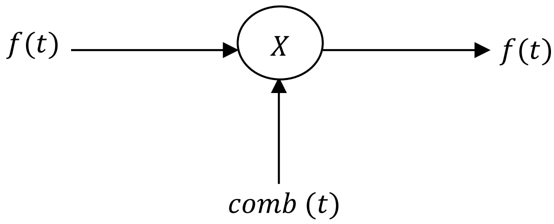
This is the discrete version of $u(t)$. An array of delta functions that extends from 0 to ∞ and that are spaced T units apart.

$$u(k) = \begin{cases} 1 & k = 0, 1, 2 \ldots \\ 0 & k = -u \ldots \ldots \end{cases}$$



(e) Arbitrary sampled function

The comb function can be used in the representation of any continuous function in the sampled or discrete form. The property of the delta function helps us to write.



$$f\ (t)\ \delta\ (t-a) = \ f\ (t=a)$$

$$f(t)d(t-t_o) = \ f\ (t=t_{o)}$$

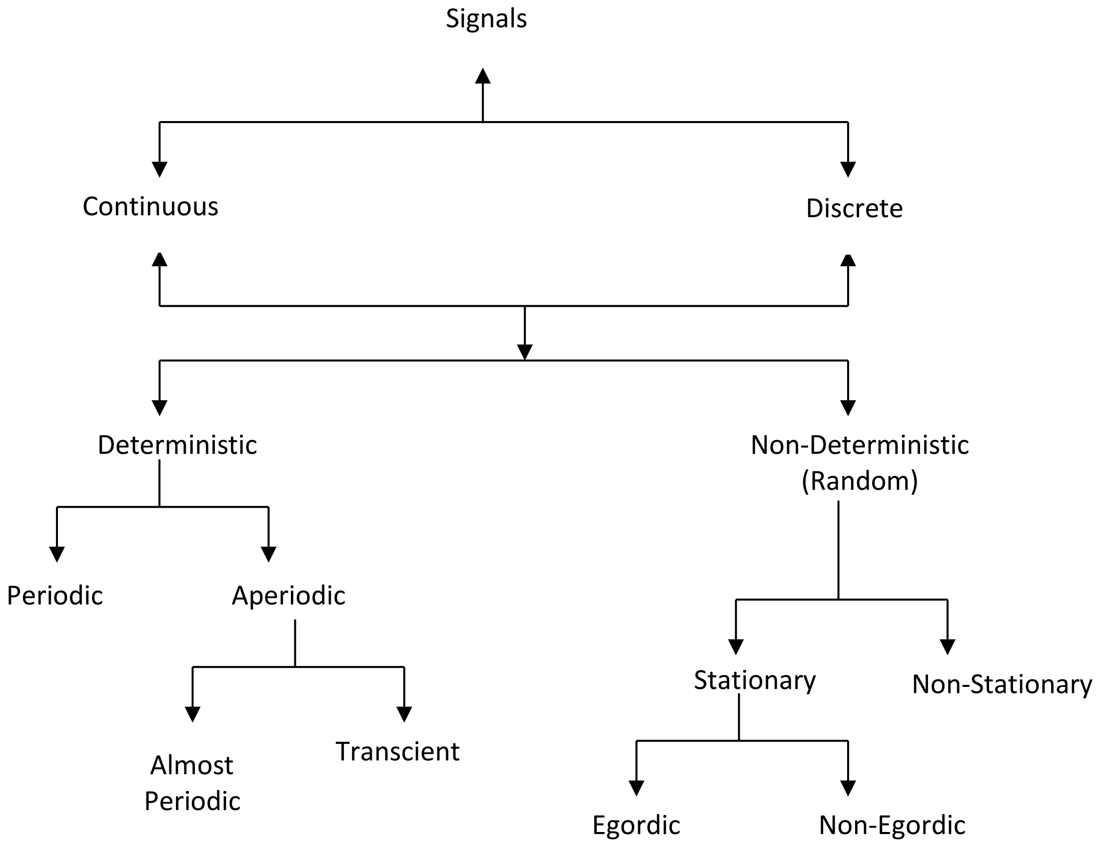
6.2 Signals and Systems

An electronic communication involves the process of generation, transmission and reception of various types of signals. However, because of the following reasons, communication process becomes very difficult:

- They will experience attenuation if the signal is to travel a long distance.
- When the medium of transmission is not in a perfect condition.
- Due to the presence of noise at the receiver input.

In some cases, the required signal strength at the receiver input might not be clearly stronger than the disturbance component seen at the point in the communication chain, left for this, communication process would have been very easy, if not trivial. To come up with an adequate signal processing techniques, which helps us to extract the desired signal from a distorted and noisy version of the transmitted signal, there is need for appropriate understanding of the nature as well as the properties of the desired and undesired signals which are present in various stages of communication system. Hence, here comes our first stage of or study of this area of communication theory. However, it is what note that signals physically exist in the time domain usually express as a function of time. Since signals exist in time domain, it is not too difficult, at least in the majority of the situation of interest to us, to visualize the signal behaviour in the time domain. The signals can even be view in an oscilloscope. Also, it is of important to study the behaviour of signals in the frequency domain or spectral domain which is analysing the signal in terms of its various frequency components i.e. its spectrum. Another powerful tool used to know the behaviour of signals of any kind is the Fourier analysis. Fourier analysis (Fourier series and Fourier transform) helps us in arriving at the spectral description of the pertinent signals.

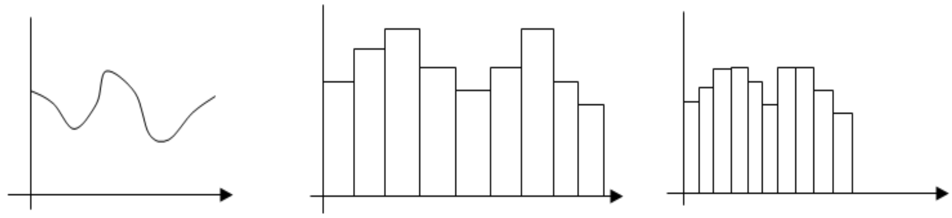
6.3 Digital and Analog signals classification



The communication process can be broadly divided into two types, namely, analog communication (analog signal) and digital communication (digital signal). The classification is mainly based on the nature of message or modulating signal. If the message to be transmitted is continuous or analog in nature, then such a communication process is termed as analog communication.

Alternatively, if the message is discrete or digital in nature, then such a communication process is termed as digital communication. The word "digital" has been used in many different context: such as digital camera, digital phone as well as digital watches etc. While at first glance these three technologies appear quite different, the digital preface means the same thing in each technology. Digital here refers to the discrete resolution of signal (information). For example, a digital watch provides the hour, the minute, and usually the second. Nevertheless, it is not possible to determine the time up to a hundredth of a second on watch a digital watch that expresses time only to the second, in an analog watch, measuring small amounts of time might be difficult, but it is possible in principle to measure the time essentially as accurately as is desired. For cameras, the difference is in the picture. A high extremely small pixels, but the image is still in discrete pieces.

A traditional camera stores the photo in continually-varying intensities on film. In digital phones, or digital music recording, the sound is broken into discrete pieces. Analog phones and LPs can transmit and store continually varying signals. The digital signal can approach the analog signal if the pieces are made arbitrarily narrow, but it will never be completely as smooth as the analog signal. The figure below illustrates the differences between analog and digital signals.



Analog signal continuosly varying   Digital signal-large time division   Digital signal-small time Division.

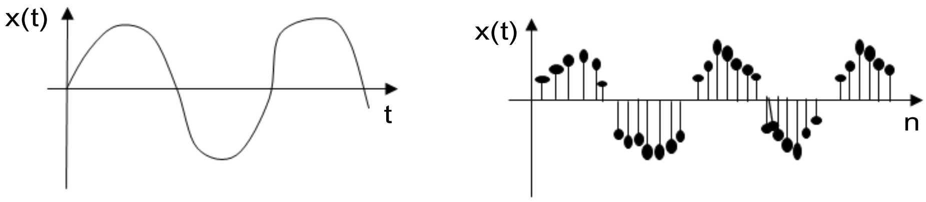
Figure 6.1. Analog and Digital signal.

From the Fig. 6.1, it gives the impression that digital is not as good as analog, but it is not necessarily true. By increasing the number 0 i.e. decreasing the size of the time division in a digital signal, can make the digital signal nearly as smooth as an analog signals and are such less prone to degradation. Each of information in digital signal is a number which is easily distinguished from other numbers. One analogy for a digital signal could be a table of numbers. A comparable analogy for an analog signal would be a graph.

6.4 Types of signal

A time varying phenomenon that can carry information is called signal. Examples of signals are human voice, electrocardiogram, sign language, videos, etc. These are can be classified as continuous time signal, discrete time signal and digital signal, random signals and non-random signals etc.

Continuous-time signal: A continuous-time signal is a signal that can be defined at every instant of time. A continuous-time signal contains values for all real numbers along the x-axis. It is denoted by $x(t)$. Fig. 6.2(a) shows continuous-time signal

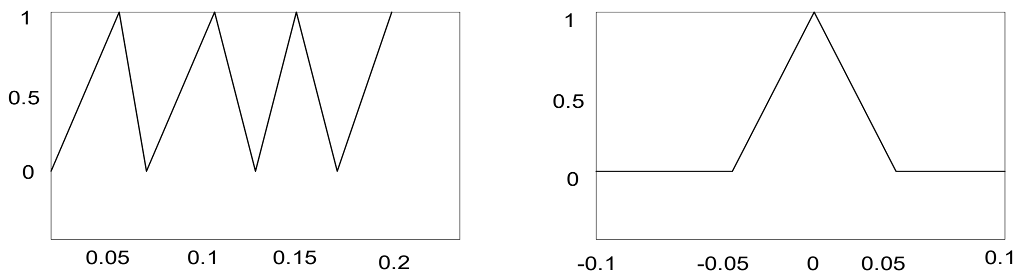


(a) Continous-time signal   (b) Discrete-time signal.

Figure 6.2. Continuous and Discrete time signals

Discrete-time signal - Signals that can be defined at discrete instant of time is called discrete time signal. Basically discrete time signals can be obtained by sampling a continuous-time signal. It is denoted as $x(n)$. Fig. 6.2(b) shows discrete-time signal.

Periodic and Aperiodic signals - A signal is said to be periodic if it repeats after some amount of time $x\ (t+T) = x(t)$, for some value of T. The period of the signal is the minimum value of time for which exactly repeats itself.

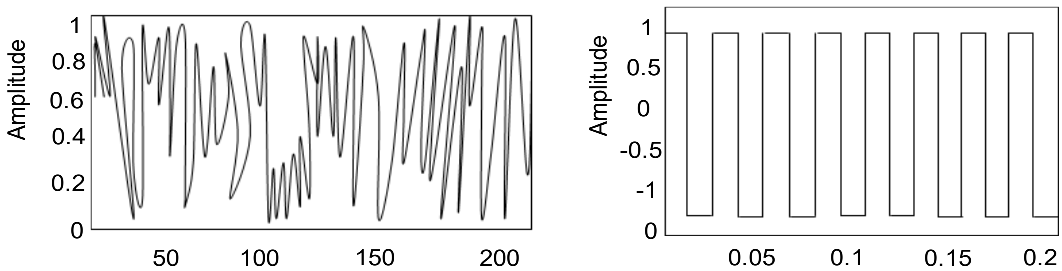


(a) Periodic signal Time (seconds) (b) Aperiodic signal

Figure 6.3. Periodic and Aperiodic signal

Signal which does not repeat itself after a certain period of time is called aperiodic signal. The periodic and aperiodic signals are shown in Figs. 6.3 (a) and 6.3 (b) respectively.

Random and Deterministic signal: A random signal cannot be described by any mathematical function. Whereas a deterministic signal is one that can be described mathematically. A common example of random signal is noise. Random and deterministic signal are as shown below.



(a) Random signal (b) Deterministic signal

Figure 6.4. Random and Deterministic signal

Causal, Non-causal and Anti-causal signal: Signal that are zero for all negative time are called causal signals, while the signals that are zero for all positive value of time are called anti-causal signal. A non-causal signal is one that has non zero values in both positive and negative time. Causal, non-causal and anti-causal signals are shown below in the Figs. 6.5(a), 6.5(b), and 6.5(c) respectively.

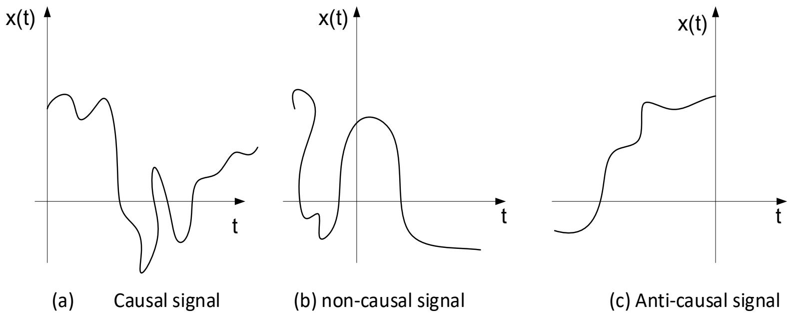


Figure 6.5. Causal, Non-causal and anti-causal signals

Even and Odd signal: An even signal is any signal 'x' which exhibit the features of the form $x(t) = x(-t)$. On the other hand, an odd signal is a signal where $x\ (t) = -x(-t)$. Even signals are symmetric around the vertical axis that they can easily be spotted.

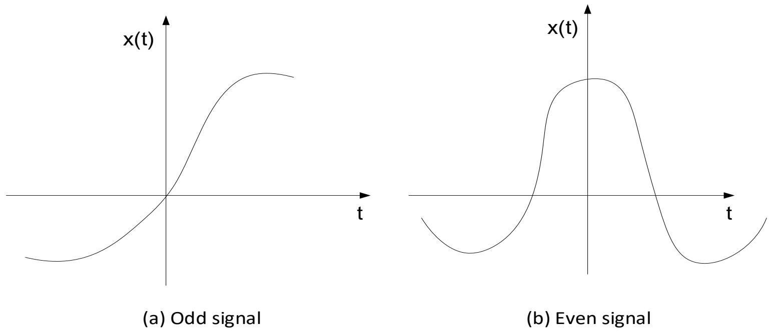


Figure 6.5. Odd and Even signal

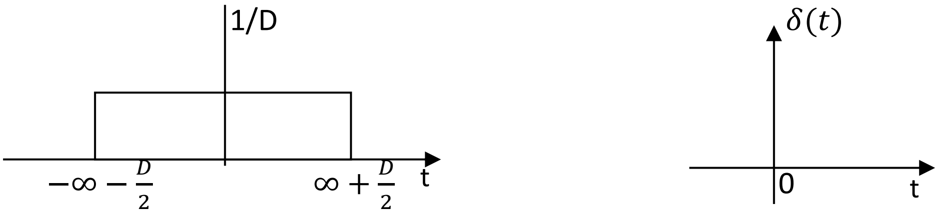
An even signal is one that is invariant under the time scaling $t\ \rightarrow -t$ and odd signal is one that is invariant under the amplitude and time scaling $x(t) \rightarrow -x(t)$. A simple way of visualizing even and odd signal is to imagine that the ordinate $[x(t)]$ axis is a mirror. For even signals, the part of $x(t)$ for $t\ < 0$ are mirror images of each other.
In case of an odd signal, the same two part of the signals are negative mirror images of each other. Some signals are odd, even and neither odd nor even. But any signal $x\ (t)$ can be expressed as a sum of its even and odd parts such as $x(t) =\ x_e(t) + x_0(t)$ or we can say that every signal is composed of the addition of an even part and odd part. The even and odd parts of a signal $x(t)$ are

$$x_e(t) = \frac{x_e(t) + x(-t)}{2} \text{ and } x_o(t) = \frac{x_0(t) - x(-t)}{2}$$

Here, $x_e\ (t)$ denotes the even part of signal $x\ (t)$ and $x_o(t)$ denotes the odd part of signal $x\ (t)$. Figs. 6.5 (a) and (b) shows the odd signal and even signal respectively.

Impulse signal: The diract delta function or unit impulse are often referred to as the delta function. This is the function that defines the idea of a unit impulse in continuous-time. Informatively, this function is one that is infinities mainly narrow, infinitely tall, yet integrates to one. Perhaps the simplest way to visualize this as a rectangular pulse from $-\infty - \frac{D}{2}$ to $\infty + \frac{D}{2}$ with a height of 1/D. As we take the limit of this setup as D approaches 0, we see that the width tends to zero and the height tends to infinity as the total area remains constant at one. The impulse function is often written as $\delta(t)$.

$$\int_{\infty}^{\infty} \delta(t)\, dt = 1$$



(a) Dirac delta function (b) Unit impulse

Figure 6.6. Delta and impulse signals

For the fact that it is quite difficult to draw something that is infinitely tall, we represent the Dirac with an arrow centered at the point it is appplied. The dirac delta function and unit impulse are as shown in the Figs. 6.6 (a) and (b) respectively.

Real and complex exponential signal: Exponential signal is of two types. These two types of signals are real exponential and complex exponential signal.
A real exponential signal is defined as: $x\,(t) = Ae^{\propto t}$ where both A and $\propto$ are real. Depending on the value of " $\propto$ ", the signals will be different. If $'\propto'$ is positive, the signal $x\,(t)$ is a growing exponential and if $\propto$ is negative, then the signal $x\,(t)$ is a decaying exponential. For $\propto = 0$, signal $x\,(t)$ will be constant. Fig. 6.7 shows a dc signal, exponentially growing signal and exponentially decaying signal respectively.

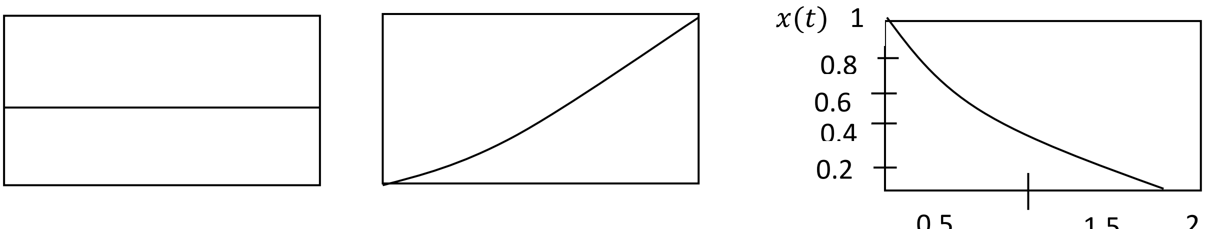


Figure 6.7. Real exponential signals

Complex exponential signals have both real and imaginary parts; in complex exponential, we know that the cosine function is the real part of a complex exponential signal. Complex exponential make dynamic systems analysis relatively sine, thus we often analyze a signals response in terms of complex exponential. Since any measurable quality is real-valued, taking the real part of the analytical result based on complex exponential will result in a cosine function. Thus, cosine becomes a natural way to express signals which vary sinusoidaly.

Sinusoidal signals: Sinusoidal signals are represented in terms of sine and/or cosine functions. In general, we represent sinusoidal as cosine functions. Our general expression for a sinusoidal signal is: $v_{(t)} = V_p \cos(\omega t + \theta)$. Where V_p is the zero to peak amplitude of the sinusoidal, w is the radian frequency of the sinusoidal whose unit is radian per seconds and θ is the phase angle of the sinusoidal in either radians or degrees (recalled that $2\pi = 360°$). A representative plot of a sinusoidal signal is provided in Fig. 6.8. In Fig. 6.8, the frequency of the sinusoidal is indicated as a period of the signal (the period is defined as the shortest time interval at which the signal repeats). The radian frequency of a sinusoidal is related to the period by: $\omega = 2\pi T$.

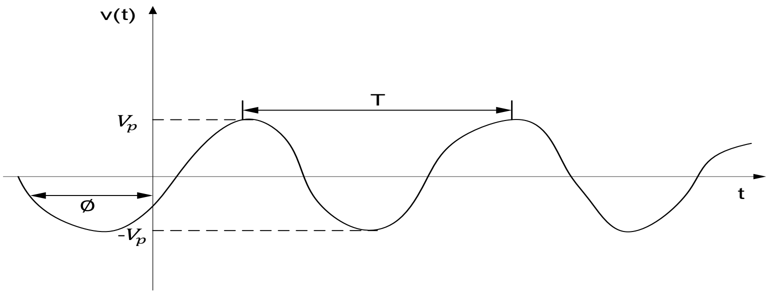


Figure 6.8. Arbitrary sinusoidal signal

The frequency of a sinusoidal signal is alternately expressed in unit of Hertz (Hz). A Hertz is the number of cycle which the sinusoidal goes through in one second. Thus, Hertz correspond to cycles/second. The frequency of a signal in Hertz is related to the period of the signal by: $f = \frac{1}{T}(\text{Hz})$.

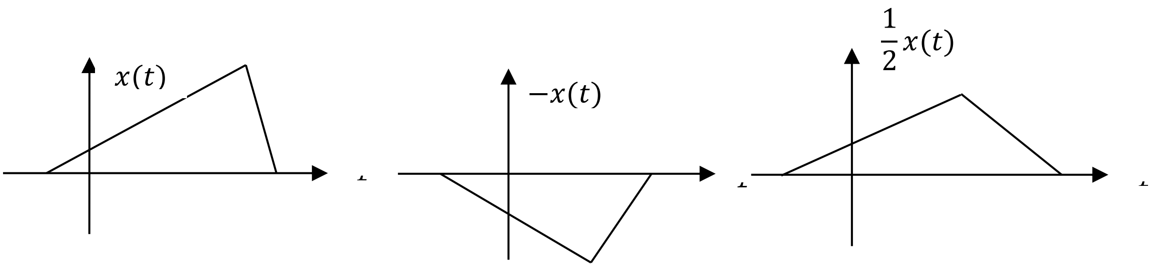
Radian frequency relate to frequencies in Hertz by $f = \frac{\omega}{2\pi}$ Hz
Although frequencies of signals are often expressed in Hertz, it is not a unit which lends itself to calculations. Thus, all our calculations will be performed in radian frequency if given a frequency in Hertz (Hz), it should be converted to Radian/second before any calculations are based on this frequency.

Energy and power signals: A signal with finite energy is an energy signal. An exponentially decaying signal that exists only for t > o is an energy signal. The periodic signals are power signals, and since these signals are everlasting, they have infinite energy. Since the periodic signals have infinite energy per second. To summarize, a power signal has infinite energy and a finite power, and an energy signal has a finite energy and zero power, where the average power is computed as energy over infinite time. It is clear that a signal cannot be both a power signal and an energy signal. It has to be either of the two. But it is possible for a signal to be neither of the two. For instance, a ramp signal is neither a power signal nor an energy signal. A ramp signal has infinite power and infinite energy. Such a signal exists in theory, but not in real world. If an everlasting exponential signal is defined as $exp(at)$i.e. e^{at}, it is neither an energy signal nor a power signal, as long as 'a' is a positive or a negative real value. If the value of 'a' is either zero or imaginary, then the the signal is a power signal. If the value of a is zero, then the signal is a dc signal, and if the value of 'a' is imaginary, then the signal is an alternating signal.

6.5 Signal properties

There are some important properties of signal such as amplitude scaling, time-scaling and time-shifting we shall consider in this section.

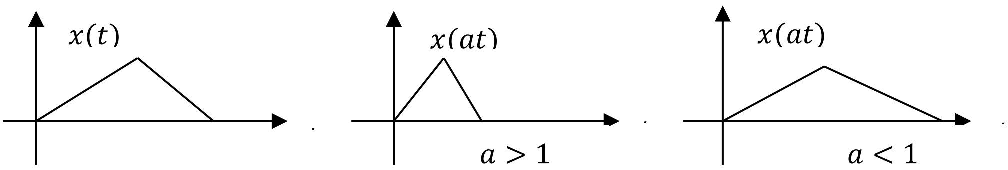
Amplitude scaling: Consider a signal $x(t)$ which is multiplying by a constant 'A' and this can be indicated by a notation $x(t) \rightarrow Ax(t)$. For any arbitrary 't', this multiplies the signal value $x(t)$ by a constant 'A'. This is called amplitude scaling. If the amplitude-scaling factor is negative then it flips the signal with the t-axis as the rotation axis of the flip. If the scaling factor is -1, then only the signal will be flip. This is shown in the Fig. 6.9 (a), (b) and (c)



(a) A signal $x(t)$ (b) A signal $x(t)$ scale by -1 (c) A signal $x(t)$ scale by ½

Figure 6.9. Amplitude scaling

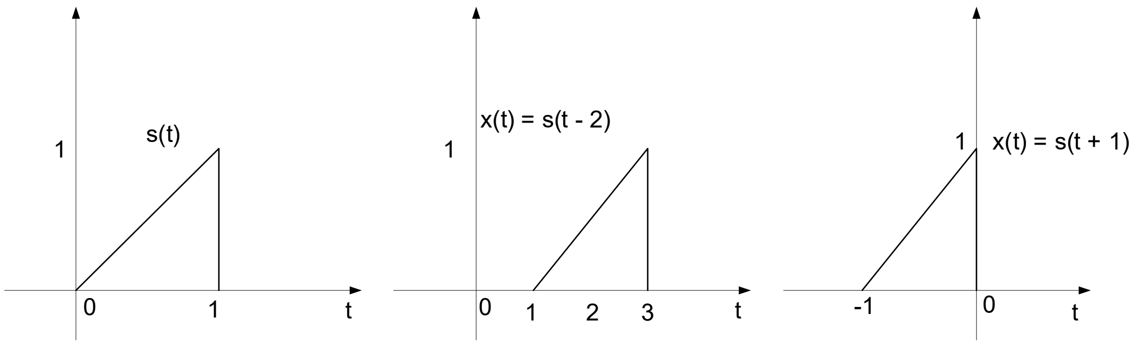
Time-Scaling of signal: Time scaling compresses or dilates a signal by multiplying the time variable by some quantity. If that quality is greater than one, the signal becomes narrower and the operation is called compression. If that quality is less than one, the signal becomes wider and the operation is called dilation. Fig. 6.10 (a) (b) and (c) shows the signal $x(t)$, compression of signal and dilation of signal respectively.



(a) Signal $x(t)$ (b) Compression of signal (c) Dilation of signal

Figure 6.10. Signal compression and Dilation

Time shifting signal: In signals and system amplitude scaling, time shifting and time scaling are some important properties. If a continuous time signal is define as $x_{(t)} = s(t - t_1)$. Then we can say that $x(t)$ is the time shifted version of $s(t)$. Consider a simple signal $s(t)$ for $0 < t < 1$.



(a) Signal s(t) (b) Signal s(t) shifted by 2 s (c) Signal s(t) shifted by -1 s

Figure 6.11 Time shifting of signals

Now shifting the function by time $t_1 = 2$ s.

$$x(t) = s(t-2) = \ t-2 \ for\ 0 \ < (t-2) < 1 = t-2 \ for\ 2 < (t-2) < 3$$

Which is simply signal s(t) with t origin delayed by 2 s. Now if we shift the signal by $t_1 = -1$ s. Then $x(t) = \ s(t+1) = \ t+1$ for $0 < (t+1) = t+1\, for - 1 < t < 0$ which is simply s(t) with its origin shifted to the left or advance in time by 1 seconds. This time-shifting property of signal is shown in the Fig. 6.11. (a), (b) and c given above.

6.6 Fourier series

In elementary analytic geometry, we were taught from vector that if you have a vector P in a two-dimensional plane x and y as given below

$$P.\hat{x} = a\,\hat{x}.\hat{x} + \ b\,\hat{y}\,.\hat{x} = a$$
$$P.\hat{y} = a\,\hat{x}.\hat{y} + \ b\,\hat{y}.\hat{y} = b$$

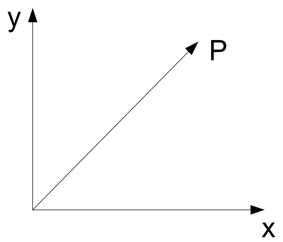


Figure 6.12.

This is possible because vector $\hat{x}$ and $\hat{y}$ are orthogonal. They are orthogonal because : $\hat{x}.\hat{y} = \hat{y}.\hat{x} = 0$ and they are normalized because $\hat{x}.\hat{x} = \hat{y}.\hat{y} = 1$.
For a given period T, the functions $\sin\left(\frac{2\pi t}{T}\right)$ and $\cos\left(\frac{2\pi t}{T}\right)$ can be considered orthogonal in a similar way.
The Fourier series is named after Jean Baptiste Joseph Fourier (1768-1830). In 1822, Fourier's genius came up with the insight that any practical periodic function can be represented as a sum of sinusoidal. Such a representation, along with the superposition theorem, allows us to find the response of circuits to arbitrary periodic inputs using phasor techniques.

6.6.1. Trigonometric Fourier series

While studying heat flow, Fourier discovered that a non sinusoidal periodic function can be expressed as an infinite sum of sinusoidal functions. Recall that a periodic function is one that repeat every T seconds. In other words, a periodic function $f(t)$ satisfies

$$f(t) = f(t + nT) \qquad 6.1$$

Where n is an integer and T is the period of the function.
According to the Fourier theorem, any practical periodic function of frequency ω_0 can be expressed as an infinite sum of sine or cosine functions that are integral multiples of ω_0. Thus, $f(t)$ can be expressed as

$$f(t) = a_0 + a_1 \cos \omega_0 t + b_1 \sin \omega_0 t + a_2 \cos 2\omega_0 t + b_2 \sin 2\omega_0 t + a_3 \cos 3\omega_0 + b_3 \sin 3\omega_0 t + \qquad 6.2$$

Which can also be written as

$$f(t) = \underbrace{a_0}_{dc} + \underbrace{\sum_{n=1}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t)}_{a.c} \qquad 6.3$$

Where $\omega_0 = \frac{2\pi}{T}$ is calleed the fundamental frequency in radians per second. The sinusoidal $\sin(n\omega_0 t)$ is called the nth harmonic of $f(t)$; it is an odd harmonic if n is odd and an even harmonic if n is even. Eq (6.3) is called the trigonometric Fourier series of $f(t)$. the constant a_n and b_n are the Fourier coefficients. The coefficient a_0 is the dc component or the average value of $f(t)$ but sinusoidal have zero average values. The co-efficient a_n and b_n (for $n \neq 0$) are the amplitude of the sinusoidal in the ac component. Thus the Fourier series of a periodic function $f(t)$ is a representation that resolves $f(t)$ into a dc component and ac component comprising an infinite series of harmonic sinusoidal.

A function that can be represented by a Fourier series as in Eq (6.3.) must meet certain requirement i.e. orthogonal, because the infinite series in Eq (6.3) may or may not converge. These conditions of $f(t)$ to yield a convergent Fourier series are as follows:

1. $f(t)$is a single value everywhere
2. $f(t)$ has a finite number of finite discontinuities in any one period.
3. $f(t)$ has a finite number of maxima and minima in any one period.
4. The integral $\int_{to}^{to+T} |f(t)|dt \; < \infty \, for \; any \; t_0$

These conditions are called Dirichlet conditions. Although they are not necessary conditions, they are sufficient condition for a Fourier series to exist. A major task in Fourier series is the determination of the Fourier coefficient a_o, a_n and b_n. The process of determining the co-efficient is called Fourier analysis. The following trigonometric integrals are very helpful in Fourier analysis. For any integers m and n

$$\int_0^T \sin n\omega_o t \, dt = 0 \qquad 6.4a$$
$$\int_0^T \cos n\omega_o t \, dt = 0 \qquad 6.4b$$
$$\int_0^T \sin n\omega_0 t \cos m\omega_0 t \, dt = 0 \qquad 6.4c$$
$$\int_0^T \sin n\omega_0 t \sin m\omega_o t \, dt = 0 \; (m \neq n) \qquad 6.4d$$
$$\int_0^T \cos n\omega_0 t \cos m\omega_o t \, dt = 0 \; (m \neq n) \qquad 6.4e$$
$$\int_0^T sin^2 n\omega_0 t dt = \frac{T}{2} \qquad 6.4f$$
$$\int_0^T cos^2 n\omega_0 t \, dt = \frac{T}{2} \qquad 6.4g$$
$$\sin n\pi = 0 \qquad 6.4h$$
$$\cos(\mathrm{n}\pi) = (-1)^n \qquad 6.4i$$

where $n \in I$

Note that in some text, instead of integrating over a period T, they integrated over 2π, which is a period for a sinusoidal functions. For so it can be interchangeably in this text. Using these identities to evaluate the Fourier coefficients. We state by finding a_0. We integrate both sides of Eq (6.3). over one period (T) and obtain

$$\int_0^T f(t)dt = \int_0^T \left[a_o + \sum_{n=1}^{\infty} (a_n \cos \, n\omega_o + b_n \sin n\omega_o t \,) \right] dt \qquad 6.5$$

$n = 1$

Involving the identities of Eqs (6.4a) and (6.4b), the two integrals involving the ac terms varnish. Hence,

$$\int_0^T f(t)dt = \int_0^T a_o dt = a_o T$$

$$\therefore\ a_o = \frac{1}{T}\int_0^T f(t)dt \ or\ a_o = \frac{1}{\pi}\int_0^{2\pi} f(t)dt \qquad 6.6$$

Note that for $T = 2\pi$, d.c component is $\frac{a_o}{2}$.
To find a_n: let us multiply both sides of Eq (6.3) by $\cos m\,\omega_o t$ and integrate over one period:

$$\int_0^T a_o \cos m\,\omega_o t dt = \int_0^T \left[a_o + \sum_{n=1}^{\infty}(a_n \cos n\omega_o t + b_n \sin n\omega_o t\right] \cos m\omega_o t\ dt$$

$$= \int_0^T a_o \cos m\omega_o t\ dt + \sum_{n=1}^{\infty}\left[\left[\int_0^T a_n \cos n\,\omega_o t \cos m\omega_o t\ dt\right.\right.$$

$$\left. + \int_0^T b_n \sin n\omega_o t \cos m\omega_o t\ dt\right] \qquad 6.7$$

The integral containing a_o is zero in view of Eq. (6.4b), while the integral containing b_n vanishes according to Eq (6.4c). the integral containing a_n will be zero except when $m = n$, in which case it is $\frac{T}{2}$, according to Eqs. (6.4e) and (6.4g). thus
$\int_0^T f(t) \cos m\,\omega_o t\ dt = a_n \frac{T}{2}\ for \quad m = n$
Such that

$$a_n = \frac{2}{T}\int_0^T f(t) \cos n\,\omega_o t\ dt\ or\ a_n = \frac{1}{\pi}\int_0^{2\pi} f(t) \cos nt\ dt \qquad 6.8$$

In a similar way, we can obtain b_n by multiplying both siddes of Eq (6.3) by $\sin m\,\omega_o t$ and integrating over the period. The result is

$$b_n = \frac{2}{T}\int_0^T f(t) \sin n\,\omega_o t\ dt\ or\ b_n = \frac{1}{\pi}\int_0^{2\pi} f(t) \sin nt\ dt \qquad 6.9$$

Be ware that since $f(t)$ is periodic, it may be more convenient to carry the integrations above from $-\frac{T}{2}$ to $\frac{T}{2}$ or generally from t_o to $(t_o + T)$ instead of 0 to T. The result will be the same.
An alternative from Eq(6.3) is the amplitude-phase form

$$f(t) = a_o + \sum_{n=1}^{\infty} A_n \cos(n\omega_o t + \emptyset_n) \qquad 6.10$$

We can apply the trigonometric identity to the a.c terms in
$\cos(A + B) = \cos A \cos B - \sin A \sin B \qquad 6.11$

Eq (6.10) so that we have

$$f(t) = a_o + \sum_{n=1}^{\infty} A_n \cos(n\omega_o t + \emptyset_n)$$

$$f(t) = a_o + \sum_{n=1}^{\infty} [A_n \cos\emptyset_n \cos n\,\omega_o t - A_n \sin\emptyset_n \sin n\,\omega_o t] \qquad 6.12$$

Equating the coefficients of the series expansion in Eqs (6.3) and (6.12) shows that

$$a_n = A_n \cos\emptyset_n,\ b_n = -A_n \sin\emptyset_n \qquad 6.13a$$

$$A_n = \sqrt{{a_n}^2 + {b_n}^2}, \emptyset_n = -\tan^{-1}\frac{b_n}{a_n} \qquad 6.13b$$

To avoid any confusion in determining $\emptyset_n$, it may be better to relate the terms in complex for as

$$A_n < \emptyset_n = a_n - jb_n \qquad 6.14$$

The convenience of the relationship will become evident in exponential Fourier series. The plot of the amplitude A_n of the harmonics versus nw_o is called the amplitude spectrum of $f(t)$; the plot of the phase $\emptyset_n$ versus $n\omega_o$ is the phase spectrum of $f(t)$. Both the amplitude and phase spectrum form the frequency spectrum for finding the spectrum of a periodic signal.

To evaluate the Fourier coefficients $a_{o,}a_n and\ b_n$, we often need to apply the following integrals:

$$\int \cos at\, dt = \frac{1}{a}\sin at \qquad 6.15a$$

$$\int \sin at\, dt = \frac{1}{a}\sin at \qquad 6.15b$$

$$\int t\cos at\, dt = \frac{1}{a^2}\cos at + \frac{1}{a}t\sin at \qquad 6.15c$$

$$\int t\sin at\, dt = \frac{1}{a^2}\sin at - \frac{1}{a}t\cos at \qquad 6.15d$$

It is needful to know the values of the cosine, sine and exponential function for integral multiples of π. These are given in table 6.1, where n is an integer.

Table 6.1. values of cosine, sine and exponential functions for integral multiples of π

Function	Value
$\cos 2n\,\pi$	1
$\sin 2n\,\pi$	0
$\cos n\,\pi$	$(-n)^n$
$\sin n\,\pi$	0

$\cos\frac{n\pi}{2}$	$\begin{cases}(-1)^{(n-2)} & , \ n = even \\ 0, & , \ n = odd\end{cases}$
$\sin\frac{n\pi}{2}$	$\begin{cases}(-1)^{\frac{(n-1)}{2}} & n = odd \\ 0 & n = even\end{cases}$
$e^{j2n\pi}$	1
$e^{jn\pi}$	$(-1)^n$
$e^{j\frac{n\pi}{2}}$	$\begin{cases}(-1)^{n/2,} & , \ n = even \\ (-1)^{\frac{(n-1)}{2}} & , \ n = odd\end{cases}$

Example 6.1:

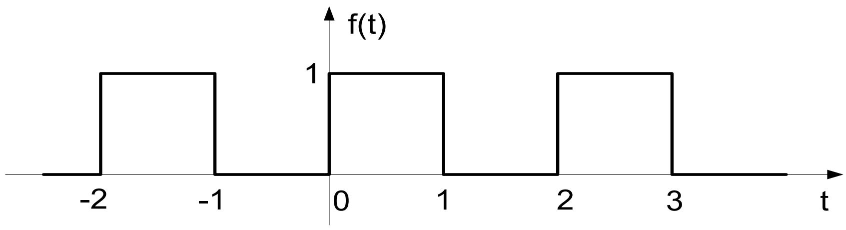


Figure. 6.13

Determine the Fourier series of the waveform shown in Fig. 6.13. obtain the amplitude and phase spectra.

Solution from

$$f(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos n\,\omega_0 t + b_n \sin \omega_0 t) \qquad 6.16$$

Our goal is to obtain the Fourier coefficient $a_{o,}a_n$, and b_n using Eqs (6.6), (6.8) and (6.9). first, we describe the waveform as

$$f(t) = \begin{cases}1 & 0 \le t \le 1 \\ 0 & 1 \le t \le 2\end{cases} \qquad 6.17$$

And $f(t) = f(t+T)$. Since $T = 2, \omega_0 = \frac{2\pi}{T}.Thus$

$$a_0 = \frac{1}{T}\int_0^T f(t)dt = \frac{1}{2}\left[\int_0^T 1\,dt + \int_1^2 0\,dt\right] = \frac{1}{2} \qquad 6.18$$

Using Eq 6.8. along with Eq 6.15a

$$a_n = \frac{2}{T}\int_0^T f(t)\cos n\omega_o t\,dt$$

$$= \frac{2}{2}\left[\int_0^1 1 \sin n\pi\, t\,dt + \int_0^2 0 \sin n\,\pi\,t\,dt\right] \qquad 6.19$$

$$= \frac{1}{n\pi} \sin n\pi \ \Big|_0^1 = \frac{1}{n\pi} \sin n\pi = 0$$

From Eq (6.9) with the aid of Eq 6.15b

$$b_n = \frac{2}{T} \int_0^T f(t) \cos n\, \omega_o t\, dt$$

$$= \frac{2}{2} \left[\int_0^1 1 \sin n\pi\, t\, dt + \int_0^2 0 \sin n\pi\, t\, dt \right] \qquad 6.20$$

$$= -\frac{1}{n\pi} \cos n\pi \ \Big|_0^1$$

$$= \frac{1}{n\pi} [1 - (-1)^n] = \begin{cases} \frac{2}{n\pi}, n = odd \\ 0, \quad n = even \end{cases}$$

Substituting the Fourier coefficient in Eqs (6.19) to (6.20) into Eq (6.16) gives the Fourier series as

$$\boldsymbol{f(t) = \frac{1}{2} + \frac{2}{\pi} \sin \pi t + \frac{2}{3\pi} \sin 3\pi t + \frac{2}{5\pi} \sin 5\pi t + \cdots} \qquad 6.21$$

Since $f(t)$ contains only the dc component and the sine terms with the fundamental component and odd harmonics, it may be written as

$$\boldsymbol{f(t) = \frac{1}{2} + \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1}{n} \sin n\pi t} \qquad \boldsymbol{n = 2k - 1} \qquad 6.22$$

By summing the terms one by one as demonstrated in Fig.6.14, we notice how superposition of the terms can evolve into the original square. As more and more Fourier components are added, the sum gets closer and closer to the square wave. However, it is not possible in practice to sum the series in Eq (6.21) or (6.22) to infinity. Only a partial sum (n=1,2,3....N where N is finite) is possible. If we plot the partial sum (or truncated series) over one period for a large N as in Fig. 6.15 we notice that the partial sum oscillates above and below the actual value of $f(t)$. At the neighborhood of the points of discontinuity $(x = 0,1,2 \ldots)$, there is overshoot and damped oscillation. In fact, an overshoot of about 9 percent of the peak value is always present, regardless of the number of terms used to approximate $f(t)$. This is called the Gibbs phenomenon.

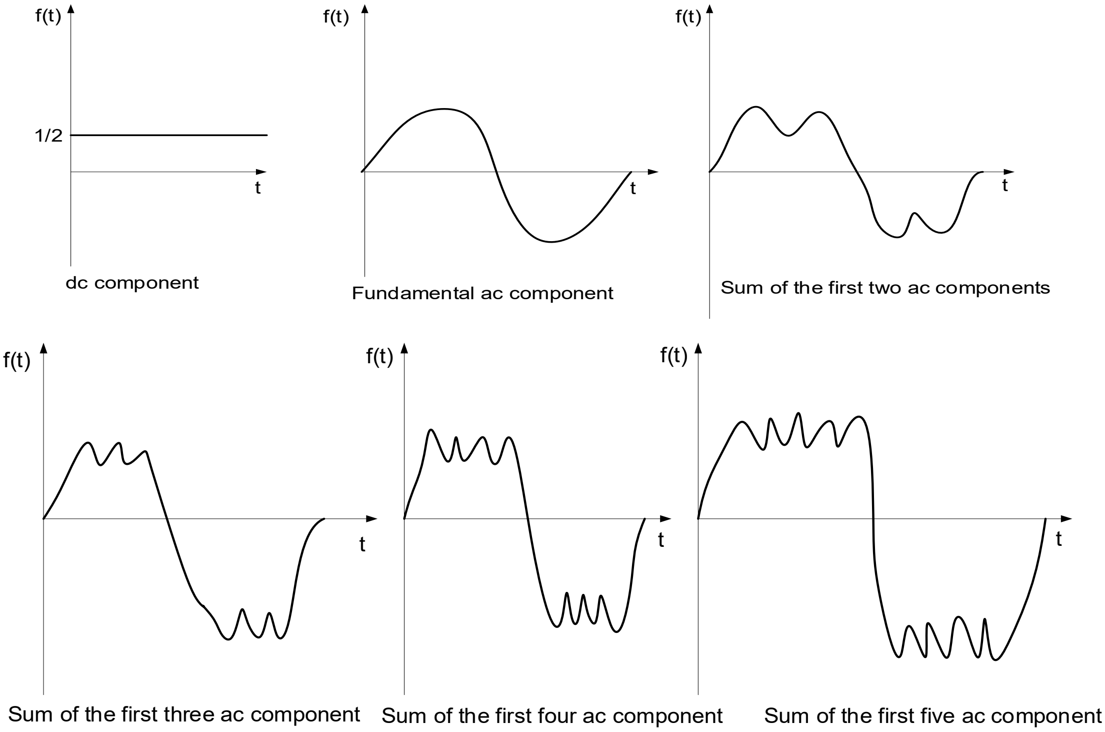


Figure 6.14. Evolution of a square waveform its Fourier components.

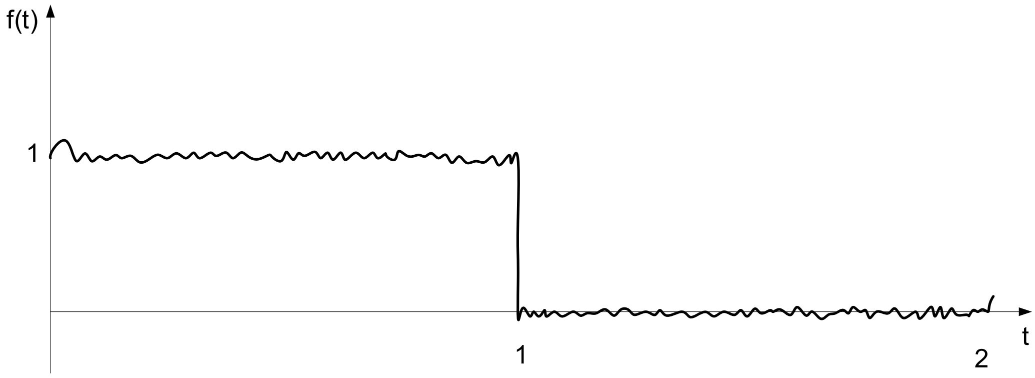


Figure 6.15. Truncating the Fourier series at N=28; Gibbs phenomenon

Let us now obtain the amplitude and phase spectral for the signal Fig. 6.13.

$$A_n = \sqrt{{a_n}^2 + {b_n}^2} \Rightarrow |b_n| = \begin{cases} \dfrac{2}{n\pi} & n = odd \\ 0 & n = even \end{cases} \quad 6.23$$

$$\text{And } \emptyset_n = -\tan {}^{b_n}/_{a_n} = \begin{cases} -90°, n = odd \\ 0, \quad n = even \end{cases} \quad 6.24$$

The plot of A_n and $\emptyset_n$ for different values of $n\omega_o = n\pi$ provide the amplitude and phase spectra in Fig. 6.16. Notice that the amplitudes of the harmonic decay very fast with frequency.

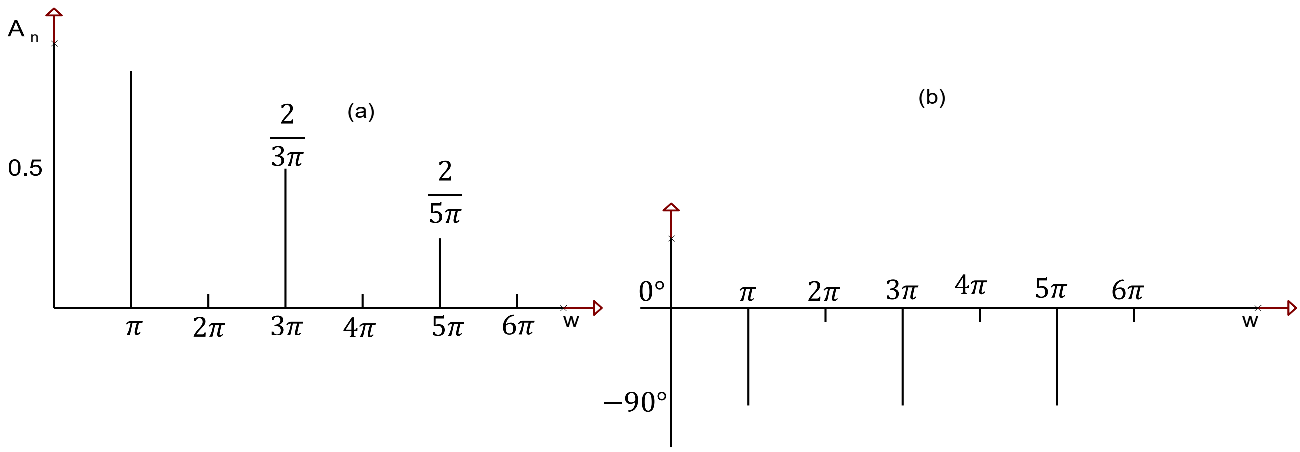


Figure 6.16. (a) Amplitude and (b) phase spectrum of the function shown in Fig. 6.13.

Example 6.2.

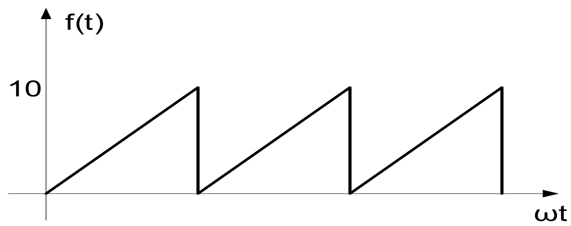


Figure 6.17

Find the Fourier series for the waveform shown in Fig. 6.17 solution: The waveform is periodic of period $\frac{2\pi}{\omega}$ in t or 2π in ωt. It is continuous for $0 \le \omega t \le 2\pi$ and given therein by $f(t) = \left(\frac{10}{2\pi}\omega t\right)$, with discontinuities at $\omega t = 2\pi n$, where n= 0, 1, 2, ... The Dirichlet conditions are satisfied. The average value of the function is 5, by inspection, and thus,

$$\frac{1}{2}a_o = 5. \quad for \quad n \ge 0,$$

$$a_n = \frac{1}{\pi}\int_0^{2\pi}\left(\frac{10}{2\pi}\right)\omega t\cos n\omega t\; d(\omega t)$$
$$= \frac{10}{2\pi^2}\left[\frac{\omega t}{n}\sin\omega t + \frac{1}{n^2}\cos n\omega t\right]_0^{2\pi}$$
$$= \frac{10}{2\pi^2 n^2}(\cos n\,2\pi - \cos 0) \;= 0$$

Thus, the series contain no cosine terms.

$$b_n = \frac{1}{\pi}\int_0^{2\pi}\left(\frac{10}{2\pi}\right)\;\omega t\cos n\omega t\; d(\omega t)\;\; d(\omega t)$$
$$= \frac{10}{2\pi^2}\left[\frac{-\omega t}{n}\cos n\omega t + \frac{1}{n^2}\sin n\omega t\;\right]_0^{2\pi} = \frac{-10}{\pi n}$$

$$\boldsymbol{f(t) = \; 5 - \frac{10}{\pi}\sin\omega t - \frac{10}{2\pi}\sin 2\omega t - \frac{10}{3\pi}\sin 3\omega t + \cdots \ldots}$$
$$\boldsymbol{= 5 - \frac{10}{\pi}\sum_{n=1}^{\infty}\frac{\sin n\omega t}{n}}$$

In the same way as the earlier example, the sine and cosine terms of like frequency can be combine as a single sine or cosine term with a phase angle. To give

$$f(t) = \frac{1}{2}a_o + \sum_{n=1}^{\infty} C_n\cos(n\omega t - \;\theta_n) \qquad 6.25$$

$$f(t) = \frac{1}{2}a_o + \sum_{n=1}^{\infty} C_n\sin(n\omega t + \;\emptyset_n) \qquad 6.26$$

Where $C_n = \sqrt{{a_n}^2 + {b_n}^2}\,, \theta_n = tan^{-1}\left(\frac{b_n}{a_n}\right) and\ \emptyset_n = \; tan^{-1}\left(\frac{a_n}{b_n}\right)$.

Note that C_n here is same as A_n earlier used. $\emptyset_n$ & θ_n are phase analysis.

6.6.2. Symmetry Considerations

We notice that the Fourier series of Example 6.1 connoted only of the sine terms. One may wonder if a method exists whereby one can know in advance that some Fourier coefficients would be zero and avoid the necessary work involved in the tedious process of calculating them. Such a method does exist; it is based on recognizing the existence of symmetry. Here we discuss three types of symmetry:

Even symmetry
Odd symmetry
Half-wave symmetry.

6.6.2.1. Even symmetry

A function $f(t)$ is even if its plot is symmetrical about the vertical axis that is;

$$f(t) = f(-t) \qquad 6.27$$

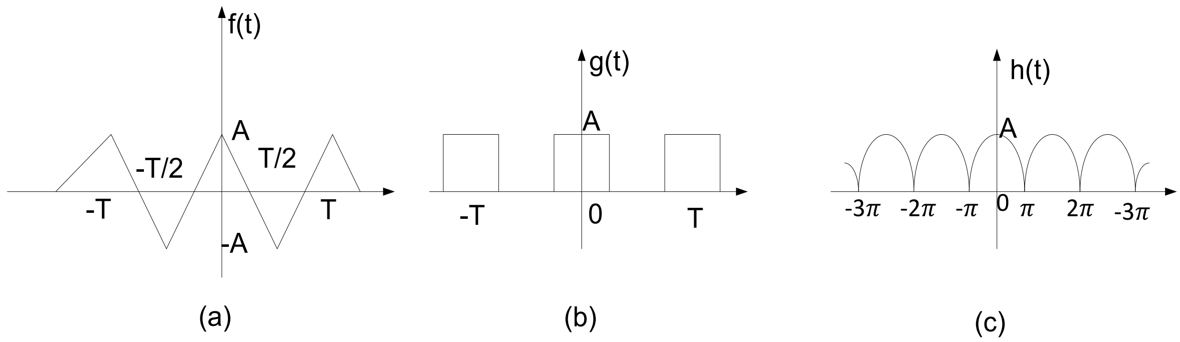


Figure 6.18. Typical examples of even periodic functions

Examples of even function are t^2, t^4, t^6, t^8, and cost. Figure 6.18 shows more examples of periodic even functions. Note that each of these examples satisfies Eq. (6.25). a main property of an even function $f_e(t)$ is that

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} f_e(t)dt = 2\int_{0}^{\frac{T}{2}} f_e(t)dt \qquad 6.28$$

Because integrating from $-\frac{T}{2}$ to 0 is the same as integrating from 0 to $\frac{T}{2}$. Utilizing this property, the Fourier co-efficient for even function becomes:

$$a_0 = \frac{2}{T}\int_{0}^{\frac{T}{2}} f(t)dt \qquad 6.29a$$

$$a_n = \frac{4}{T}\int_{0}^{\frac{T}{2}} f(t)\cos n\,\omega_o t dt \qquad 6.29b$$

$$b_n = 0$$

Since $b_n = 0$, Eq (6.3) becomes a Fourier cosine series. This makes sense because the cosine function is itself even. It also makes intuitive sense that an even function contain no sine terms since the sine function is odd.

To confirm Eq (6.27). Quantitatively, we apply the property of an even function in Eq (6.26) in evaluating the Fourier coefficients in Eqs (6.6), (6.8) and (6.9). it is convenient in each case to integrate over the interval $-\frac{T}{2} \le t \le \frac{T}{2}$, which is symmetrical about the origin. Thus

$$a_o = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} ft\, dt = \frac{1}{T}\left[\int_{-\frac{T}{2}}^{0} f(t)dt + \int_{0}^{\frac{T}{2}} f(t)dt\right] \qquad 6.30$$

We change variables for the integral over the interval $-\frac{T}{2} < t < 0$ by letting $t = -x$, so that $dt; f(t) = f(-t) = f(x)$, since $f(t)$ is an even function, and when $t = -\frac{T}{2}, x = \frac{T}{2}, then,$

$$a_0 = \frac{1}{T}\left[\int_{\frac{T}{2}}^{0} f(x)(-dx) + \int_{0}^{\frac{T}{2}} f(t)dt\right] \qquad 6.31$$

$$= \frac{1}{T}\left[\int_{0}^{\frac{T}{2}} f(x)dx + \int_{0}^{\frac{T}{2}} f(t)dt\right]$$

Showing that the two integrals are identical. Hence,

$$a_0 = \frac{2}{T}\int_{0}^{\frac{T}{2}} f(t)dt \qquad 6.32$$

As expected. Similarly, from Eq (6.8)

$$a_n = \frac{2}{T}\left[\int_{-\frac{T}{2}}^{0} f(t)\cos n\,\omega_o t\, dt + \int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t\, dt\right] \qquad 6.33$$

We make the same change of variables that led to Eq (6.31) and note that both $f(-t) = f(t)\ and \cos(-n\,\omega_0 t) = \cos n\,\omega_0 t$. Eq (6.33), then

$$a_n = \frac{2}{T}\left[\int_{-\frac{T}{2}}^{0} f(-x)\cos(-n\omega_0 x)(-dx) + \int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t\, dt\right]$$

$$= \frac{2}{T}\left[\int_{-\frac{T}{2}}^{0} f(x)\cos(n\omega_0\, x)\,(-dx) + \int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t\, dt\right]$$

$$= \frac{2}{T}\left[\int_{0}^{\frac{T}{2}} f(x)\cos(n\omega_0 x)\, dx + \int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t\, dt\right] \qquad 6.34a$$

Or

$$a_n = \frac{4}{T}\int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t dt \qquad 6.35b$$

As expected. For b_n, we apply Eq (6.9)

$$b_n = \frac{2}{T}\left[\int_{-\frac{T}{2}}^{0} f(t)\sin n\omega_0 t\, dt + \int_{0}^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt\right] \quad 6.36$$

We make the same change of variable but keep in mind that $f(-t) = f(t) but \sin(-n\omega_0 t) = -\sin n\omega_0$. Eq (6.35) yields

$$b_n = \frac{2}{T}\int_{\frac{T}{2}}^{0} f(-x)\sin(-n\omega_0 x)(-dx) + \int_{0}^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt$$

$$\frac{2}{T}\left[\int_{\frac{T}{2}}^{0} f(x)\sin n\omega_0 x\, dx + \int_{0}^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt\right]$$

$$= \frac{2}{T}\left[-\int_{0}^{\frac{T}{2}} f(x)\sin(n\omega_0 x)\, dx + \int_{0}^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt\right] = 0 \quad 6.37$$

Odd symmetry

A function $f(t)$ is said to be odd if its plot anti symmetrical about the vertical axis:

$$f(-t) = -f(t) \quad 6.38$$

∴ Examples of odd functions are $t, t^3, t^5, \sin t$.

Fig. 6.19. shows more examples of periodic odd functions. All these examples satisfy Eq (6.37). An odd function $f_o(t)$ has its major characteristic:

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} f_0(t)dt = 0 \quad 6.39$$

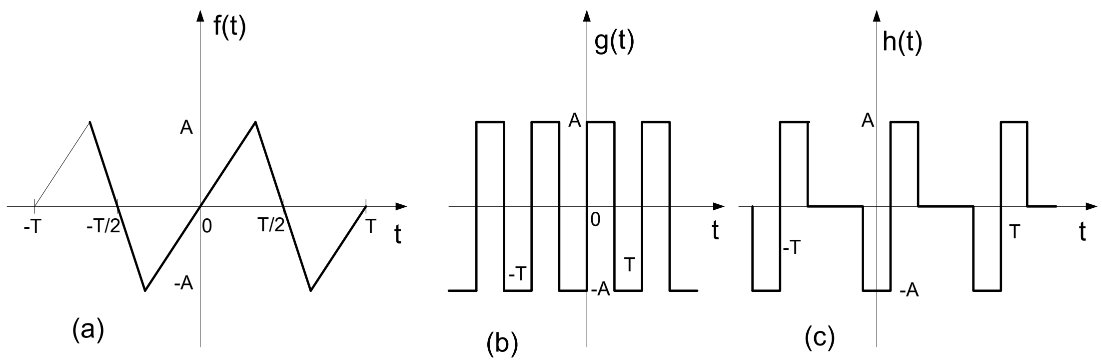


Figure 6.19. Odd periodic function

Fig. 6.19. Typical examples of odd periodic functions because integrating from $-\frac{T}{2}\ to\ 0$ is the negative of that from 0 to $\frac{\pi}{2}$. with this property, the Fourier coefficients for an odd function becomes

$$a_o = 0, a_n = 0$$

$$b_n = \frac{4}{T}\int_0^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt \qquad 6.40$$

While give us a Fourier sine series. Again this make sense because the sine function is itself an odd function. Also, note that there is no dc term for the Fourier series expansion of an odd function.

The quantitative proof of Eq (6.39) follows the same procedure taken to prove Eq (6.29) except that $f(t)$ is now odd, so that $f(t) = -f(t)$. With this fundamental but simple difference, it is easy to see that $a_0 = 0$ in Eq (6.36), $a_n = 0$ in Eq (6.34a) and b_n in Eq (6.35) becomes

$$b_n = \frac{2}{T}\left[\int_{\frac{T}{2}}^{0} f(-t)\sin(-n\omega_0 t)(-dt) + \int_0^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt\right]$$

$$= \frac{2}{T}\left[-\int_{\frac{T}{2}}^{0} f(t)\sin n\omega_0 t\, dt + \int_0^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt\right]$$

$$= \frac{2}{T}\left[\int_0^{\frac{T}{2}} f(t)\sin(n\omega_0 t)\, dt + \int_0^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt\right]$$

$$b_n = \frac{4}{T}\int_0^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt \qquad 6.41$$

As expected, it is interesting to note that any periodic function $f(t)$ with neither even nor odd symmetry may be decomposed into even and odd parts. Using the properties of even and odd function from Eq. 6.27 and 6.37, we can write

$$f(t) = \underbrace{\frac{1}{2}[f(t) + f(-t)]}_{\text{Even}} + \underbrace{\frac{1}{2}[f(t) - f(-t)]}_{\text{Odd}} = f_e(t) + f_o(t) \qquad 6.42$$

Notice that $f_e(t) = \frac{1}{2}[f(t) + f(-t)]$ satisfies the property of an even function in Eq (6.27), while $f_o = \frac{1}{2}[f(t) - f(t)]$ satisfies the properties of an odd function in Eq (6.37). the fact that $f_e(t)$ contains only the dc term and the cosine terms, while $f_o(t)$ has only the sine terms can be explained in grouping the Fourier series expansion of $f(t)$ as

$$f(t) = \underbrace{a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega_0 t}_{\text{Even}} + \underbrace{\sum_{n=1}^{\infty} b_n \sin n\omega_0 t}_{\text{Odd}} = f_e(t) + f_o(t)$$

Follow readily from Eq (6.42) that when $f(t)$ is even, $b_n = 0$ and when $f(t)$ is odd, $a_0 = 0 = a_n$.

Also note the following properties of odd and even functions

1. The product of two even functions is also an even function.
2. The product of two odd functions is an even function.
3. The product of an even function and an odd function is an odd function.
4. The sum or difference of two even function is an even function.
5. The sum or difference of two odd function is an odd function.
6. The sum or difference of an even function and an odd function is neither even nor odd.

6.6.2.2. Half wave symmetry

A function is half-wave (odd) symmetry if

$$f\left(t - \frac{T}{2}\right) = f(t) \qquad 6.43$$

Which means that each half-cycle is the mirror image of the next half-cycle. Notice that functions $\cos nw_0 and \sin nw_0 t$ satisfy Eq. (6.43) for odd values of n and therefore possess half-wave symmetry when n is odd. Fig.6.20 shows other examples of half-wave symmetric functions. The functions in Fig.6.19 a&b are half-wave symmetric. Notice that for each function, one half-cycle is the inverted version of the adjust half-cycle. The Fourier coefficients becomes

$$a_0 = 0$$

$$a_n = \begin{cases} \frac{4}{T}\int_0^{\frac{T}{2}} f(t)\cos n\omega_0 t\, dt & for\ n\ odd \\ 0, & for\ n\ even \end{cases} \qquad 6.44a$$

$$b_n = \begin{cases} \frac{4}{T}\int_0^{\frac{T}{2}} f(t)\sin n\omega_0 t\, dt & for\ n\ odd \\ 0, & for\ n\ even \end{cases} \qquad 6.44b$$

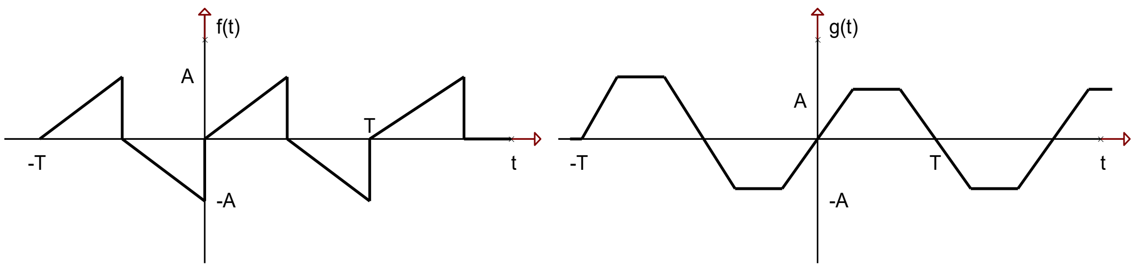


Figure 6.20: Typical Example of half-wave odd symmetric function showing that the Fourier series of a half-wave symmetric function contains only odd harmonics.

To derive Eq (6.44a) and b we apply the property of half-wave symmetric functions in Eq (6.44a) and b we apply the property of half-wave symmetric functions in Eq (6.43) in evaluating the Fourier co-efficient in Eq (6.6), (6.8) and (6.9). thus

$$a_0 = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} f(t)dt = \frac{1}{T}\left[\int_{-\frac{T}{2}}^{0} f(t)dt + \int_{0}^{\frac{T}{2}} f(t)dt\right] \qquad 6.45$$

We change variables for the integral over the interval-$\frac{T}{2} \leq\ t\ \leq 0$ by shifting $x = t + \frac{T}{2}$, so that $dx = dt; when\ t = -\frac{T}{2}, x = 0; and\ when\ t = 0, x = \frac{T}{2}$. Also, we keep Eq 6.1.26 in mid; that is, $f\left(x - \frac{T}{2}\right) = -f(x).\ then$

$$a_o = \frac{1}{T}\left[\int_{0}^{T/2} f\left(x - \frac{T}{2}\right)dx + \int_{0}^{\frac{T}{2}} f(t)dt\right]$$

$$= \frac{1}{T}\left[-\int_{0}^{\frac{T}{2}} f(x)dx + \int_{0}^{\frac{T}{2}} f(t)dt\right] = 0 \qquad 6.46$$

Confirming the expression for a_0 in Eq 6.44. Similarly,

$$a_n = \frac{2}{T}\int_{-\frac{T}{2}}^{0} f(t)\cos n\ \omega_0 t\ dt + \int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t\ dt + \int_{0}^{\frac{T}{2}} f(t)\cos n\omega_o t\ dt\Bigg] \qquad 6.47$$

We make the same change of variables that led to Eq 6.46 so that Eq 6.47 becomes

$$a_n = \frac{2}{T}\left[\int_{0}^{\frac{T}{2}} f\left(x - \frac{T}{2}\right)\right]\cos n\omega_0\left(x - \frac{T}{2}\right)dx + \int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t\ dt\Bigg] \qquad 6.48$$

Since $f\left(x - \frac{T}{2}\right) = -f(x) and \cos n\omega_0\left(x - \frac{T}{2}\right) = \cos(n\omega_0 t - n\pi)$

$$\cos nw_0\left(x - \frac{T}{2}\right) = \cos n\omega_0 t\cos n\pi + \sin n\omega_0 t\sin n\pi =\ (-1)^n\ \cos n\omega_0 t \qquad 6.49$$

Substituting these in Eq (6.48) leads to

$$a_n = \frac{2}{T}[1 - (-1)^n]\int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t\ dt$$

$$= \begin{cases} \frac{4}{T}\int_{0}^{\frac{T}{2}} f(t)\cos n\omega_0 t\ dt, & for\ \ n\ \ odd \\ 0, & for\ n\ even \end{cases} \qquad 6.50$$

Example 6.3.

Determine the Fourier series for the half-wave rectified cosine function in Fig. 6.21

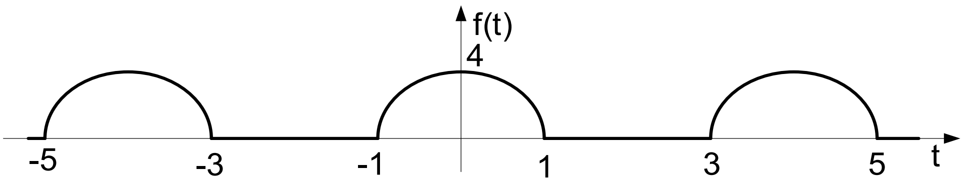


Figure 6.21 A typical Even function

Solution: This is an even function so that $b_n = 0$. Also $T = 4, \omega_0 = \frac{2\pi}{T} \Longrightarrow \ \omega_0 = \frac{\pi}{2}$. Over a period.

$$f(t) = \begin{cases} 0, & -2 \leq t \leq -1 \\ 4\cos\frac{\pi}{2}t, & -1 \leq t \leq 1 \\ 0, & 1 \leq t \leq 2 \end{cases}$$

$$a_0 = \frac{2}{T}\int_0^{\frac{T}{2}} f(t)dt = \frac{2}{4}\left[\int_0^1 4\cos\frac{\pi}{2}\ t\ dt + \ \int_1^2 0dt\right.$$
$$= \frac{1}{2}.\frac{8}{\pi}\sin\frac{\pi}{2}t\ \Big|_0^1 = \frac{4}{\pi}$$
$$a_n = \frac{4}{T}\int_0^{\frac{T}{2}} f(t)\cos nw_0 t\ dt = \frac{4}{4}\left[\int_0^1 4\cos\frac{\pi}{2}t\cos\frac{n\pi t}{2}\ dt + 0\right]$$
$$But \cos A\cos B = \frac{1}{2}[\cos(A+B) + \cos(A-B). Then$$
$$a_n = \frac{4}{2}\left[\int_0^1 \left[\cos\frac{\pi}{2}(n+1)t + \cos\frac{\pi}{2}(n-1)t\right]\right] dt$$
$$a_n = 2\int_0^1 \left[\cos\frac{\pi}{2}t + \cos\frac{\pi}{2}(n-1)t\right] dt$$

For $n = 1$

$$a_1 = 2\int_0^1 [\cos\pi t + 1]\ dt = 2\left[\frac{sin\pi t}{\pi} + \ t\right]\Big|_0^1 = 2$$
$$for\ n > 1$$
$$a_n = \frac{2}{\pi(n+1)}\sin\frac{\pi}{2}(n+1) + \frac{2}{\pi(n-1)}\sin\frac{\pi}{2}(n-1)$$
$$for\ n = odd\ (n = 1,3,5,\ldots), \qquad (n+1)\ and\ (n-1)\ are\ both\ odd.$$
$$Also, \qquad \sin\frac{\pi}{2}(n+1) = -\sin\frac{\pi}{2}(n-1) = \cos\frac{n\pi}{2} = (-1)^{\frac{n}{2}}, \qquad n = even.$$

Hence

$$a_n = \frac{2(-1)^{\frac{n}{2}}}{\pi(n+1)} + \frac{-2(-1)^{\frac{n}{2}}}{\pi(n-1)} = \frac{4(-1)^{\frac{n}{2}}}{\pi(n^2-1)} n = even$$

Thus

$$\boldsymbol{f(t) = \frac{4}{\pi} + \cos\frac{\pi}{2}t - \frac{8}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{(4k^2-1)}\cos k\pi t}$$

Which is a Fourier cosine series.

6.7 Systems

An important aspect of an engineer's work is a preoccupation with the prediction of the behaviour of some system in response to an excitation or disturbance. An excitation or disturbance is a signal and response is also a signal. A system is a physical object or assembly of physical objects for which there is a point at which excitation can be applied and another point at which the response can be observed.

For each system there are mathematical models for each elements of the system, Mathematical model for the interconnected elements which can be reduced to input-output relationship. A system may be represented by a block diagram that describes the terminal properties of the system i.e. relationship between its input-output. The relationship may be in the time domain or frequency domain.

A system can be categorized as continuous time system if its input and output signals are analog signals. A system is a discrete time system if its input and output are discrete signals.

A continuous time systems may deal with random signals, and such systems are called stochastic systems meanwhile, continuous systems may deal with deterministic signals and such a system is a deterministic system.

A system can be linear or non-linear. A linear relation exist when there are linear relationship between input and output and a non-linear system has a non-linear input-output relation

$f_1(t) \rightarrow g_3(t) \quad (Linear \ \rightarrow proportional\ relationship)$

$$f_2(t) \rightarrow g_2(t)$$
$$a\, f_1(t) \rightarrow a g_1(t)$$
$$b\, f_2(t) \rightarrow b g_2(t)$$
$$a\, f_1(t) + b f_2(t) \rightarrow a g_1(t) + b g_2(t)$$

Where a and b are constants

$$f_1(n) \rightarrow g_1(n)$$
$$f_2(n) \rightarrow g_2(n)$$
$$a f_1(n) \rightarrow g_1(n)$$
$$b f_2(n) \rightarrow b g_2(n)$$
$$a\, f_2(n) \pm b f_2(n) \rightarrow a g_1(n) + b g_2(n)$$

Systems whose characteristics vary with time are called time-varying systems while systems whose characteristics do not change with time are known as time invariant system.

Today input $f_2 \rightarrow g_1$,

Tomorrow input $f_1 \rightarrow g_1$, time invariant

Systems for which an output at any time ‘t_o’ is a function only of those input values that have occurred at $t \leq t_0$ are said to be causal when the present output does not depend on any future input. Most systems are causal. Non-causal systems are also possible.

A system is memory-less if the present output only depends on the present input and not on any past input. A system has memory if the present output depends on a past input or past inputs.

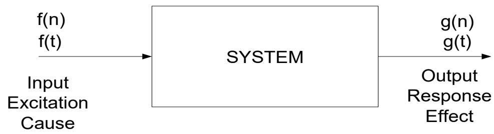


Figure 6.22 A block diagram of system

6.7.1. Time domain analysis of systems

1. Differential and Difference equations

For continuous time systems, the input output relation is a differential equation such as

$$g'(t) + \propto g(t) = f(t)$$

Substitute for $f(t)$, solve the differential equation to get $g'(t)$ for discrete time systems, the input output relation is a difference equation to get $g'(t)$.

Example 6.4

$6g(n) + 5g(n-1) + g(n-2) = f(n)$, Use iteration to obtain $g(4)$ given Fig. 6.23

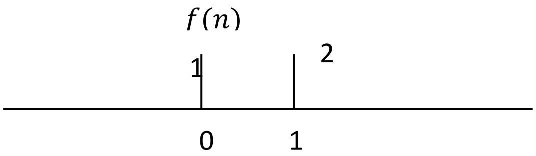


Figure 6.23

Solution

Assume $g(n) = 0$ for $n < 0$

Therefore, given: $6g(n) + 5g(n-1) + g(n-2) = f(n)$

for $n < 0, g(n) = 0$

$6g(0) + 5g(-1) + g(-2) = f(0)6g(0) = f(0)$

$6\,g(0) = f(0), \qquad g(0) = \dfrac{f(0)}{6}$

for $n = 1, 6g(1) + 5g(0) + g(-1) = f(1)$

$6g(1) + \dfrac{5}{6}f(0) = f(1)$

But $f(0) = 1, f(1) = 2$

$$6g(1) + \frac{5}{6} = 2$$

$$6g(1) = 2 - \frac{5}{6} = \frac{7}{6}$$

$$g(1) = 7/36$$

for $n = 2, 6g(2) + 5g(1) + g(0) = f(2)$

$$6g(2) = \frac{-35}{36} - \frac{1}{6} = \frac{-41}{36}$$

for $n = 3,\ 6g(3) + 5g(2) + g(1) = f(3)$

$$6g(3) = \frac{-7}{36} + \frac{205}{36} = \frac{198}{36}$$

$$g(3) = \frac{33}{36} = \frac{11}{12}$$

for $n = 4,\ 6g(4) + \ 5g(3) + \ g(2) = f(4)$

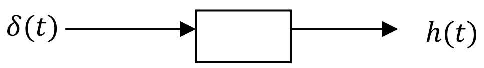
$$6g(4) = \frac{41}{36} - \frac{55}{12} = \frac{-124}{36}$$

$$g(4) = \frac{31}{54}$$

6.8. Convolution

Convolution plays a significant role in communication signal detection and in other fields such as spectral analysis of atomic emission in physics and chemistry. The importance of the convolution in system studies stems from the fact that a knowledge of the output of a linear time-invariant (LTI) system to an impulse (delta) function excitation allows us to find its output to any input function. Convolution is also used in finding the Fourier transform of signals.

The impulse response of a system is defined as the response of a system when the input to the system is an impulse function $\delta(t)$.



For a causal system

$$\delta(t - t_0) \to h(t - t_0)$$

$$\delta(t - K\Delta T) \to h(t - K\Delta t)$$

Remember the property of delta function

$$f(t_0) = \int_{-\infty}^{\infty} f(t)\delta(t - t_0)dt$$

$$f(t)\ change\ of$$

$$f(t) = \int_{-\infty}^{\infty} \delta(T)\delta(t - t)dt$$

Since for the system $\delta(t) \to h(t)$

$$h(t) = e^{-0.5t}u(t)$$

$$\delta(t - \tau) = h(t - \tau)$$

$$f(\tau)\delta(t - \tau) \to f(t)h(t - \tau)$$

$$= g(t)\int_{-\infty}^{\infty} f(t)h(t-\tau)dt$$

6.8.1 Continous Signal Convolution Integral

$$g(t) = f(t) * h(t)$$

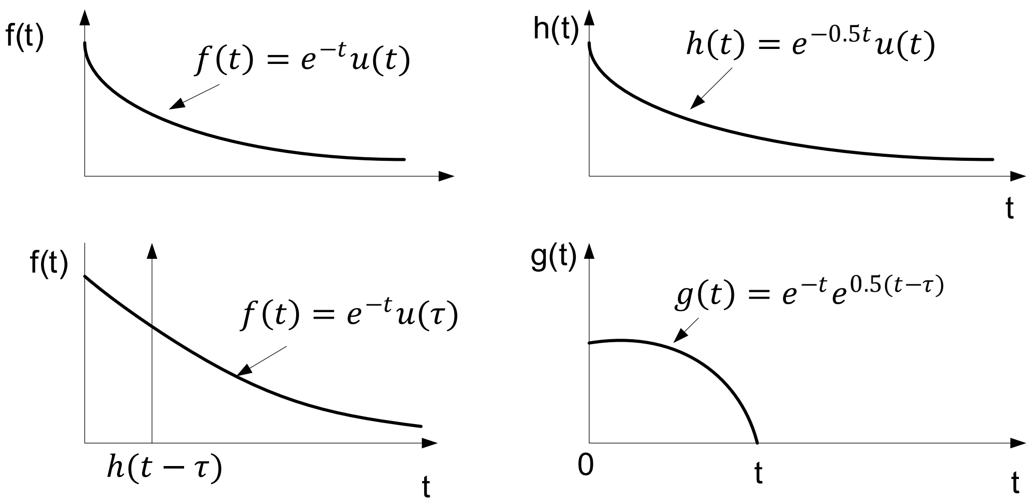
Convolution is commutative i.e.

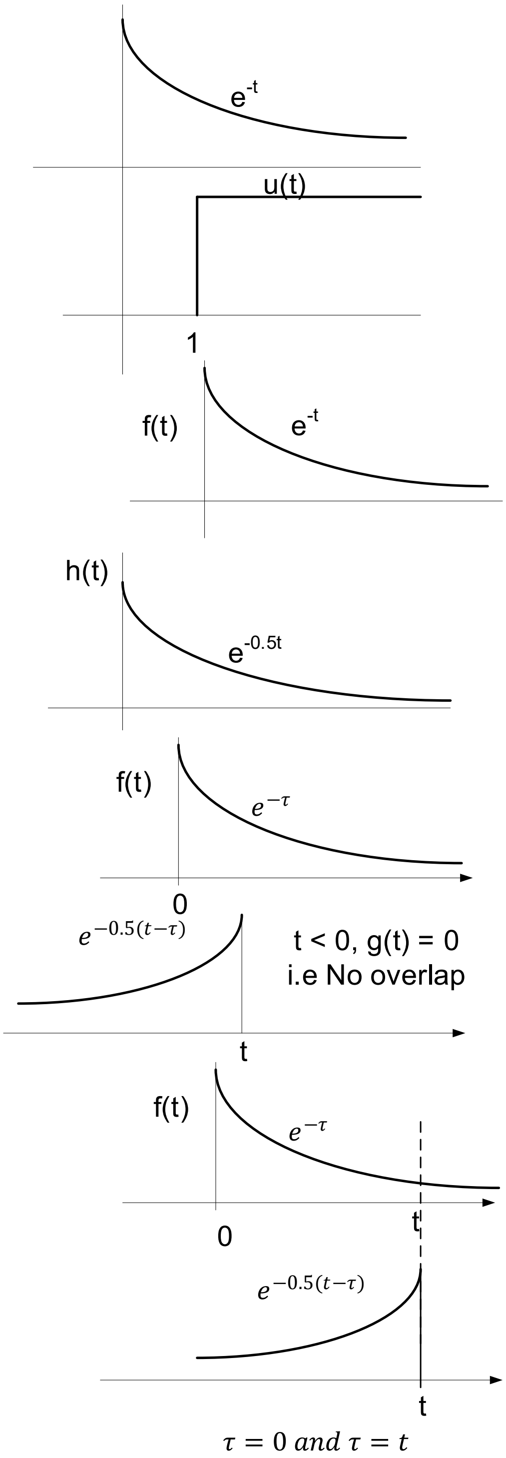
$$g(t) = \ \delta(\tau) * h(t) = h(t) * \delta(t)$$

$= \int_{-\infty}^{\infty} f(\tau)h(t-\tau)dt = \int_{-\infty}^{\infty} f(t-\tau)h(t)dt$

Example 6.5

Using convolution integral, combine the graph of $h(t) = e^{-0.5t}u(t)$ and $f(t) = e^{-t}u(t)$





$$g(\tau) = \int_0^t e^{-\tau} e^{-0.5(t-\tau)d\tau}$$

$$g(\tau) = \int_0^t e^{-\tau-0.5t+0.5\tau} d\tau$$

$$= \int_0^t e^{-0.5(t+\tau)} d\tau$$

$$= \left[\frac{e^{-0.5(\tau+t)}}{-0.5}\right]_0^t$$

$$= \frac{-e^{-t}}{0.5} - \frac{-e^{-0.5t}}{0.5}$$

$$= -2e^{-t} + 2e^{-0.5t}$$

The combined graph of convoluted process

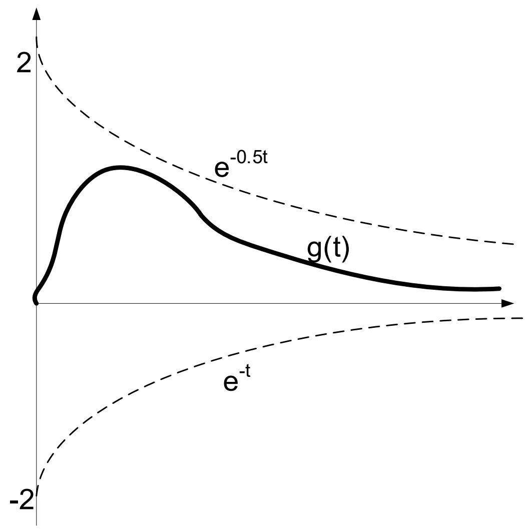


Figure 6.24

Example 6.6: Convolution $Pa(t)$ and $P_{2a}(t)$

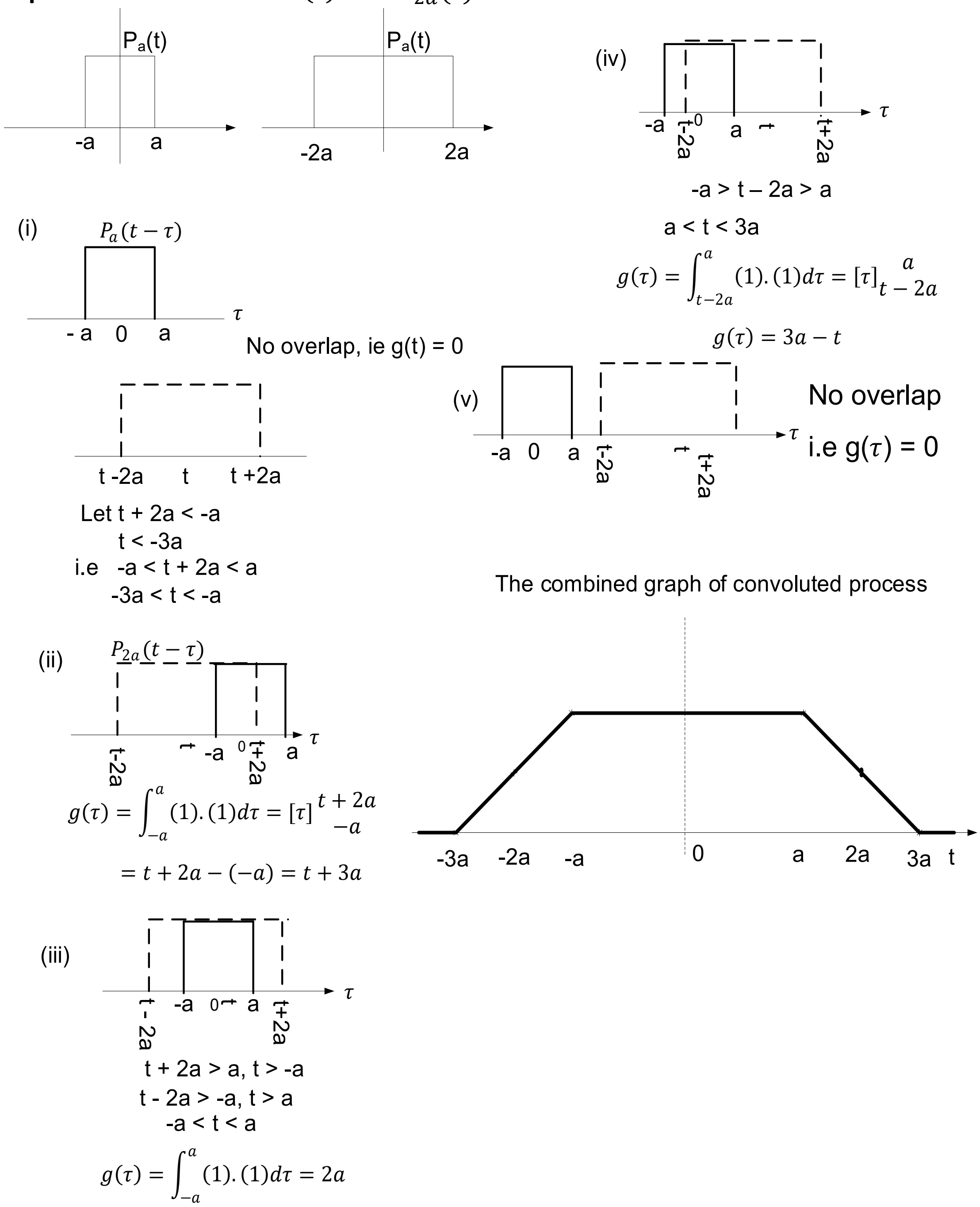


Figure 6.25

Example 6.7

Convolution the graphs of Figs. (a) and (b) below using convolution integral.

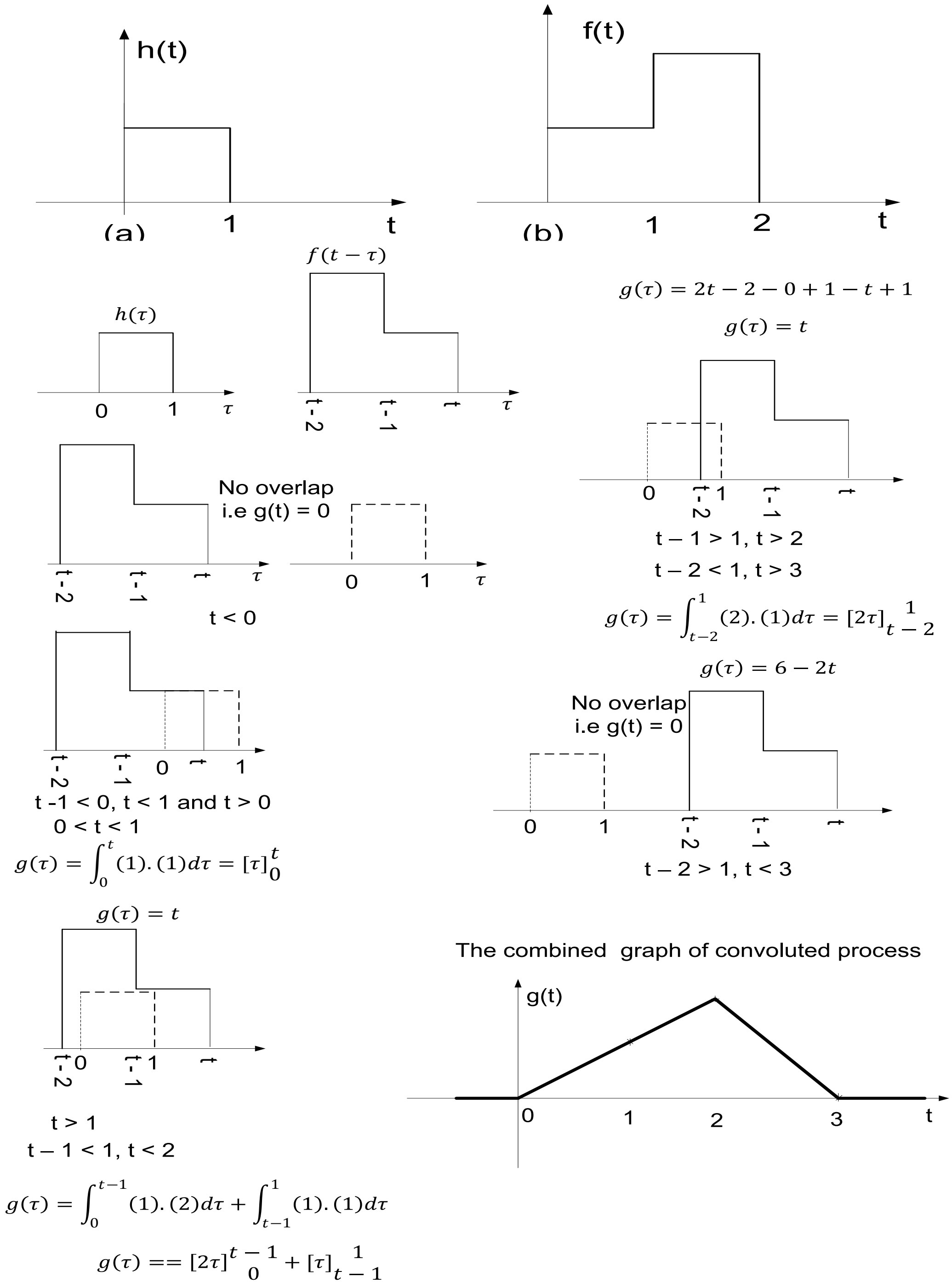


Figure 6.26

6.8.2 Discrete Signal Convolution

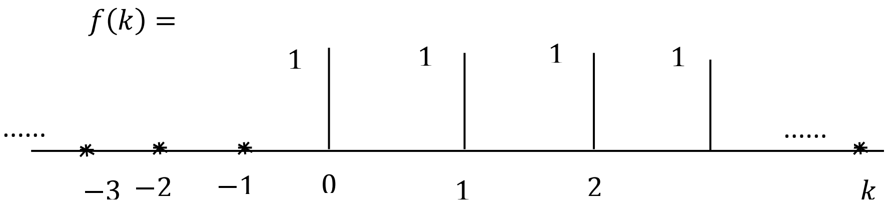
Suppose $f(k) = e^{-0.5k}u(k)$

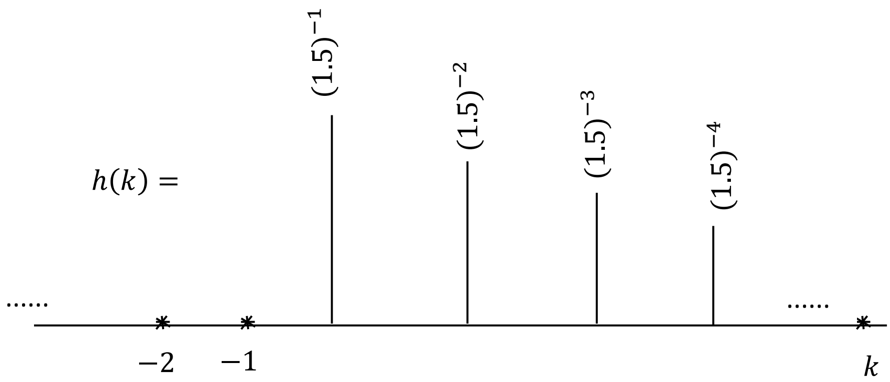
$$h(k) = \delta(k)$$
$$f(n) = e^{-0.5n}u(n)$$
$$h(k-n) = \delta(k-n)$$
$$g(k) = \sum_{n=-\infty}^{\infty} f(n)h(k-n)$$
$$= \sum_{n-\infty}^{\infty} e^{0.5n}\delta(k-n)$$
$$k = 0 \quad g(0) = 1 + 0 + 0 = 1$$
$$k = 1\,, \quad g(1) = 0 + e^{-0.5} + 0 \ldots = e^{-0.5}$$
$$k = 2\,, \quad g(2) = 0 + 0 + e^{-1} + 0 - = e^{-1}$$
$$k = k\,, \quad g(k) = e^{-0.5K}$$
$$\boldsymbol{g(k) = e^{-0.5k}u(k)}$$

Example 6.8:

Determine the convolution of the functions $f(k) = u(k)\ and\ h(k) = \frac{1}{1.5^{k+1}}$

Solution

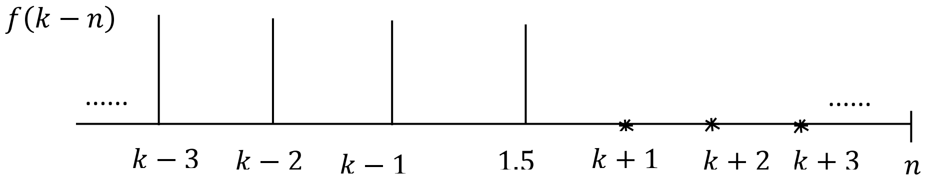


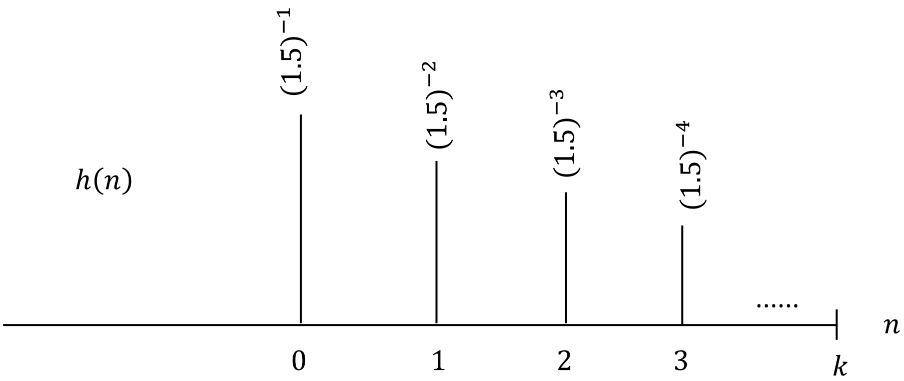


To convoluted $g(k) = f(k) * h(k)$

$$g(n) = \sum f(k-n) * h(n)$$

Graph of $f(k-n)$





$$when\ k = 0$$
$$g(k) = g(0) = (1.5)^{-1} * 1$$
$$= 1.5^{-1}$$

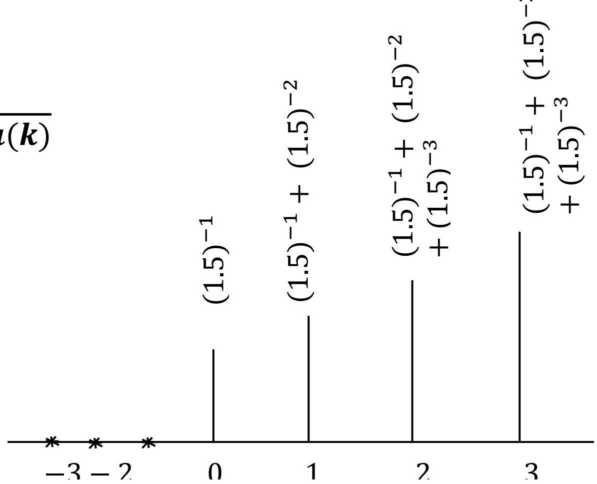
$$when\ k = 1$$
$$g(k) = g(1) = [1 * (1.5)^{-2}] + [(1.5)^{-1} * 1]$$
$$= (1.5)^{-2} + (1.5)^{-3}$$

$$when\ k = 2$$
$$g(k) = (1.5)^{-1} + (1.5)^{-2} + (1.5)^{-3}$$

$$when\ k = 3$$
$$g(k) = g(3) = (1.5)^{-1} + (1.5)^{-2} + (1.5)^{-3} + (1.5)^{-4}$$

Generally:

$$\boldsymbol{g(k)} = \sum_{i=1}^{k+1} \frac{\mathbf{1}}{(\mathbf{1.5})^{i}\, \boldsymbol{u(k)}}$$



GRAPH of $g(k)$

6.9 Chapter Review Problems

6.1 Show that if $h(t)$ (a) Using the factors $f_1(t) = p_1(t-1)\ and\ f_2(t) = e^{-t}$ verify the connotative property of the convolution. (b) Convolute (i) $tu(t)\ \&\ t^2u(t)$ (ii) $t^2u(t)\ \&\ t^3u(t)$

6.2 Show that if $h(t)$ is a real function and are defined $[f_1(t) * h(t) = s_1(t)] * h(t) = s_1(t) * g_2(t) then\ f_1(t) * h(t) = g_1(t)$ and $f_2(t) * h(t) = g_2(t) if\ g(t) = f_1(t) * f_2(t), show\ that\ \frac{dg}{dt}(t) = \frac{df_1}{dt}(t) * f_2(t) * \frac{df_2}{dt}(t)$.

6.3 (a) Explain the following terms:

(i) Signal (ii) System with memory (iii) Causal , Linear, Time-invariant System

(b) Draw the waveform of the following Signals

(i) $x(t) = 3e^{-3t}\,u(t+8)P_7\,(t+8)$

(ii) $z(t) = 4P_{10}\,(t-8)$sin(t) $\delta\,(t)$

(iii) $y(n) = -4u(n) + 4u(n-5) - 3\delta(n-7)$

(University of Ibadan, TEL525- Signal Processing 2010/2011 BSc degree Examination)

6.4 (a) Explain the following term:

(i) Signal (ii) Memory less System

(iii) Impulse Response (iv) Causal, Linear, Time-invariant System

(b) Sketch the waveform of the following Signals:

(i) $x(t) = 2e^{-2t}u(t+7)P_7(t+7)$

(ii) $x(t) = 2u(3-t)\ sinc(t)$

(iii) $y(n) = -3u(n) + 3u(n-5) - 3\delta(n-7)$

(iv) $p(n) = 2\delta(n) + 3u(n+2) - 2u(n-5) - u(n-9)$

(University of Ibadan, TEL525- Signal Processing 2013/2014 BSc degree Examination)

6.5 (a) Explain the following terms:

(i) System (ii) Causal, Linear, Time-invariant System (iii) Impulse Response

(b) Sketch the waveforms for the following Signals:

(i) $x(t) = 3u(3-t)P_{10}\,(t-3) + 4\delta(t-7)$

(ii) $x(t) = 4\delta(t-7)\ sinc(t-2) + P_{10}(t-3)$

(iii) $y(n) = 3u(n) - 3u(n-5)$

(iv) $y(n) = 2u(n) - u(n-7) + 2\delta(n+3)$

(University of Ibadan, TEL525- Signal Processing 2011/2012 BSc degree Examination)

6.6 (a) Explain the following terms:

(i) System (ii) Causal, Linear, Time-invariant System (iii) Impulse Response

(b) Sketch the waveform for the following Signals:

(i) $x(t) = 3u(3-t)\,P_{10}\,(t-3) + 4\delta(t-7)$

(ii) $x(t) = 4\delta(t-7)\;sinc(t-2) + P_{10}(t-3)$

(iii) $y(n) = 3u(n) - 3u(n-5)$

(iv) $y(n) = 2u(n) - u(n-7) + 2\delta(n+3)$

(University of Ibadan, TEL525- Signal Processing 2012/2013 BSc degree Examination)

6.7 (a) Explain briefly the following system types:

(i) Discrete-Time (ii) Analog (iii) A system with Memory

(iii) Linear

(b) Draw the waveform representing:

(i) $y(t) = 3u(2-t)\,sinc(t)$

(ii) $z(t) = 3\delta(t)P_{10}\,(t)$

(iii) $x(\pi) = 2u(\pi) - u(\pi) - u(n-4) - u(\pi-7) + 2\delta(\pi-3)$

(iv) $p(\pi) = 2u(\pi+2) - 2u(\pi-5) + \delta(\pi-7)$

6.8 Draw the waveform of the following Signals

(i) $z(t) = 3e^{-3t}\,u(t+8)\,P_7\,(t+8)$

(ii) $z(t) = 4P_{10}\,(t-8)\,sinc(t)\,\delta\,(t)$

(iii) $y(n) = -4U(n) + 4u(n-5) - 3\delta(n-7)$

(iv) $p(n) = \delta(n) + 3u(n-1) - 3u(n-4) - 2u(n-7)$

(University of Ibadan, TEL525- Signal Processing 2004/2005 BSc degree Examination

6.9 Find the trigonometric Fourier series for the triangular wave shown in fig 6Q.1 and plot the line spectrum.

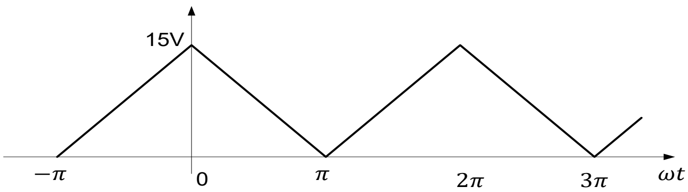


Figure 6Q.1

6.10 Find the trigonometric Fourier series for the triangular wave shown in Fig. 6Q.2 and plot the line spectrum.

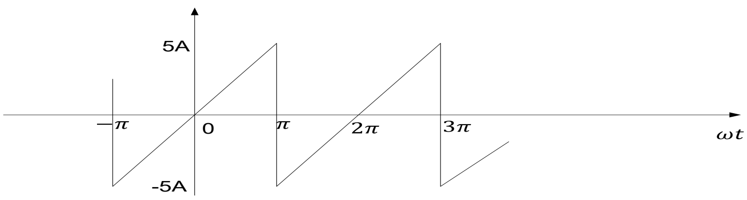


Figure 6Q.2

6.11 Find the trigonometric Fourier series for the half-wave rectified sine wave shownin Fig. 6Q.3 and sketch the line spectrum.

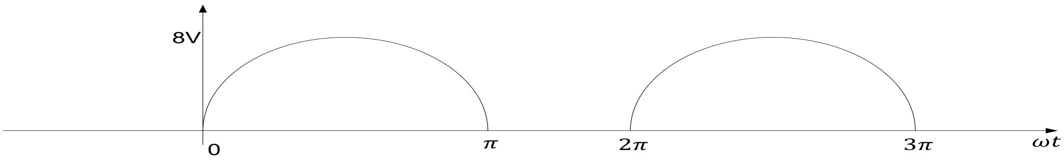


Figure 6Q.3

6.12 Find the trigonometric Fourier series for the half-wave rectified sine wave shown in Fig. 6Q.4, where the vertical axis is shifted from its position 6Q.3

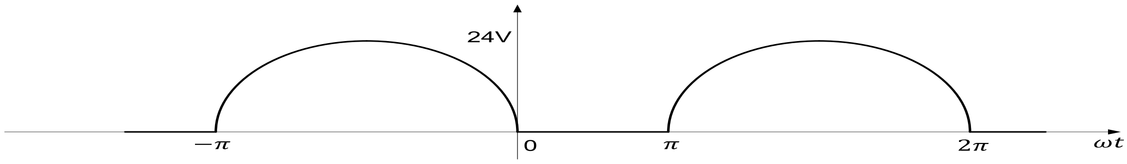


Figure 6Q.4

6.13 The voltage wave shown in figure 6Q.5 is applied to a series circuit of $R = 4\ \text{k}\Omega$ and $\text{L} = 20\ \text{H}$. Use the trigonometric Fourier series to obtain the voltage across the resistor. Plot the line spectra of the applied voltage and v_R to show the effect of the inductance on the harmonics$\omega = 377\ rad/s$

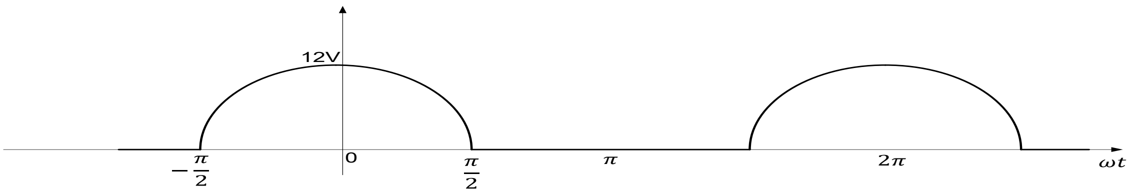


Figure 6Q.5

6.14 Find the trigonometric Fourier series for the half-wave rectified sine wave shown in Fig. 6Q.6 and sketch the line spectrum.

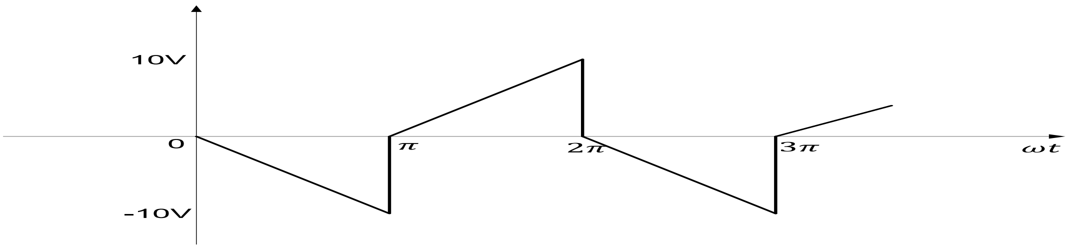


Figure 6Q.6

CHAPTER 7
SAMPLING THEORY

7.0 Introduction

As we know that broadly, there are two types of signals, continuous time signal and discrete-time signals. Due to some recent advance development in digital technology over the past few decades, the inexpensive, light weight programmable and easily reproducible discrete-time systems are available. Therefore, the processing of discrete-time signals is more flexible and is also-preferable to processing of continuous-time signals.

This means that in practice, although we have a large number of continuous time signals, but we prefer processing of discrete-time signals. For this purpose, we should be able to convert a continuous-time signal into discrete-time signal.

This problem is solved a fundamental mathematical tool known as sampling theorem. The sampling theorem is extremely important and useful in signal processing. With the help of sampling theorem, a continuous-time signal *may* be completely represented and recovered from the knowledge of samples taken uniformly. This means that sampling theorem provides a mechanism for representing a continuous-time signal by a discrete-time signal. Therefore, sampling theorem may be viewed as a bridge between continuous-time signal and discrete-time signals.

The concept of sampling provides a widely used method for using *discrete* time system technology to implement continuous-time systems and process the continuous-time signals. We utilize sampling to convert a continuous signal to a discrete-time signal, process the discrete-time signal using *a discrete* time system and then convert back to continuous-time signals.

7.1 The Sampling Theorem

Sampling of the signals is the fundamental operation in signal-processing. A continuous time signal is first converted to discrete signal by sampling process. The sufficient number of samples of the signal must be taken so that the original signal is represented in its samples completely. Also, it should be possible to recover or reconstruct the original signal completely from its samples. The number of samples to be taken *depends* on maximum signal frequency present in the signal. Sampling theorem *gives* the complete idea about the sampling of signals. Different types *of samples* are also taken like ideal samples, natural samples and flat-top *samples.*

Let us discuss the sampling theorem first and then we shall discuss *different* types of sampling processes. The statement of sampling theo*rem can be given in two parts as:*

(i) A band-limited signal of finite energy, which has no frequency-component higher than f_m Hz, is completely described by its sample values at uniform intervals less than or equal to $\frac{1}{2}\mathrm{f_m}$ second apart.

(ii) A band-limited signal of finite energy, which has no frequency components higher than f_m Hz, may be completely *recovered* from the knowledge of its samples taken at the rate of $2f_m$ samples per second.

The first part represents the representation of the signal in its samples and *minimum* sampling rate required to represent a continuous-time signal into its *samples.*

The second part of the theorem represents reconstruction of the original signal from its samples. It gives sampling rate required for satisfactory reconstruction of signal from its samples.

Combining the two parts, the sampling theorem may be stated as under:

"A continuous-time signal may be completely represented in its samples and recovered back if the sampling frequency is $f_s > 2f_m$. Here f_s is the sampling frequency and f_m is the maximum frequency present in the signal".

7.2 Proof of Sampling Theorem

To prove the sampling theorem, we shall show t*hat a signal whose spectrum is ban*d-limited to f_m Hz, can be reconstructed exactly without any error from its Samples taken uniformly at a rate $\mathrm{f_s} > 2\mathrm{f_m}$ Hz.

Let us consider a continuous time signal x(t) *whose spectrum is band-limited* to f_m Hz. This means that the signal x(t) has no frequency components *beyond* f_m Hz. Therefore, X(w) is zero for $|\omega| > \omega_m$ i.e.,

$$\mathrm{X}(\omega) = 0 \text{ for } |\omega| > \omega_m$$

Where $\omega_m = 2\pi f_m$

Fig. 7.1 (*a*) shows this continuous-time signal x(t). Let X(ω) represents Courier transform or frequency spectrum as shown in Fig. *7*.1(*b*). Sampling of x(t) at a rate of f_s in Hz (f_s samples per second) may be achieved by multiplying x(t) by an impulse train $\delta_{T_s}(t)$. The impulse train $\delta_{T_s}(t)$. consists of P impulses repeating periodically every T_s seconds, where $T_s = \frac{1}{f_s}$

Fig. *7*.1(c) shows this impulse train. This multiplication results in the *sam*pled signal g(t) shown in Fig. *7*.1(e).

This sampled signal consists of impulses spaced every T_s seconds (the sam*pling* interval).

The resulting or sampled signal may be written as

$$\mathrm{g(t)} = \mathrm{x(t)}\delta_{T_s}(t). \qquad 7.1$$

Again, since the impulse train $\delta_{\mathrm{T_s}}(\mathrm{t})$ is a periodic signal of period T_s it may be expressed as a Fourier series. The trigonometric Fourier series expansion of impulse-train $\delta T_s(t)$ is expressed as

$$\delta_{T_s}(t) = \frac{1}{T_s}[1 + 2\cos \omega_s t + 2\cos 2\omega_s t + 2\cos 3\omega_s t + \dots\dots\dots] \qquad 7.2$$

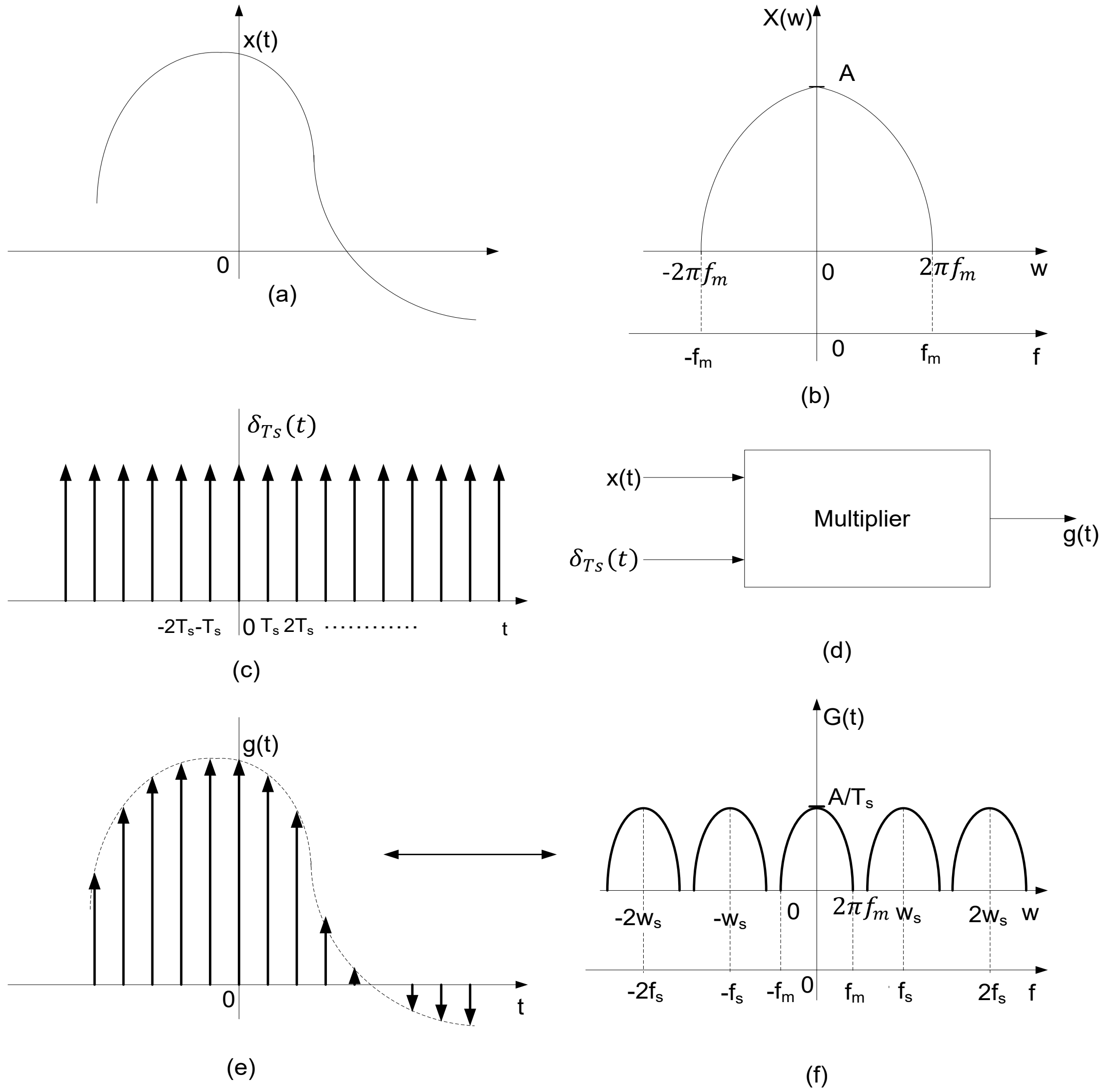


Figure 7.1 (a) A continuous-time signal (b) Spectrum of continuous-time signal (c) Impulse train as sampling function (d) Multiplier (e) Sampled signal (f) Spectrum of sampled signal.

Here $\omega_s = \frac{2\pi}{T_s} = 2\pi f_s$

Putting the values of $\delta T_s(t)$ from Eq (7.2) in Eq (7.1), the sampled signal is

$$g(t) = \frac{1}{T_s}[x(t) + 2x(t)\cos \omega_s t + 2\cos 2x(t)\cos 2\,\omega_s t + 2x(t)\cos 3\omega_s t + \dots] \qquad 7.3$$

Now, to obtain G(w), the Fourier transformation of g(t), we will have to take the Fourier transform of right hand side.

Fourier transform of x(t) is X(ω)

Fourier transform of $2x(t) \cos \omega_s t$ is $[X(\omega - \omega_s) + X(\omega + \omega_s)]$.
Fourier transform of $2x(t) \cos 2\omega_s t$ is $[X(\omega - 2\omega_s) + X(\omega - 2\omega_s)]$ and so on. Therefore, on taking Fourier transformation, the equation (7.5) becomes

$$G(\omega) = \frac{1}{T_s} s[X(\omega) + X(\omega - \omega_s) + X(\omega + \omega_s) + X(\omega - 2\omega_s) + X(\omega + 2\omega_s) + X(\omega - 3\omega_s) + X(\omega + 3\omega_s) + \cdots \ldots \ldots \ldots] \qquad 7.4$$

$$G(\omega) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} x(\omega - n\omega_s) \qquad 7.5$$

From Eqs (7.4) & (7.5), it is clear that the spectrum G(ω) consist of X(ω) repeating periodically with period $\omega_s = \frac{2\pi}{T_s}$ rad/s. or $f_s = \frac{1}{T_s}$ Hz as shown in Fig. 7.1(f)
Now if have to reconstruct x(t) from g(t), we must be able to recover X(ω) from G(ω). This is possible if there is no overlap between succession cycles of G(ω) Fig. 7.1(f) shows that this requires.

$$f_s > 2f_m \qquad 7.6$$

But the sampling interval $T_s = \frac{1}{f_s}$

Hence, $$T_s < \frac{1}{2f_m} \qquad 7.7$$

Therefore, as long as the sampling frequency f_s is greater than twice the maximum signal frequency f_m, G(ω) will consist of non-overlapping repetitions of X(ω). If this is true, Fig. 7.1(f) shows that $x(t)$; can be recovered from its samples g(t) by passing the sampled signal g(t) through
on ideal law-pass filter of bandwidth f Hz. This proves the sampling theorem.

7.2.1. Few points about sampling theorem

(i). Fig. 7.1 (f) shows the spectrum of sampled signal. According to the figure, as long as, the signal is sampled at rate $f_s > 2f_m$, the spectrum G(ω) will repeat periodically without overlapping.

(ii). The spectrum of sampled signal extends up to infinity and the ideal bandwidth of sampled signal is infinite. But here our purpose is to extract our original spectrum X(ω) out of the spectrum G(ω).

(iii). The original or desired spectrum X(ω) is centered at ω = 0 and is having bandwidth or maximum frequency equal to ω_m. The desired spectrum may be recovered by passing the sampled signal with spectrum G(ω) through a low pass filter with cut-off frequency w_m. This means that since a low-pass filter allows to pass only low frequencies up to cut-off frequency (ω_m) and rejects all other higher frequencies, the original spectrum X(ω)extended up to w_m will be selected and all other successive higher frequency cycles in the sampled-spectrum will be rejected. Therefore, in this way, original spectrum X(ω) will be extracted out of spectrum $G(\omega)$. This original spectrum $X(\omega)$ can now be converted into time-domain signal $x(t)$.

(iv). It may also be observed from Fig. 7.1 that for the case $f_s > 2f_m$, the successive cycles, of G(ω) are not overlapping each other. Hence in this case, there is no problem in recovering the original spectrum $X(\omega)$.

(v). For the case $f_s = 2f_m$, although the successive cycles of $G(\omega)$ are not overlapping each other, but they are touching each other. In this case also, the original spectrum $X(\omega)$ can be 'recovered from the sampled spectrum G(ω) using a low-pass filter with cut-off frequency ω_m.

For the case $f_s < 2f_m$, the successive cycles, of the sampled spectrum will overlap each other and hence in this case, the original spectrum $X(\omega)$ cannot be extracted out of the spectrum $G(\omega)$. Hence, For reconstruction without distortion, we must have

$f_s > 2f_m$

7.3 Nyquist Rate and Nyquist Interval

When the sampling rate becomes exactly equal to $2f_m$ samples per second, then it is called Nyquist rate. Nyquist rate is also called the minimum sampling rate. It is given by

$$f_s = 2f_m \qquad 7.8$$

Similarly, maximum sampling interval is called Nyquist interval. It is given by Nyquist Interval

$$T_s = \frac{1}{2f_m}\ \text{s} \qquad 7.9$$

When the continuous-time band limited signal is sampled at Nyquist rate ($f_s = 2f_m$), the sampled-spectrum $G(\omega)$ contains non-overlapping $G(\omega)$ repeating periodically. But the successive cycles of $G(\omega)$touch each other as shown in Fig.7.2. Therefore, the original spectrum $X(\omega)$ can be recovered from the sampled spectrum by using low pass filter with a cut-off frequency ω_m.

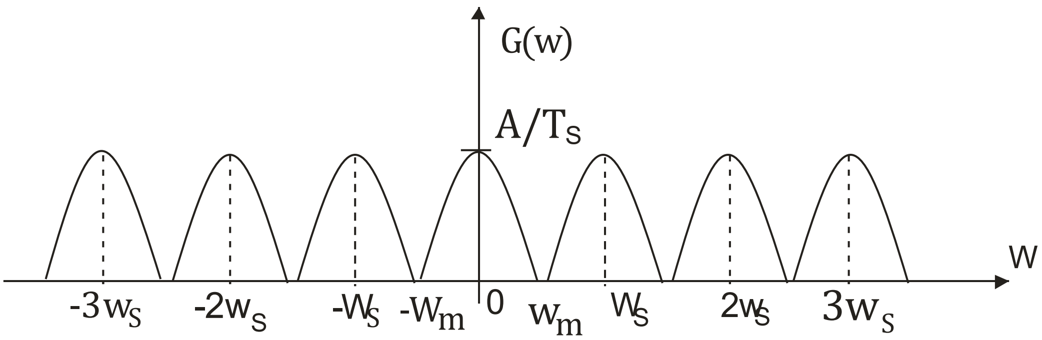


Figure 7.2. Sampled spectrum at Nyquist rate

Example 7.1 An analog signal is expressed by the equation $x(t) = 3cos\ 50\pi t + 10\ sin\ 300\pi t - \cos 100\pi t$. Calculate the Nyquist rate for this signal.

Solution: The given signal is expressed as

$$x(t) = 3cos\ 50\pi t + 10sin\ 300\pi t - cos\ 100\pi t$$

Let three frequencies present be $\omega_1, \omega_2\ \&\ \omega_3$

So that the new equation for signal,

$$x(t) = 3cos\ \omega_1 t + 10sin\ \omega_2 t - cos\ \omega_3 t$$

Comparing equations (i) and (ii) we have

$$\omega_1 t = 50\pi t;\ \omega_1 = 50\pi$$

Or $2\pi f_1 = 50\pi$

Or $2f_1 = 50$

$f_1 = 25\text{ Hz}$

Similarly, for second factor

$\omega_2 t = 300\pi t\ or\ \omega_2 = 300\pi$

Or $2\pi f_2 = 300\pi\ or\ 2\pi f_2 = 300\pi$

Therefore: $f_2 = 150\text{ Hz}$

Again, for third factor

$\omega_3 t = 100\pi t\ or\ 2\pi f_3 t = 100\pi t$

Or $2\pi f_3 = 100\pi$

Therefore: $f_3 = 5\text{ Hz}$

Therefore, the maximum frequency present in $x(t)$ is,

$f_2 = 150\text{ Hz}$

Nyquist rate is given as

$f_s = 2f_m$

Where f_m = maximum frequency present in the signal

Here $f_m = f_2 = 150\text{ Hz}$

Therefore, Nyquist rate

$f_s = 2f_2 = 2 \times 150$

Or

$\boldsymbol{f_s = 300\text{ Hz}}$ Ans

Example 7.2 Find the Nyquist rate and the Nyquist interval for the signal

$$x(t) = \frac{1}{2\pi} \cos(4000\pi t) \cos(1000\pi t)$$

Solution: Given signal is

$$x(t) = \frac{1}{2\pi} \cos(4000\pi t) \cos(1000\pi t)$$

$$x(t) = \frac{1}{4\pi} [2 \cos(4000\pi t) \cos(1000\pi t)]$$

Or $$x(t) = \frac{1}{4\pi} [\cos(4000\pi t + 1000\pi t) + \cos(4000\pi t - 1000\pi t)]$$

$$2 \cos A \cos B = \cos(A + B) + \cos(A - B)$$

Or

$$x(t) = \frac{1}{4\pi}\ [\cos\ (5000\pi t\ +\ \cos 3000\pi t)] \qquad .....(i)$$

Let the two frequencies present in the signal be ω_1 & ω_2 so that the new equation for the signal will be

$$x(t) = \frac{1}{4\pi}\ [\cos \omega_1 t + \cos \omega_2 t] \qquad ......(ii)$$

Comparing Eqs (i) & (ii), we have

$\omega_1 t = 5000\pi t$

Or $2\pi f_1 t = 5000\pi t$

Or $2f_1 = 5000$

Therefore: $f_1 = 2500\ \text{Hz}$

Similarly, for second factor

$\omega_2(t) = 3000\pi t$

or $2\pi f_2 t = 3000\pi t$

or $2\pi f_2 = 3000$

Therefore, $f_2 = 1500\ \text{Hz}$

Therefore, the maximum frequency present in $x(t)$ is

$f_1 = 2500\ \text{Hz}$

Nyquist rate is given as

$f_s = 2f_m$

Where fm = maximum frequency present in the signal

Here $f_m = f_1 = 2500\ \text{Hz}$

Therefore Nyquist rate

$$f_s = 2f_m = 2 \times 2500$$

Or $\boldsymbol{f_s = 5000\ \text{Hz} = 5\ \text{kHz}\ Ans}$

Nyquist interval is given as

$$T_s = \frac{1}{2f_m} = \frac{1}{2 \times 2500}$$

Or $T_s = \frac{1}{5000}$

Or $T_s = 0.2 \times 10^{-3}$ s

Or $\boldsymbol{T_s = 0.2\ \text{ms}}$ Ans

Example 7.3 A continuous-time signal is given below: $x(t) = 8\ cos\ 200\pi t$ Determine the following;

i. Minimum sampling rate i.e., Nyquist rate required to avoid aliasing.
ii. If sampling frequency $f_s = 400\ \text{Hz}$. What is the discrete-time signal $x(n)\ or\ x(nT_s)$ obtained after sampling?
iii. If sampling frequency $f_s = 400\ \text{Hz}$. What is the discrete-time signal $x(n)\ or\ x(nTs)$ obtained after sampling?

iv. What is the frequency $0 < f < \frac{f_s}{2}$ of sinusoidal that cycles samples identical to those obtained in part (iii)?

Solution:

The highest frequency component of continuous-time signal is $f = 100\text{ Hz}$. Hence minimum sampling rate required to avoid aliasing is called Nyquist rate and is given as

Nyquist rate $2f = 2 \times 100 = 200\text{ Hz}$

The continuous-time signal x(t) is sampled at $f_s = 400Hz$. The frequency of the discrete –time signal will be

$$\mathrm{F} = \frac{\text{Frequency of continuous - time signal f}}{\text{sampling frequency, } \mathrm{f_s}} = \frac{100}{400} = \frac{1}{4}$$

Then the discrete-time signal will be given as

$$X[n] = 8cos\ 2\pi n\mathrm{F} = 8cos\ 2\pi \times \frac{1}{4}n$$

or $X[n] = 8cos\ \frac{n\pi}{2}$

The continuous-time signal $x(t)$ is sampled at $f_s = 150\text{ Hz}$. the frequency of discrete-time will be

$$\mathrm{F} = \frac{\mathrm{f}}{\mathrm{f_s}} = \frac{100}{150} = \frac{2}{3}$$

Then the discrete-time signal will be given as

$$x(n) = 8\,cos\,2\,\pi f n = 8\cos 2\pi\,(\tfrac{2}{3})n = 8\cos \tfrac{4\pi}{3}n$$

$$= 8\cos\,[2\pi - \frac{2\pi}{3}]n = 8\cos \frac{2\pi n}{3}$$

$$x(n) = 8\,cos\ \frac{2\pi n}{3}$$ Ans

For sampling rate of $f_s = 150\text{ Hz}$

$$\mathrm{F} = \frac{\mathrm{f}}{\mathrm{f_s}} \text{ or } \mathrm{f} = \mathrm{f_s} \times \mathrm{F}$$

$$\mathrm{F} = \frac{1}{3} \times 150 = \mathbf{50\ Hz}$$

Then the sinusoidal signal will be

$$y(t) = 8\,cos\,2\pi ft = 8\,cos\,2\pi \times 50t$$
$$= 8\,cos\,100\pi t$$

Sampling at $f_s = 150\text{ Hz}$, yields identical samples hence $f = 100\text{ Hz}$ is an alias of $f = 50\text{ Hz}$ for sampling rate $f_s = 150\text{ Hz}$

Example 7.4

Determine the Nyquist rate for a continuous-time signal

$$x(t) = 6\cos 50\pi t + 20\, sin\, 300\pi t - 10\cos \pi t$$

Solution: In a general form, any continuous-time signal may be expressed as

$$x(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t + A_3 \cos \omega_3 t \qquad \text{(i)}$$

And the signal is

$$x(t) = 6\cos 50\pi t + 2\sin 300\pi t - 10\cos 100\pi t \qquad \text{(ii)}$$

On comparing given signal (ii) with standard from of a signal (i), we obtain the frequencies for the given signal is

$$f_1 = \frac{\omega_1}{2\pi} = \frac{50\pi}{2\pi} = 25 \text{ Hz}$$

$$f_2 = \frac{\omega_2}{2\pi} = \frac{300\pi}{2\pi} = 150 \text{ Hz}$$

$$f_3 = \frac{\omega_3}{2\pi} = \frac{100\pi}{2\pi} = 50 \text{ Hz}$$

Thus, the highest frequency component of the given message signal will be

$$f_{max} = 150 \text{ Hz}$$

Therefore, $Nyquist\ rate = 2f_{max}$

$$= 2 \times 150 = 300 \text{ Hz}$$

7.4 Reconstruction Filter (Low Pass Filter)

The low pass filter is used to recover original signal from its samples. This is also known as *interpolation filter*.

A low pass filter is that type of filter which passes only low-frequencies upto a specified cut-of frequency and rejects all other frequencies above cut off frequency. Fig.7.3 shows the frequency response of low pass filter

From Fig.7.3, it may be observed that in case of low-pass filter, there is sharp change in response at cut-off frequency, that is amplitude response becomes suddenly zero at cut-off frequency which is not possible practically. This means that an ideal low-pass filter is not physically realized. In place of ideal-low pass filter, we use practical filter.

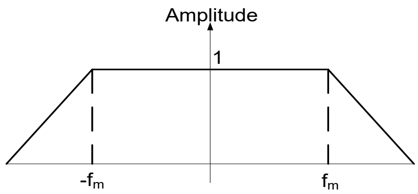


Figure 7.4: Ideal low-pass filter

Fig. 7.4 shows the frequency response of practical low-pass. From Fig. 7.4, it may be observed that in case of practical filter, the amplitude response decreases slowly to become zero. This means that there is a transition band in case of practical filter. Fig. 7.5 shows the use of practical low-pass filter in reconstruction of original signal from its sample.

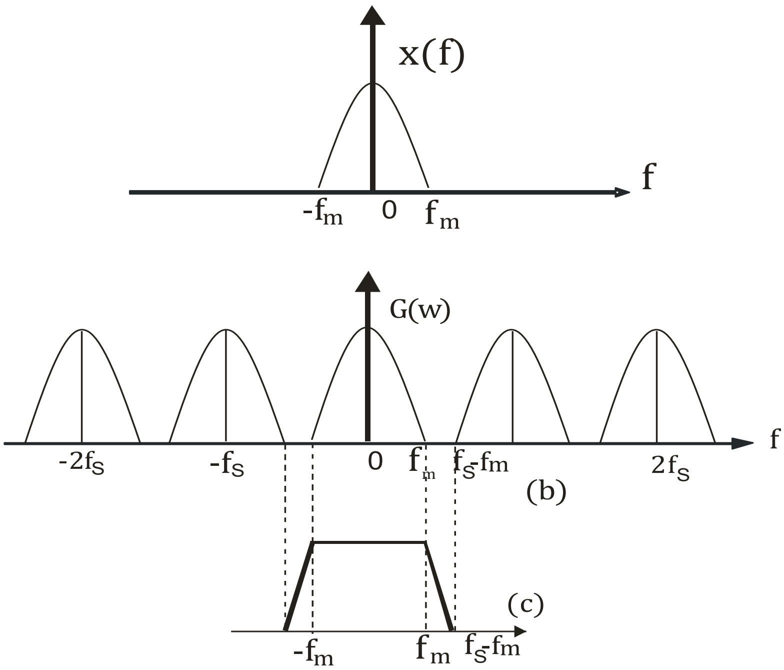


Figure 7.5: (a) Spectrum of original signal (b) Spectrum of sampled signal (c) Amplitude response of practical low-pass filter

7.5 Signal Reconstruction: The Interpolation Formula

The process of reconstruction a continuous-time signal $x(t)$ from its samples is called as interpolation.

As discussed earlier, a signal $x(t)$ band-limited to f_m Hz can be reconstructed (interpolated) completed from its samples. This is achieved by passing the sampled signal through an ideal low-pass filter of cut-off frequency f_m Hz

The expression for sampled signal is written as

$$g(t) = x(t).\delta T_1 (t) \qquad 7.10$$

$$g(t) = 1/T_s\, [x(t) + 2\, x(t)\, cos\, \omega_s t + 2\, x(t)\, cos\, 2\, \omega_s t + \ldots\ldots \qquad 7.11$$

From the above equation, it may be observed that the sampled signal contains a component 1/T_s x $x(t)$

To recover x(t) or X(w), the sampled signal must be passed through an ideal low-pass filter of bandwidth of f_mHz and gain T_s

Therefore, the reconstruction or interpolation filter transfer function may be expressed as

$$H(\omega) = T_s \times\ rect\left(\frac{\omega}{4\pi f_m}\right) \qquad 7.12$$

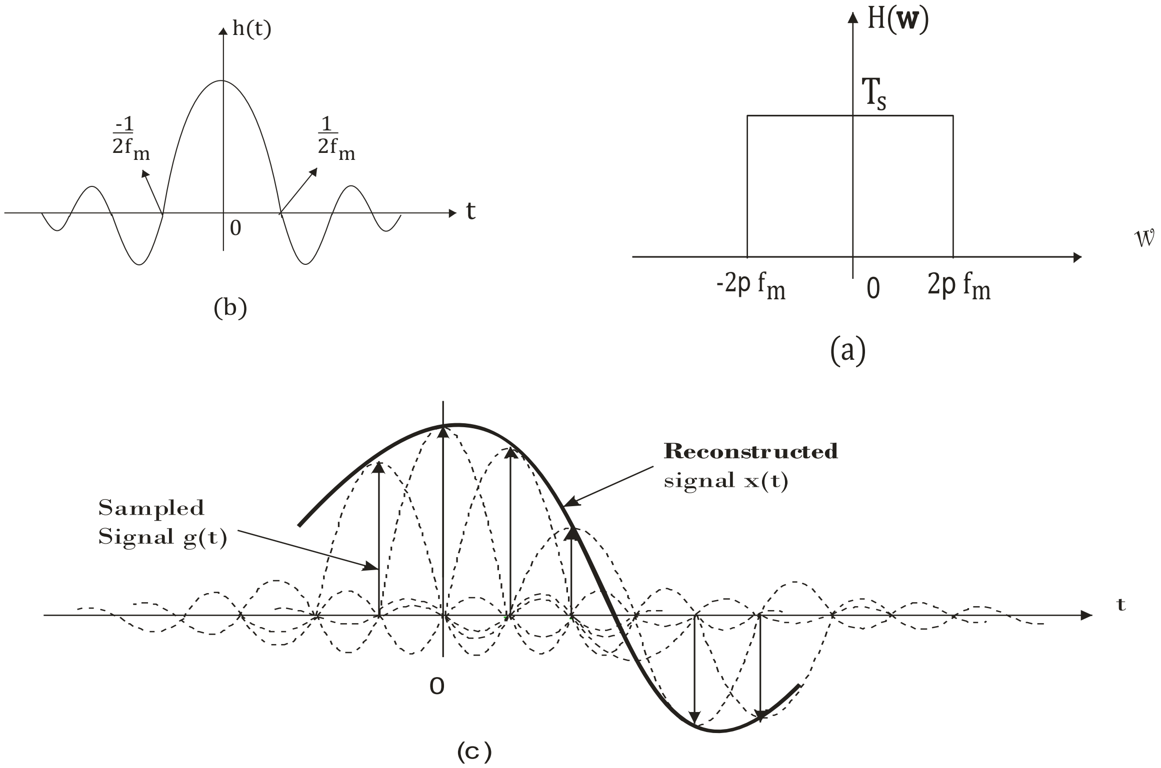


Figure 7.6. The impulse response h(t) of this filter is the inverse Fourier transform of H(ω)

$$h(t) = F^{-1}[H(\omega)]$$

$$h(t) = F^{-1}\left[T_s\, rect\left(\frac{\omega}{4\pi f_m}\right)\right] \qquad 7.13$$

$$h(t) = 2f_m, \qquad T_s = 1$$

Assuming that sampling is done at Nyquist rate than

$$T_s = \frac{1}{2\text{f}_m}$$

So that

$$2f_sT_s = 1 \qquad 7.13a$$

Putting this value of $2f_mT_s$ in Eq (7.13) we have

$$h(t) = 1\,\text{sinc}(2\pi f_m t)$$

or

$$h(t) = \text{sinc}\,(2\pi f_m t) \qquad 7.14$$

Fig. 7.6(b) shows the graph of $h(t)$.

From figure, it may be observed that $h(t) = 0$ at all Nyquist sampling instants $t = \pm\frac{n}{2}f_m$ except $t = 0$.

Now, when the sampled signal $g(t)$ is applied at the input of this filter, the output will be $x(t)$

Each sample in $g(t)$, being an impulse, produces a sinc pulse of height equal to the strength of the sample.
Addition of the sinc pulses produced by all the samples results in s(t). for instant, the k^{th} sample of the input g(t) is the impulse $x\ (kT_s)\ \delta\ (t - kT_s)$
The filter output of this impulse will be $x(kT_s)\ h(t - kT_s)$.
Therefore, the filter output to $g(t)$, which is $x(t)$, may be expressed as a sum

$$x(t) = \sum_k x\,(\mathrm{kT}_s)\ \mathrm{h}\ (\mathrm{t} - \mathrm{kT}_s) \qquad 7.15$$

$$= \sum_k x\,(\mathrm{kT}_s)\ \mathrm{sinc}\ [2\pi f_m\ (\mathrm{t} - \mathrm{kT}_s)] \qquad 7.16$$

$$x(t) = \sum_k x\,(\mathrm{kT}_s)\ \mathrm{sin\ c}\ (2\pi f_m - \mathrm{K}_m\ )\ \mathrm{T}_s = \frac{1}{2\mathrm{f}_m} \qquad 7.17$$

Eq (7.17) is know as the Interpolation formula, which provides values of x(t) between samples as a weighted sum of all the sample values.
In the proof of sampling theorem, it is assumed that the signal x(t) is strictly band-limited. But, in general, an information signal may contain a wide range of frequencies and cannot be strictly band-limited. This means that the maximum frequency f_m in the signal x(t) cannot be predictable. Therefore, it is not possible to select suitable sampling frequency f_s.

7.6 Effect of Under Sampling: Aliasing

When a continuous band-limited signal is sampled at a rate lower than Nyquist rate $f_s < 2f_m$, then the successive cycles of the spectrum $G(\omega)$ of the sampled signal g(t) overlap with each as shown in Fig.7.7.

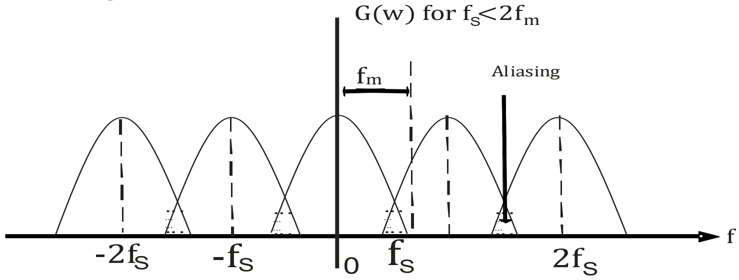


Figure 7.7: Spectrum of the sampled signal for the case $f_s < 2f_m$

Hence, the signal is under-sampled in this case *($f_s < 2f_m$)* and some amount of aliasing is produced in this under-sampling process, In fact, aliasing is the phenomenon in which a high frequency component in the frequency-spectrum of the signal takes identity of a lower-frequency component in the spectrum of the sampled signal.

From Fig. 7.7 it is clear that because of the overlap due to aliasing phenomenon, it is not possible to recover original signal $x(t)$ from sampled signal $g(t)$ by low-pass filtering since the spectral components in the overlap regions add and hence the signal is distorted.

Since any information signal contains a large number of frequencies so to decide a sampling frequency is always a problem. Therefore, a signal is first passed through a low-pass filter. This low-pass filter blocks all the frequencies which are above f_m Hz. This process is known as band limiting of the original signal $x(t)$. This low-pass filter is called prealias filter because it is used to prevent aliasing effect, After band-limiting, it becomes easy to decide sampling frequency since the maximum frequency is fixed at f_m Hz.

In short, to avoid aliasing:

Preaias filter must be used to limit band of frequencies of the signal to f_m Hz

Sampling frequency $f_s > 2f_m$ must be selected.

7.7 Sampling of BandPass Signals

In previous sections, we discussed sampling theorem for low-pass signals. However, when the given signal is a bandpass signal, then a different criteria must be used to sample the signal. Therefore, the sampling theorem for band pass signals may be expressed as under:

The bandpass signal $x(t)$ whose maximum bandwidth is $2f_m$ can be completely represented into and recovered from its samples if it is sampled at the minimum rate of twice the bandwidth. Here, f_m is the maximum frequency component present, in the signal.

Hence if the bandwidth is $2f_m$, then the; minimum sampling rate for bandpass signal must be $4f_m$ samples per second. Fig. 7.8 shows the spectrum of an arbitrary bandpass signal.

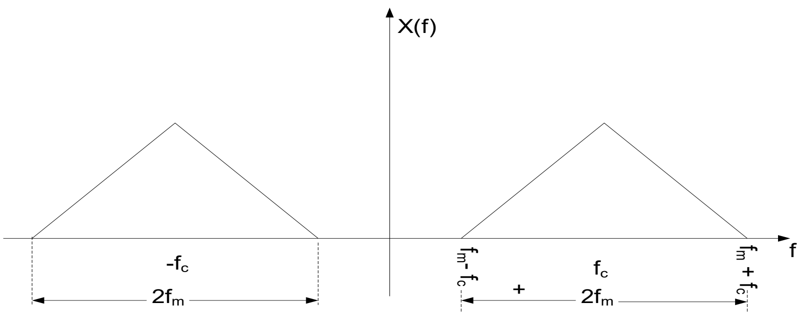


Figure 7.8: Spectrum of an orbitary bandpass signal

The spectrum in Fig. 7.8 is centred around frequency f_c. The bandwidth is $2f_m$. Thus, the frequencies present in the bandpass signal are from $f_c - f_m \; to \; f_c + f_m$. This means that the highest frequency present in the bandpass signal is $f_c + f_m$. Generally the centre frequency $f_c > f_m$.

This bandpass signal is first represented in terms of its in phase and quadrature components

Let $x_I(t)$ = Inpahse component of $x(t)$

And $x_Q(t)$ = Quadrature component of $x(t)$

Thus, the signal $x(t)$ in terms of inphase and quadrature components will be expressed as

$$x(t) = x_I(t)\cos\ (2\pi f_c t)\ \ - x_Q(t)\sin\ (2\pi f_c t) \qquad 7.18$$

The inphase and quadrature components are obtained by multiplying $x(t)$ by $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$ and then suppressing the sun frequencies by means of low-pass filters. Thus, inpahse $x_I(t)$ and quadrature $x_Q(t)$ components contain only low frequency components. The spectrum of these components is limited between $-f_m\ \ to\ +f_m$. This is shown in Fig. 7.9.

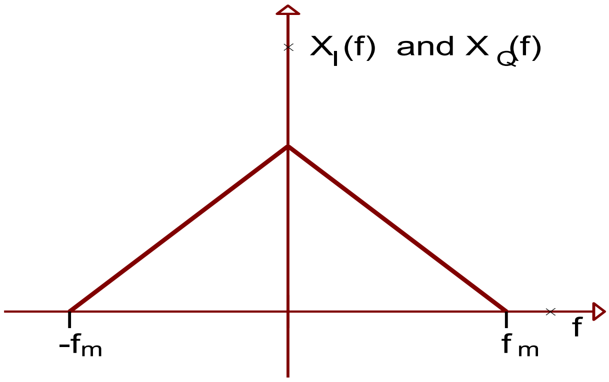


Figure 7.9: Spectrum of inphase and Quadrature components of bandpass signal $x(t)$

After few mathematical manipulation in Eq (7.18), we obtain the reconstruction formula as

$$x(t)\ =\ \sum_{n\,=\,-\infty}^{\infty} x\ (\frac{n}{4f_m})\ \text{sinc}\ (2f_m t - \frac{n}{2})\ cos\left[2nf_c\left(t - \frac{n}{4f_m}\right)\right] \qquad 7.19$$

Comparing this reconstruction formula with that of low-pass signal given in Eq (7.17), we observe that $x(t)$ is replaced by $x\left(\frac{n}{4f_m}\right)$

$$\text{Here,}\qquad x\left(\frac{n}{4f_m}\right) = \ x(nT_s) = \text{sampled version of bandpass signal}$$

and

$$\mathrm{T}_s = \frac{1}{4f_m}$$

Thus, if $4f_m$ samples per second are taken, then the bandpass signal of bandwidth $2f_m$ can be completely recovered from its samples. Hence, for bandpass signals of bandwidth $2f$, Minimum Sampling rate = Twice of bandwidth $2f_m$

Minimum sampling rate = Twice of bandwidth

= 4 f_m samples per second.

Example 7.5: Show that a bandlimited signal of finite energy which has no frequency components higher than $\mathrm{f_m}$ Hz is completely described by specifying values of the signals at instants of time separated by $\frac{1}{2f_m}$ seconds and also show that if the instantaneous values of the signal are separated by intervals larger than $1/2f_m$ seconds, they fail to describe the signal. A bandpass signal has spectral range that extends from 20 to 82 kHz.
Find the acceptable range of sampling frequency f_s.

Solution: Let $x(t)$be the band limited signal which has no frequency components higher than f_m Hz. Let this signal be sampled by a sampling function given as

$$\delta T_s(t) = \sum_{n=-\infty}^{\infty} \delta\,(t - nT_s) \qquad \text{i}$$

The sampling function is the train of impulses with T_s as distance between successive impulses. Let $x(nT_s)$be the instantaneous amplitude of signal $x(t)$ at instant $t = T_s$. The sampled version of $x(t)$ may be given as multiplication of $x(nT_s)$ *and* $\delta T_s(t)$ *i.e.,*

$$g(t) = \sum_{n=-\infty}^{\infty} \delta\,(t - nT_s)$$

Now, Fourier transform of this sampled signal may be obtained as

$$G(f) = \mathrm{FT}\{g(t)\} = \mathrm{f}_s \sum_{n=-\infty}^{\infty} X\,(f - nf_s) \qquad \text{ii}$$

Here, f_s is the sampling rate which is given as

$$f_s = \frac{1}{T_s}$$

and $X(f - nf_s) = X(f)\; at\; nf_s = 0, \pm f_s, \pm 2f_s \pm 3f_s\; \ldots \ldots$

Hence, the same spectrum $X(f)$ appears at $f = 0,$

$f = \pm f_s,\quad f = \pm 2f_s$ etc.

this means that a periodic spectrum with period equal to fs is generated in frequency domain because of sampling $x(t)$ in time-domain.

Thus, equation (i) may be written as

$$G(f) = f_s \times (f) + f_s \times (f \pm f_s) + f_s \times (f + 2f_s) + f_s \times (f + 3f_s) + f_s \times (f \pm 4f_s) + \ldots\ldots \qquad ii$$

Or

$$G(f) = f_s \times (f) + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} f_s \times (f \pm nf_s) \qquad \text{iii}$$

By definition of Fourier transform, we have

$$X\,(f) = \int_{-\infty}^{\infty} X\,(t)e^{-j2st}\;dt$$

For sampled version of $x(t)$, we have $t = nT_s$
Then the equation (iii) becomes

$$G\,(f) = \int_{n-\infty}^{\infty} X\,(n\,Ts)e^{j2\pi nT} \qquad \text{iv}$$

Now, given that the signal is band limited to f_m Hz and
, given that the signal is band limited to f_m Hz and

$$T_s = \frac{1}{2f_m}\text{ s}$$

Therefore $$f_s = \frac{1}{T_s} = 2f_m \qquad \text{v}$$

From equation (ii) may be observed that $G(f)$ is periodic with a period f_s. Thus the spectrum $X(f)$ and $G(f)$ are shown in Fig. 7.10.

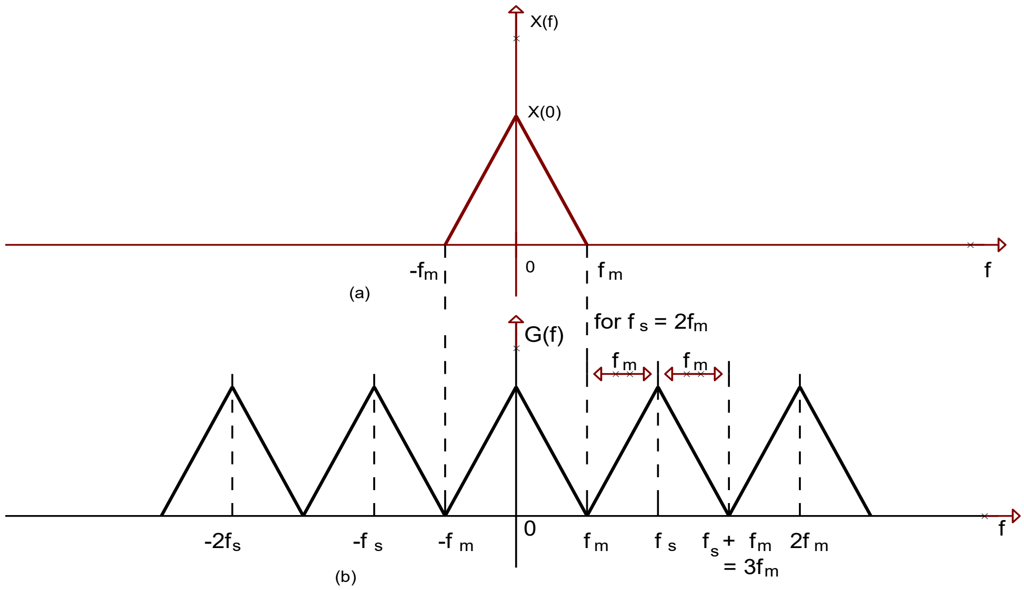


Fig 7.10: (a) Spectrum of x(t) (b) spectrum of g(t) with $f_s = 2f_m$

Now, since $f_s = 2f_m$
Therefore $f_s - f_m = f_m$
$f_s + f_m = 3f_m$

Hence, the periodic spectrum $x(f)$ just touch$\pm f_m, \pm 3f_m, \pm 5f_m$ …. etc
Thus, there is no aliasing
Using Eq (iii), we write

$$X(f) = \frac{1}{f_s}\,\mathrm{G}(f) - \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \mathrm{X}\,(\mathrm{f} - \mathrm{nf}_s) \qquad \text{vi}$$

Substituting $f_s = 2f_m$ in Eq (vi) we get

$$X(f) = \frac{1}{2\mathrm{f}_m}\,\mathrm{G}(f) - \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \mathrm{X}\,(\mathrm{f} - \mathrm{nf}_s)$$

$$X(f) = \frac{1}{2\mathrm{f}_m}\,\mathrm{G}(f) \text{ for } -\mathrm{f}_m \leq \mathrm{f} \leq \mathrm{f}_m \qquad \text{vii}$$

Now putting the value of G(f) from Eq (iv) to Eq (vii), we get

$$X(f) = \frac{1}{2f_m} \sum_{n=-\infty}^{\infty} x(nT_s)e^{-j2\pi f_n T_s}$$

Since $T_s = \frac{1}{2\mathrm{f}_m}$

$$i.e. \qquad X(t) = \frac{1}{2f_m} \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) e^{-j2\pi f_n T_m} \qquad \text{viii}$$

$x(t)$ may be recovered from $X(f)$ by taking Inverse Fourier transform of last equation i.e.,

$$i.e. X(t) = \mathbf{IFT}\left\{\frac{1}{2f_m} \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) e^{-j2\pi f_n T_m}\right.$$

This equation indicates that $x(t)$ is represented completely by its samples $x(\frac{n}{2f_m})$ for $-\infty < n < \infty$. This means that the sequence $x(\frac{n}{2f_m})$ has all the information contained in $x(t)$.

Reconstruction of signal from samples:
Let us consider equation (ix) as

$$x\,(t) = \mathbf{IFT}\left\{\frac{1}{2\,f_m} \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2\,f_m}\right) e^{-\,j2\,\pi f_n f_m}\right. \qquad \text{ix}$$

$$X\,(f) = \int_{-\,f_m}^{f_m} \{\frac{n}{2\,f_m} \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2\,f_m}\right) e^{-\,j2\,\pi f_n f_m}\; dt$$

Interchanging the order of summation and integration, we get

$$x\,(t) = \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2f_m}\right) \frac{1}{2\,f_m} \int_{-\,f_m}^{f_m} e^{-\,j2\,\pi f\left(t - \frac{n}{2\,f_m}\right)}\; dt$$

$$x\ (t) = \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2f_m}\right) \frac{\sin(2\pi\ f_m\ t\ - n\pi)}{(2\ \pi f_m\ t\ -\ n\pi)}$$

$$x\ (t) = \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2f_m}\right) \frac{\sin(2\pi\ f_m\ t\ - n\pi)}{(2\ \pi f_m\ t\ -\ n)}$$

Since,

$$x\ (t) = \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2f_m}\right) \frac{\sin\ (2\pi\ f_m\ t\ - n\pi)}{(2\ \pi f_m\ t\ -\ n\pi)}$$

Therefore,

$$x\ (t) = \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2f_m}\right)\ \sin\ (2\pi\ f_m\ t - n) - \infty < n < \infty$$

Hence, this is the interpolation formula to reconstruct $x(t)$ from its samples $x(nT_s)$. Therefore, from all above, it is clear that the signal may be completely represented into and recovered its samples between the successive samples is

$\frac{1}{2f_m}$ seconds i.e., $f_s = 2f_m$ samples per seconds

Sampling frequency for Bandpass signal

Since the spectral range of the bandpass signal is 20 kHz to 82 kHz

Therefore,

$Bandwidth = 2f_m,\ and\ the\ range\ is\ \ 82\ kHz\ - 20\ kHz\ =\ 62\ kHz$

Hence, minimum $sampling\ rate\ =\ 2\ x\ bandwidth$

$= 2 \times 62$

$= \mathbf{124\ kHz}$

Generally, the range of minimum sampling frequencies is specified for bandpass signals. It lies between $4f_m$ to $8f_m$ samples per second

Therefore,

Range of minimum sampling frequencies

$= (2 \times bandwidth)\ to\ (4 \times bandwidth)$

$= 2 \times 62\ \text{kHz}\ to\ 4 \times 62\ \text{kHz}$

` $= \mathbf{124\ kHz}\ \boldsymbol{to}\ \mathbf{248\ kHz}$

7.8 Sampling Techniques

In the last article, we discussed how sampling of a continuous-time signal is done. This sampling of a signal is done in several ways. Therefore ion this section we shall discuss different types of sampling i.e., sampling techniques.

Basically, there are three types of sampling techniques are under:

I. Instantaneous sampling
II. Natural sampling
III. Flat top sampling

Out of these three, instantaneous sampling is called ideal sampling whereas natural sampling and flat-top sampling are called practical sampling methods. Now, let us discuss three different types of sampling techniques in detail.

7.8.1. Ideal Sampling or Instantaneous Sampling or Impulse Sampling

In the proof of sampling theorem, we used ideal or impulse sampling. In this type of sampling, the sampling function is a train of impulses, Fig.7.11b shows this sampling function.

$x(t)$ is the input signal *(i.e.,* signal to be sampled) as shown in Fig.7.11a.

Fig. 7.11c shows a circuit to produce instantaneous or ideal sampling. This circuit is known as the switching sampler.

The working principle of this circuit is quite easy. The circuit simply consists of a switch. Now if we assume that the closing time *'t'* of the switch approaches zero, then the output $g(t)$ of this circuit will contain only instantaneous value of the input signal $x(t)$. Since the width of the pulse approaches zero, the instantaneous sampling gives a train of impulses of height equal to the instantaneous value of the input signal $x(t)$ at the sampling instant.

We know that the train of impulses may be represented as

$$\delta T_s(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \qquad 7.20$$

This is know as sampling function and its waveform is shown in figure (7.11b)

The sampled signal g(t) is expressed as the multiplication of x(t) ans $\delta T_s(t)$

Thus:

$$g(t) = x(t)T_s(t) \qquad 7.21$$

$$x(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \qquad 7.22$$

$$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s)\,\delta(t - nT_s) \qquad 7.23$$

The Fourier transform of the ideally sampled signal given by above equation may be expressed as

$$G(t) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \qquad 7.24$$

Note: This equation gives the spectrum of ideally sampled signal. It shows that the spectrum $X(f)$ is periodic in f_s and weighted by f_s. However, it may be noted that ideal or instantaneous sampling is possible only in theory since it is impossible to have a pulse whose width approaches zero. Ideal sampling was used in last article to prove sampling theorem. Practically flat-top sampling and nature sampling are used.

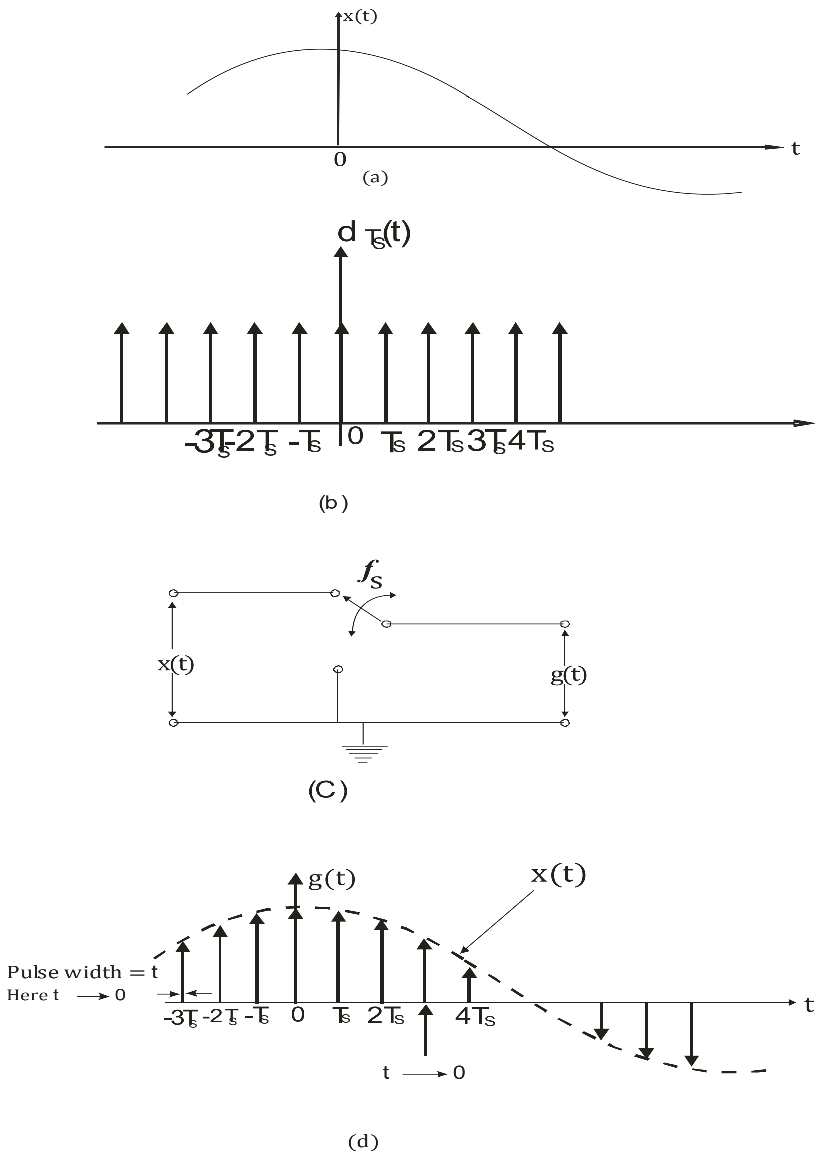


Figure 7.11: (a) baseband signal (b) impulse train (c) functional diagram of a switching sampler (d) sampled signal

7.8.2. Natural sampling

As discussed in last article, the instantaneous sampling results in the sampling whose width τ approaches zero. Due to this, the power content in the instantaneously sampled pulse is negligible. Thus, this method is not suitable for transmission purpose. Natural sampling is a practical method and will be discussed in this section.

In natural sampling the pulse has a finite width equal to τ.
Let us consider an analog continuous-time signal $x(t)$ to be sampled at the rate of f_s Hz. Here it is assumed that f_s is higher that Nyquist rate such that sampling theorem is satisfied.
Again, let us consider a sampling function $c(t)$ which is a train of periodic pulses of width τ and frequency equal to f_s Hz.

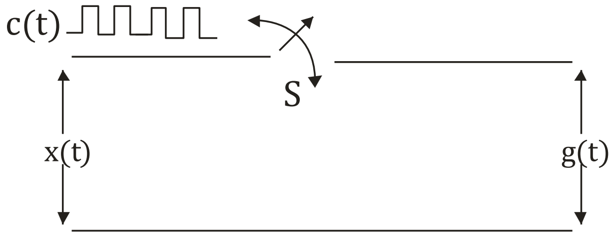


Figure 7.12. A functional diagram of Natural Sampler

Fig. 7.12 shows a functional diagram of a natural sampler. With the help of this natural sampler, a sampled signal $g(t)$ is obtained by multiplication of sampling function $c(t)$ and input signal $x(t)$.
Now according to Fig. 7.12, we have
When $c(t)$ goes high the switch 'S' is closed.
Therefore

$$g(t) = x(t) \text{ when } c(t) = A \qquad 7.25$$

And $$g(t) = 0 \, when \; c(t) = O \qquad 7.26$$

Where A is the amplitude of c(t)
The waveform of signal $x(t), c(t)$ and $g(t)$ have been illustrated in Figs. 7.13(a), 7.13(b) and 7.13(c) respectively.
Now, the sampled signal g(t) may also be described mathematically as

$$g(t) = c(t) * x(t) \qquad 7.27$$

Here, $c(t)$ is the periodic train of pulse of width τ and frequency f_s.
We know that the Exponential Fourier series for any periodic waveform is expressed as

$$s(t) = \sum_{n=-\infty}^{\infty} c_n \, e^{j2\pi f n T_s} \qquad 7.28$$

Also, for periodic pulse train of c(t), we have

$$T_0 = T_s = \frac{1}{f_s} = periodic \; of \; c(t)$$

$$f_0 = f_s = \frac{1}{T_0} = \frac{1}{T_s} = frequency \; of \; c(t)$$

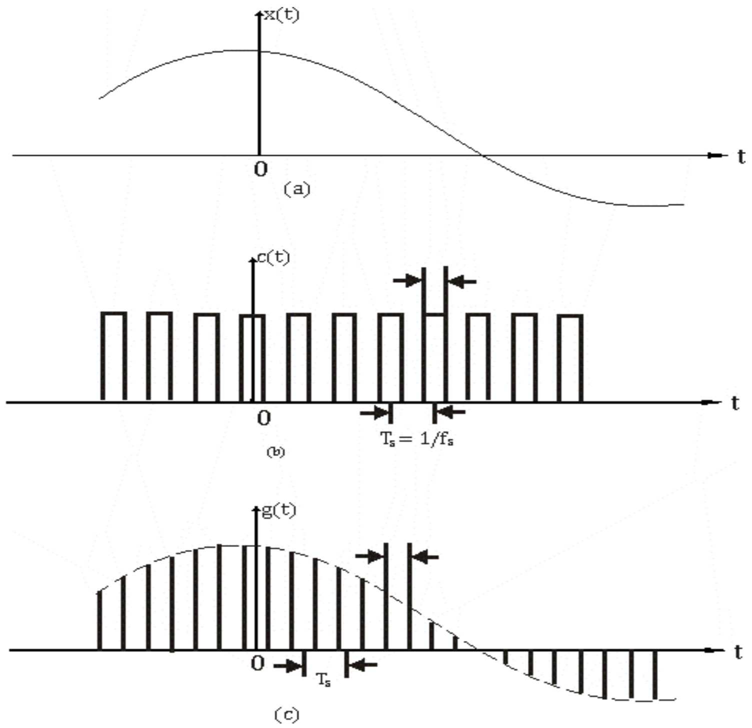


Figure 7.13: (a) Continuous time signal $x(t)$. (b) Sampling function wave form i.e. periodic pulse train (c) Naturally sampled signal waveform $s(t)$

Therefore according to Eq (7.28) for periodic pulse train $c(t)$, we have

$$c(t) = \sum_{n=-\infty}^{\infty} C_n e^{j2\pi f n T_s} \ with \ \frac{1}{T_0} = f_s \qquad 7.29$$

Now, it may be noted that since c(t) is a rectangular pulse train, therefore C_n for this waveform will be expressed as

$$C_n = \frac{T_A}{T_0} sinc(f_n . T) \qquad 7.30$$

Here $T = pulse\ width = \tau$
And $f_n = harmonic\ frequency$
But here, $f_n = nf_s$
Or $f_s = nf_0$
Hence,

$$C_n = \tau . \frac{A}{T_s} sinc(f_n . \tau) \qquad 7.31$$

Therefore, using Eqs. (7.29) and (7.30) the Fourier series representation for $c(t)$ will be given as

$$c(t) = \sum_{n=-\infty}^{\infty} \frac{\tau\,.A}{T_s} \sin c\ (f_n\,.\tau)\, e^{j2\pi f_s.nt} \quad with\ \frac{1}{T_0} = f_s \qquad 7.32$$

Now, substituting the value of c(t) from Eq (7.32) to Eq (7.27) we get

$$g(t) = \frac{\tau A}{T_s} \sum_{n=-\infty}^{\infty} \sin c\ (f_n\,.\tau)\, e^{j2\pi f_s.nt}\,.x(t) \qquad 7.33$$

This is required time-domain representation for naturally sampled signal g(t)

Now, to get the frequency-domain representation of the naturally sampled signal g(t), let us take its Fourier transform as

$G(f) = FT[g(t)]$

$$G(t) = \frac{\tau A}{T_s} \sum_{n=-\infty}^{\infty} \mathrm{sinc}(f_n\,.\tau)\, FT[e^{j2\pi f_s.nt}\,.x(t)] \qquad 7.34$$

Recall the frequency-shifting property of Fourier transform which states that

$$e^{j2\pi f_s.n(t)}.x(t) \leftrightarrow X(f - fs.n) \qquad 7.35$$

Therefore,

$$G(f) = \frac{\tau A}{T_s} \sum_{n=-\infty}^{\infty} sinc(f_n\,.\tau) X(f - f_s.n) \qquad 7.36$$

Now, since $f_n = f_s n\ =\ harmonic\ frequency$

Therefore, equation (7.36) becomes

$$G(t) = \frac{\tau A}{T_s} \sum_{n=-\infty}^{\infty} \sin c\ (f_n\,.\tau)\, X(f - nf_s) \qquad 7.37$$

Hence, we write

Spectrum of Naturally Sampled signal:

$$G(f) = \frac{\tau A}{T_s} \sum_{n=-\infty}^{\infty} \sin c\ (f_n\,.\tau)\, X(f - nf_s) \qquad 7.38$$

This equation shows that the spectra of $x(f)\ i.e., X(f)$ are periodic in f_s and are weighed by the sin$c$ function.

Fig. 7.14, illustrates some arbitrary spectra for $x(t)$ and corresponding spectrum $G(f)$.

Note : Thus form Fig. 7.14 it may be noted that unlike the spectrum of instantaneously sampled signal shown in Fig. 7.1f, the spectrum of a naturally sampled signal is weighted by a sin c function whereas the spectrum of an instantaneously sampled signal (Fig. 7.1f) remains constant throughout the frequency range.

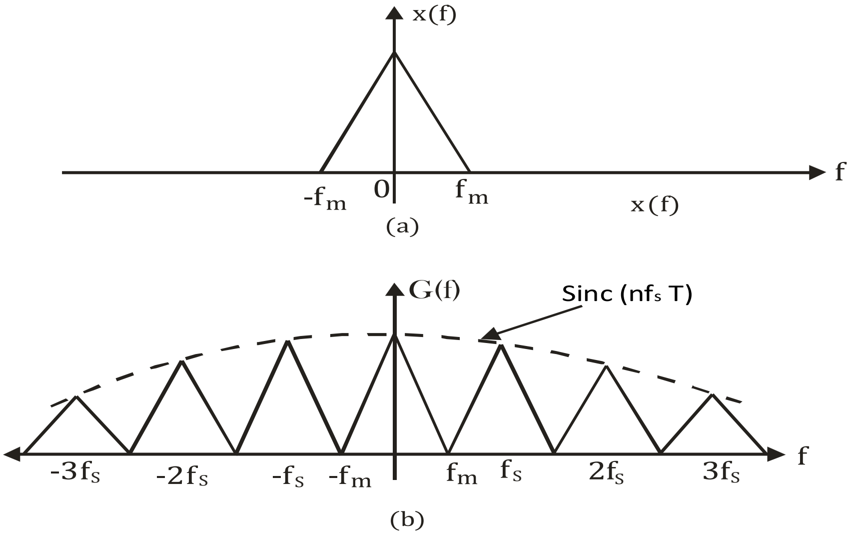


Figure 7.14: (a) Spectrum of continuous-time signal x(t) (b) Spectrum of naturally sampled signal

7.8.3. Flot Top Sampling or Rectangular Pulse Sampling

Flot top sampling like natural sampling is also a practically possible sampling method. But natural sampling is little complex whereas it is quite easy to get flat top samples. In flat-top sampling or rectangular pulse sampling, the top of the samples remains constant and is equal to the instantaneous value of the baseband signal $x(t)$at the start of sampling. The duration or width of each sample is τ and sampling rate is equal to $f_s = \frac{t}{T_s}$. Fig.7.15(a) shows the functional diagram of a sample and hold circuit which is used to generate the flot top samples Fig.7.15(b) shows the general waveform of flat top samples.

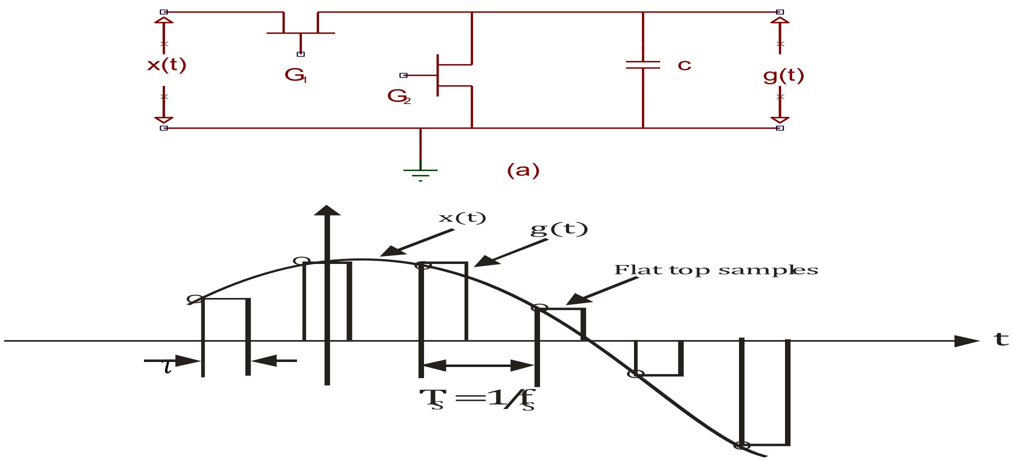


Figure 7.15: (a) A sampled and hold circuit to generate flat top samples. (b) A general waveform of flat top sampling.

From Fig. 7.15(b), it may be noted that only starting edge of the pulse represents instantaneous value of the baseband signal $x(t)$. Also the flat top pulse of $g(t)$is mathematically equivalent to the convolution of instantaneous sample and a pulse $h(t)$ as depicted in Fig. 7.16.
This means that the width of the pulse in *g(t)* is determined by the width of h(f) and the sampling instant is determined by delta function.

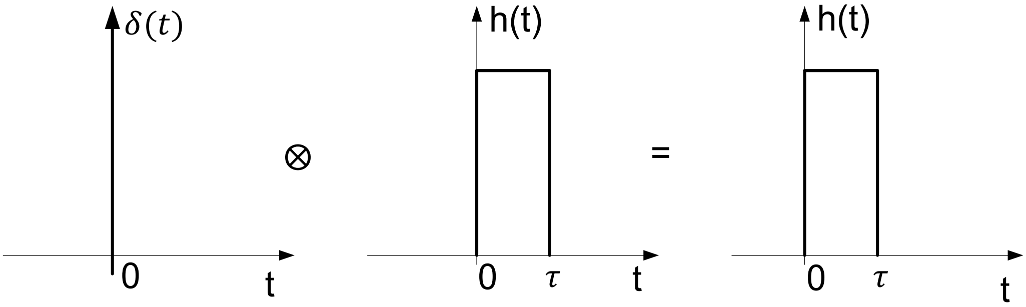


Figure 7.16: Convolution of any function with delta function is equal to that function

In Fig. 7.15(b), the starting edge of the pulse represents the point where baseband signal is sampled and width is determined by function $h(t)$. Therefore g(t) will be expressed as

$$g(t) \ = \ s(t) \otimes h(t) \qquad 7.39$$

This equation has been explained in Fig.7.17
Now, from the property of delta function, we know that for any function

$$f(t)\otimes \delta(t) = f(t) \qquad 7.40$$

This property is used to obtain flat top samples. It may be noted that to flat top sampling, we are not applying the Eq (7.40) directly here i.e we are applying a modified form of Eq (7.40). This modified equation is Eq (7.39).
Thus, in this modified equation, we are taking $s(t)$in place of delta function $\delta(t)$ observe that $\delta(t)$ is a constant amplitude delta function whereas $s(t)$is a varying amplitude train of impulses. This means that we are taking $s(t)$which is an Instantaneously sampled signal and this is convolved with function $h(t)$ as in Eq (7.39). Therefore, on convolution of s(t) and h(t), we get a pulse whose duration is equal to $h(t)$ only but amplitude is defined by $s(t)$
Now, we know that the train of impulses may be represented mathematically as

$$\delta T_s(t) = \sum_{n=-\infty}^{\infty} \delta(\tau - n\, T_s) \qquad 7.41$$

The signal $s(t)$ is obtained by multiplication of baseband signal $x(t)\ and\ \delta T_s(t)$. Thus

$$S(t) = x(t).\,\delta_{T_S}(t) \qquad 7.42$$

$$= X(t) = \sum_{n=-\infty}^{\infty} \delta(\mathrm{t} - n\,T_s)$$

$$s(t) = \sum_{n=-\infty}^{\infty} x\,(nT_s)\,\delta(\mathrm{t} - n\,T_s) \qquad 7.43$$

Now, sampled signal g(t) is given as (Eq 7.39)

$$g(t) \;=\; s(t) \otimes h(t) \qquad 7.44$$

$$\text{Or} \quad \mathrm{g(t)} = \int_{-\infty}^{\infty} \mathrm{s(\tau)h(\,t - \tau)d\tau}$$

$$\text{Or} \quad \mathrm{g(t)} = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} X(nT_s)\,\delta(\mathrm{t} - n\,T_s)h(t - \tau)\mathrm{d\tau} \qquad 7.45$$

$$\text{Or} \quad \mathrm{g(t)} = \sum_{n=-\infty}^{\infty} X(nT_s) \int_{-\infty}^{\infty} \delta(\mathrm{t} - n\,T_s)h(t - \tau)\mathrm{d\tau} \qquad 7.46$$

According to shifting property of delta function, we know that

$$\int_{-\infty}^{\infty} (t)\,\delta(\mathrm{t} - t_0) = f(t_0) \qquad 7.47$$

Using equation (7.46) and (7.47), we get

$$\mathrm{g(t)} = \sum_{n=-\infty}^{\infty} x(nT_s)\,h(\mathrm{t} - n\,T_s)$$

This equation represent value of g(t) in terms of sampled value $x(nT_s)$ and function $h(t - nT_s)$ for flat top sampled signal.

Now, again from equation (7.39), we have

$$\mathrm{g(t) = s(t)\otimes h(t)}$$

Taking Fourier transform of both sides of above equation, we get

$$G(f) = S(f)H(f) \qquad 7.48$$

We know that $S(f)$ is given as

$$S(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \qquad 7.49$$

Therefore, Eq (7.48) becomes

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)\,H(f) \qquad 7.50$$

Thus, spectrum of flat top sampled signal:

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)\, H(f) \qquad 7.51$$

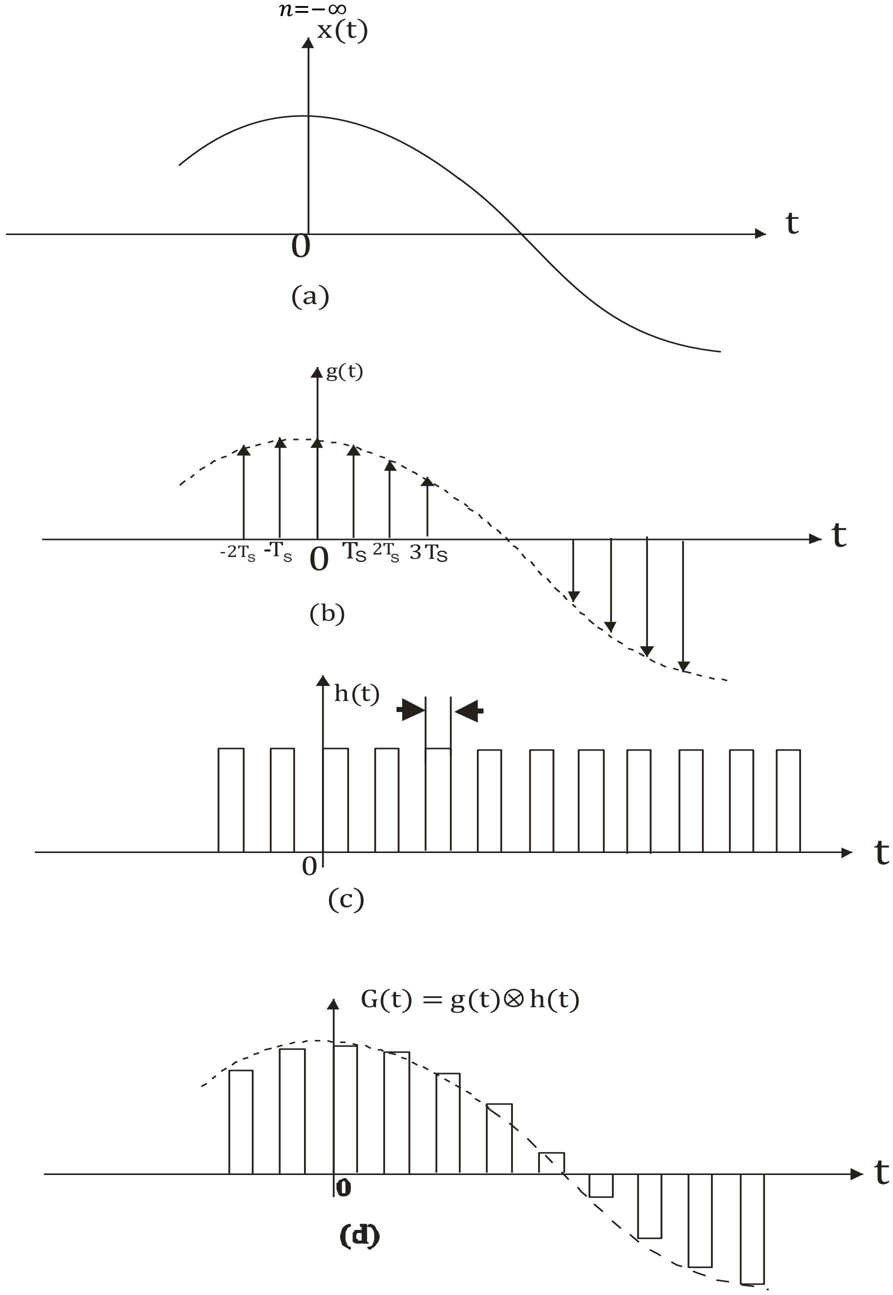


Figure 7.17: (a) Baseband signal $x(t)$ (b) Instantaneous sample single $s(t)$ (c) Constant pulse width function $h(t)$ (a) Flat top sampled singl g(t) obtained through convolution of $\mathbf{h(t)}$ and $\mathbf{s(t)}$

7.9 Aperture Effect

The spectrum of flat top sampled signal is expressed as

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)\, H(f) \qquad 7.52$$

The equation shows that the signal $\mathrm{g(t)}$ is obtained by passing the signal $\mathrm{s(t)}$ through a fitter having PLACE function $\mathrm{h(t)}$. The corresponding impulse response $\mathrm{h(t)}$ in time-domain is as shown in Fig. 7.18(a). This is one pulse rectangular shown in Fig. 7.17(c). Each sample of $\mathrm{x(t)}$ $[\mathrm{i.e.}, s(\mathrm{t})]$ is convolved with this pulse. Eq 7.52 represents that the spectrum of this rectangular pulse is multiplied with that of $\mathrm{s(t)}$
Fig. 7.18(b) shows the spectrum of one rectangular pulse of $\mathrm{h(t)}$

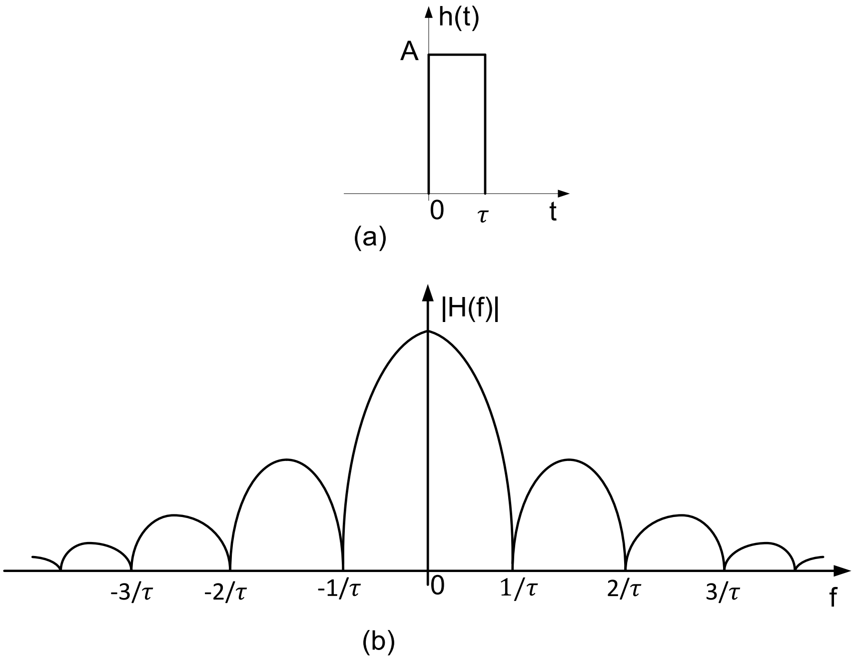


Figure 7.18: (a) One pulse of rectangular pulse train (b) Spectrum of the pulse shown in figure (a)

We know that the spectrum of a rectangular pulse is expressed as
$H(f) = \tau . sinc\,(f.t)e^{-j\pi f\tau}$ $(\therefore A = 1)$ 7.53
Hence, from Fig. 7.18(b), it may be observed that by using flat top samples in amplitude distortion is introduced in the reconstructed signal x(t) from g(t). In fact, the high frequency roll-off of H(f) acts like a low-pass filter and thus attenuates the upper portion of message signal spectrum. These high frequencies of x(t) are affected. This type of effect is known as ***aperture effect*.**

Now, as the duration 'τ' pulse increases , the aperture effect is more prominent Hence, during reconstruction an equalizer is needed to compensate for this effect .As depicted in Fig.7.19, the receiver contains a low-pass reconstruction filter with cutoff frequency slightly high than the maximum frequency present in message signal .The equalizer compensates. For aperture effect. It also compensates for the attenuation by a low-pass reconstruction filter.

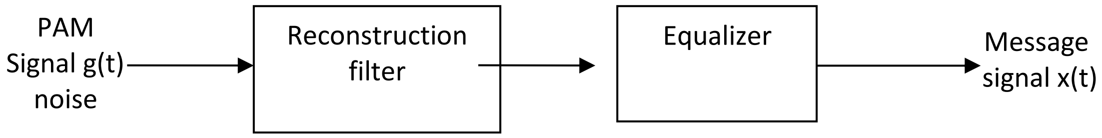


Figure 7.19: Recovering x(t) at receiver

From Eq (7.53), it may be noted that the sample function h(t) acts like a low-pass filter where Fourier transform as express as

$$\mathrm{H(f)} = \tau.\sin \mathrm{c}\ (\mathrm{f.t})\ \mathrm{e}^{-\mathrm{j}\pi \mathrm{f}\tau} \qquad 7.54$$

This spectrum has been plotted in Fig. 7.18.

Equalizer used in casecade with the reconstruction filter as the frequency increases in such a way as to compensate for the aperture effect.

Also, the transfer function of the equalizer is express as

$$H_{eq}(f) = \frac{K.\mathrm{e}^{-\mathrm{jzf\,x}}}{H(f)} \qquad 7.55$$

Here 'τ_d' is known as the delay introduction by low-pass filter which is equal to $\frac{\tau}{2}$.

Therefore,

$$H_{eq}(f) = \frac{K.\mathrm{e}^{-\mathrm{j}\pi \mathrm{f}\tau}}{\tau sinc(f\tau)\mathrm{e}^{-\mathrm{j}\pi \mathrm{f}\tau}} \qquad 7.56$$

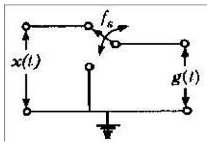
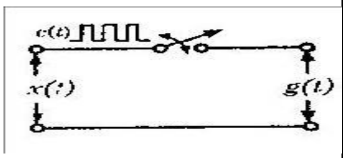
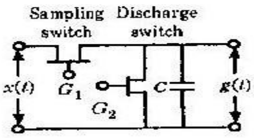
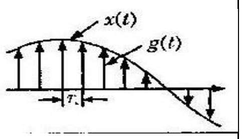
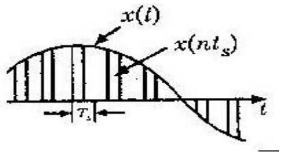
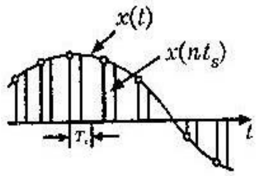
$$H_{eq}(f) = \frac{K.\mathrm{e}^{-\mathrm{j}\pi \mathrm{f}\tau}}{\tau sinc(f\tau)} \qquad 7.57$$

Which is the transfer function of an equalizer.

7.10 Comparison of Various Sampling Techniques

We can compare various sampling techniques on the basis of their method, noise interference and spectral properties etc.The table (7.1) lists some of the important points of comparison of generation.

Table 7.1. Comparison of three sampling techniques

S/N	Parameter of comparison	Idea or instantaneous sampling	Natural sampling	Flat top sampling
1	Sampling principle	It uses multiplication	It uses chopping principle	It uses sample and hold circuit
2	Generation			
3	Waveforms involved			
4	Feasibility	This is not a practically possible method	This method is used practically	This method is also used practically
5	Sampling rate	Sampling rate tends to infinity	Sampling rate satisfies Nyquist criteria	Samples rate satisfies Nyquist criteria
6	Noise interference	Noise interference is maximum	Noise interference is maximum	Noise interference is maximum
7	Time domain representation	$g(t) = \sum_{n=-\infty}^{\infty} x\,(nT_s)\,\delta\,(t - nT_s)$	$g(t) = \frac{\tau A}{T_s}\sum_{n=-\infty}^{\infty} \sin c\,(nf_s\tau)\, e^{j2\pi n f_s \cdot}$	$g(t) = \sum_{n=-\infty}^{\infty} x\,(nT_s)\delta\,(t - nT_s)$
8	Frequency domain representation	$G(f) = f_s \sum_{n=-\infty}^{\infty} X\,(f - nf_s)$	$G(f) = \frac{\tau A}{T_s}\sum_{n=-\infty}^{\infty} \sin c\,(nf_s.\tau) X\,(f - nf_s.)$	$G(f) = f_s \sum_{n=-\infty}^{\infty} X\,(f - nf_s.)\ H(f)$

Example 7.6: Fig. 7.20 shows the spectrum of an arbitrary signal x(t). this signal is sampled at the Nyquist rate with a periodic train of rectangular pulses of duration 50/3 milliseconds. Determine

The spectrum of the sampled signal for frequencies up to 50 Hz giving relevant expression.

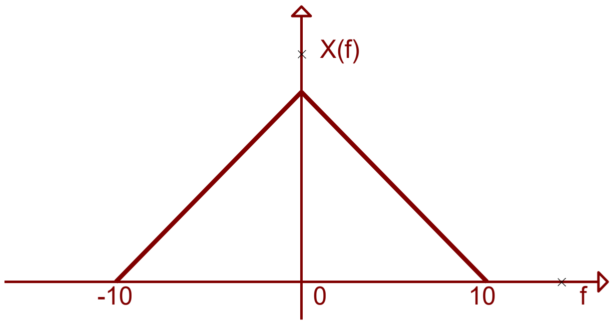


Figure 7.20

Solution: From Fig. 7.20, it may be observed that the signal is band limited to 10 Hz.

$$f_m = 10 \text{ Hz}$$

So the Nyquist rate is $2f_m$

= 2 x 10 = 20 Hz

Since the signal is sampled at the Nqyuist rate, the sampling frequency would be f_s = 20 Hz

Given that the rectangular pulses are used for sampling i.e, flat top sampling is used

The spectrum of the flat top sampled signal is given as

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)\,H(f)$$

Value of H(f) is expressed as

$$H(f) \;=\; \tau\, sinc(f\tau)e^{-j\pi f\tau}$$

Here τ is the width of the rectangular pulse for sampling

The given value of rectangular sampling pulse duration is 50/3 milliseconds i.e

$$\tau = \frac{50}{3} \times 10^{-3} = \frac{0.05}{3}\text{ s}$$

Substituting the value of τ in equation (ii), we get

$$H(f) = \frac{0.05}{3} sinc\{\frac{0.05f}{3}\}\, e^{-j0.05\pi f/3}$$

Again, putting this value of H(f) and fs in equation (i), we get

$$G(f) = 20 \sum_{n=-\infty}^{\infty} X(f - 20n)\frac{0.05}{3} sinc\{\frac{0.05f}{3}\}\, e^{-j0.05\pi f/3}$$

$$G(f) = \frac{1}{3} \sum_{n=-\infty}^{\infty} X(f - 20n)\frac{0.05}{3} sinc\left(\frac{0.05f}{3}\right)\, e^{-j0.05\pi f/3} \qquad (\because \text{f}_\text{s} = 20 \text{ Hz})$$

This expression gives the spectrum up to 60 Hz (since n=$\pm$ 3) for the sampled signal

Example 7.7: A flot top sampling system samples a signal of maximum frequency 1Hz with 9.5 Hz sampling frequency. The duration of the pulse is 0.2

seconds. Compute the amplitude distortion due to aperture effect at the highest signal frequency. Also determine the equalization characteristic.

Solution: Given that sampling frequency
$f_s = 2.5\,\text{Hz}$
Maximum signal frequency
$f_{max} = 1\text{ Hz}$
And pulse width $\tau = 0.2\text{ s}$
We know that the aperture effect is expressed by a transfer H(f) as

$$H(f) = \tau\, sinc(f\tau)e^{-j\pi f\tau}$$

The magnitude of this equation will be

$$|\mathrm{H(f)}| = \tau\,\mathrm{sinc}\;(\mathrm{f}\,\tau)$$

$$|\mathrm{H(f)}| = 0.2\mathrm{sinc}\;(\mathrm{f}\times 0.2)$$

Now, aperture effect at the highest frequency will be obtained by putting $f = f_{max} = 1\text{Hz}$ in equation (i) i.e.,

$$|H(1)| = 0.2\,\mathrm{sinc}(0.2)$$

$$= 0.18709$$

Or $|H(1)| = 18.70\%$ Ans

Also, the equalizer characteristics is expressed as

$$H_{eq}(f) = \frac{K}{\tau sinc(f\tau)}$$

Substituting, τ = 0.2 second and assuming
K = 1, the last equation becomes

$$H_{eq}(f) = \frac{1}{0.2\,\mathrm{sinc}(0.2f)}$$

This equation is the plot of $H_{eq}(f)$ versus f and it represents the equalization characteristics to overcome aperture effect

7. 11 Analog Pulse Modulation Methods

We know that in analog modulation systems, some parameter of a sinusoidal carrier is varied according to the instantaneous value of the modulating signal.

7.11.1. Sampling Process

Sampling is a signal processing operation that helps in sensing the continuous time signal values with instant of time. The sampling sequence will have amplitudes equal to signal values at the sampling instants and undefined at all other times
This process can be conveniently performed using PAM. The sampling process can be treated as an electric switching action as shown in Fig. 7.21

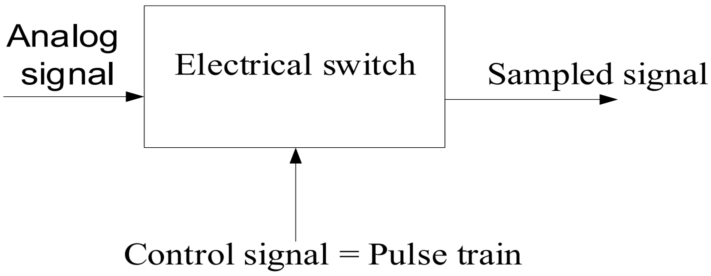


Figure 7.21 Illustration of sampling process

The continuous time signal to be sampled is applied to the input terminal. The pulse train is applied as the control signal of the switch when the pulse occurs; the switch is in ON condition, that is, acts as short circuit between input and output terminals. The output value will therefore be equal to input. During the other intervals of the pulse train, the switch is in OFF condition that is acts as open circuit. The output is therefore undefined. The output of the switch will be essentially a PAM signal. Any active device like diode, transistor or FET can be use as a switch

However, the sampling theorem will decide the periodicity associated with the pulse train, the width of the pulse does not influences the amplitudes of the sampled value. Even though this is not obvious in the time domain, it can be understood by observing the frequency domain behaviour of the PAM process due to the convolution of sinc function of pulse train with input signal spectrum (The sampled signal can be seen clearly as the dot in Fig. 7.11). To minimize this effect for all practical processing as the pulse width approaches zero so that the pulse train becomes ON impulse train. The Fourier transform of an impulse train is also an impulse train in the frequency domain hence convolution does not affect the sampled signal, rather it leads only to the periodicity of the spectrum.

7.11.2. Signal Sampling and Reconstruction Using Fourier Series

An analog signal can be converted to a digital signal by the process of sampling. When a continuous or analog signal is represented by a set of sampled values, it is said to have been sampled. The basic process of sampling is the getting of an analog signal by a periodic pulse train which will only allow the signal through when each pulse is ON.

The analog signal being sampled is referred to as the baseband signal or the modulating signal $f(t)$ while the getting signal is referred to as the sampling finite or carrier $s(t)$**.** $s(t)$ has pulse of constant high, length $'\tau'$ second and separation T_s second. Sampling intervals is T_s sampling frequency $f_S = \frac{1}{T_S}$.

The sampling process is carried out by multiplying the two signals to give the sampling signal or digital signal $v_s(t)$. Sampling however, is the process of converting a continuous signal into a discrete signal and the circuit that dose this is sampling & hold circuit. Given a system with an input signal $v_m(t)$ to sample the signal, we have to use the Fourier series mathematical analysis to achieve our desire as shown in the analysis below.

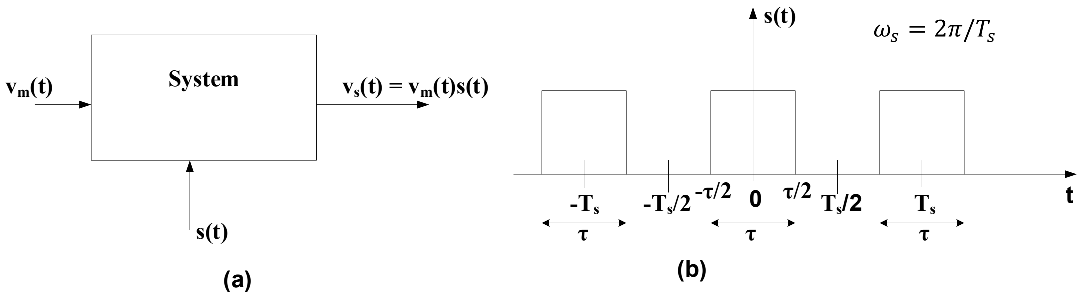


Figure 7.22 Sample and Hold circuit (a), Sampling signal (b)

$v_s(\mathrm{t}) \; = \; v_m(\mathrm{t}) \times \mathrm{s(t)}$

Therefore the sampling signal pulse can be modified shown in the Fig.7.23

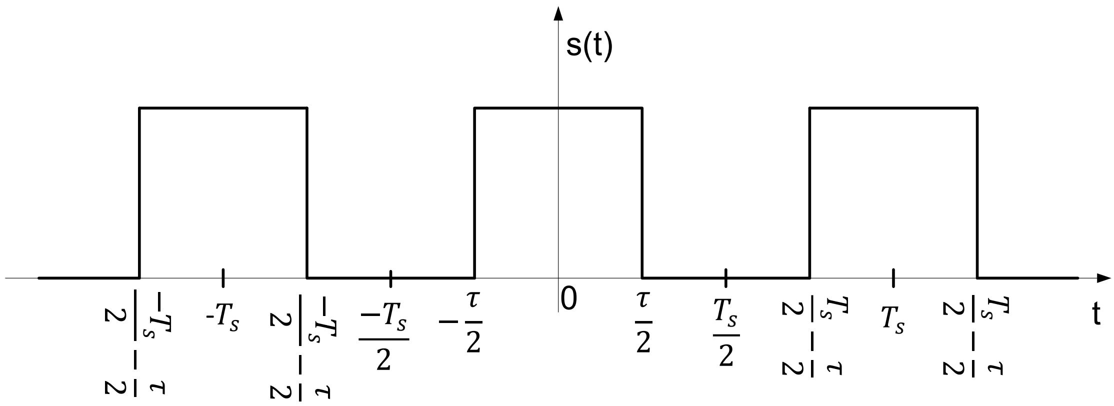


Figure 7.23 Sampling signal

Expressing the sampling signal above in Fourier series, from 1st principle, s(t) can be summed up as shown in Eq (7.58).

$$s(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos n\omega_s t + \sum_{n=1}^{\infty} b_n \sin n\omega_s t \qquad 7.58$$

$$But \qquad a_n = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} s(t)\, cos\, n\omega_s tdt \qquad 7.79$$

Because the sampling signal is a periodic function, assuming the amplitude is unity, we can express its Fourier series for a periodic as follows

$$s(t) = \begin{cases} 0 & \frac{-T_s}{2} \leq t \leq \frac{-\tau}{2} \\ 1 & \frac{-\tau}{2} \leq t \leq 0 \\ 1 & 0 \leq t \leq \frac{\tau}{2} \\ 0 & \frac{\tau}{2} \leq t \leq \frac{T_s}{2} \end{cases}$$

$$a_n = \frac{1}{T_s}\int_{\frac{-\tau}{2}}^{\frac{\tau}{2}} \cos n\omega_s t \, dt \qquad 7.60$$

$$= \frac{1}{T_s} \sin \frac{n\omega_s t}{n\omega_s}\Big|_{\frac{-\tau}{2}}^{\frac{\tau}{2}}$$

$$= \frac{1}{T_s}\left[\sin \frac{\left(n\omega_s\left(\frac{\tau}{2}\right)\right)}{n\omega_s} - \sin \frac{\left(n\omega_s\left(\frac{-\tau}{2}\right)\right)}{n\omega_s}\right]$$

$but \quad \omega_s = \frac{2\pi}{T_s}$

7.61

$$a_n = \frac{1}{T_s}\left[\sin \frac{\left(2\pi n\left(\frac{\tau}{2T_s}\right)\right)}{n\omega_s} - \sin \frac{\left(-2\pi n\left(\frac{\tau}{2T_s}\right)\right)}{n\omega_s}\right]$$

$$a_n = \frac{1}{T_s}\left[\frac{\sin n\pi \frac{\tau}{T_s}}{n\omega_s} - \frac{\sin\left(-n\pi \ \frac{\tau}{T_s}\right)}{n\omega_s}\right]$$

$$= \frac{1}{T_s}\left[\frac{\sin n\pi \frac{\tau}{T_s}}{n\omega_s} + \frac{\sin\left(n\pi \ \frac{\tau}{T_s}\right)}{n\omega_s}\right]$$

$$= \frac{1}{T_s}\left[\frac{2\sin n\left(\pi \ \frac{\tau}{T_s}\right)}{n\omega_s}\right] \qquad 7.62$$

But from equation 7.61

$$a_n = \frac{1}{T_s}\left[\frac{2\sin\left(n\pi\ \frac{\tau}{T_s}\right)}{n\pi\ \frac{2\,\pi}{T_s}}\right]$$

$$= \frac{2\sin\left(n\pi\frac{\tau}{T_s}\right)}{2n\pi}$$

$$= \frac{\sin\left(n\pi\ \frac{\tau}{T_s}\right)}{n\pi} \qquad 7.63$$

Multiplying numerator and denominator of Eq (7.63) by $\frac{\tau}{T_s}$

$$a_n = \frac{\tau}{T_s} \times \frac{\sin\left(n\pi\ \frac{\tau}{T_s}\right)}{n\pi\ \frac{\tau}{T_s}} = \frac{\tau}{T_s}\text{sinc}\left(\frac{n\pi\tau}{T_s}\right) \qquad 7.64$$

$$Because\ \frac{\sin x}{x} = sinc(x) \qquad 7.65$$

$$Also\ b_n = \frac{1}{T_s}\int_{-\frac{T_s}{2}}^{\frac{T_s}{2}} sin(n\omega_s)t\ dt = 0 \qquad 7.66$$

Eq (7.66) is true because s(t) is an even function as discuss in chapter 6
Hence putting n = 1, 2, 3, ……… and noting that sinc (0) = 1

$$s(t) = \frac{\tau}{2T_s} + \frac{\tau}{T_s} sinc\frac{\pi\tau}{T_s}\cos\omega_s t + \frac{\tau}{T_s} sinc\frac{2\pi\tau}{T_s}\cos 2\omega_s t + \frac{\tau}{T_s} sinc\frac{3\pi\tau}{T_s}\cos 3\omega_s t + \cdots \ldots\ldots\ldots\ldots \qquad 7.67a$$

Since $\tau << T_s$ it means that $\frac{\tau}{T_s} \to 0$

Which means that Eq (7.67) a will be reduced to

$$s(t) = \frac{\tau}{2T_s} + \frac{\tau}{T_s}\cos\omega_s t + \frac{\tau}{T_s}\cos 2\omega_s t + \frac{\tau}{T_s}\cos 3\omega_s t \ldots. \qquad 7.67b$$

If by assumption we have the baseband signal $v_m(t)$ such that $v_m(t) = A\cos\omega_m t$
Then the frequency spectrum will be as shown in Fig. 7.24

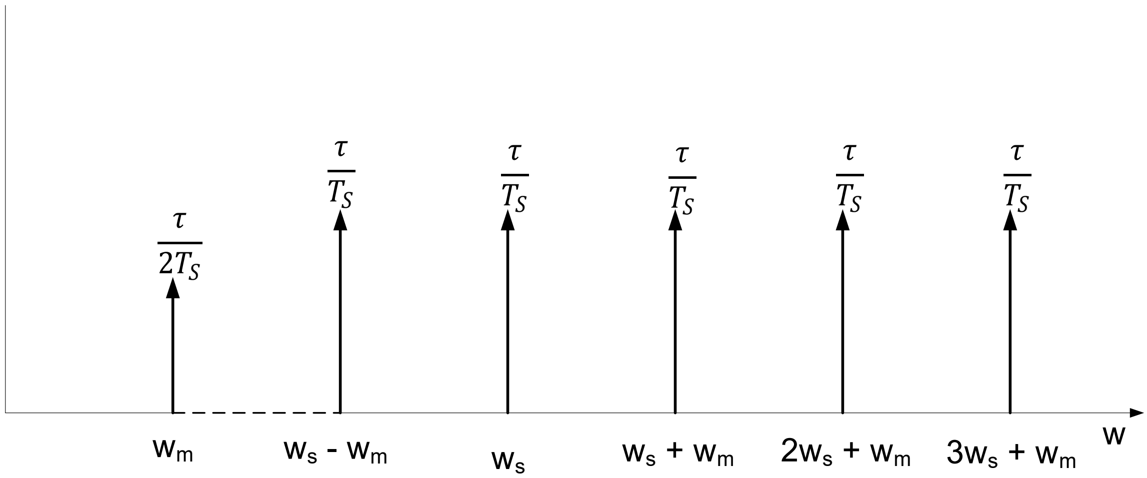


Figure 7.24 Frequency spectrum of $v_s(t)$ with its radian frequency

Because revealed that $v_s(t) = v_m(t) \times s(t)$, therefore

$$v_s(t) = A\cos\omega_m t\left[\frac{\tau}{2T_s} + \frac{\tau}{T_s}\cos\omega_s t + \frac{\tau}{T_s}\cos 2\omega_s t + \cdots \ldots\right] \qquad 7.71$$

$$v_s(t) = \frac{A\tau}{2T_s}\cos\omega_m t + \frac{A\tau}{2T_s}\cos\omega_m t\cos\omega_s t + \frac{A\tau}{T_s}\cos 2\omega_m t\cos 2\omega_s t + \qquad 7.72$$

Recalled that from elementary trigonometry identities that

$$\left.\begin{array}{l}[\cos(A+B) + \cos(A-B)] = 2\cos A cos\ B \\ \frac{1}{2}[\cos(A+B) + \cos(A-B) = \cos AcosB\end{array}\right\} \qquad 7.73$$

$$v_s(t) = \frac{A\tau}{2T_s}\cos\omega_m t + \frac{A\tau}{2T_s}\cos(\omega_s+\omega_m)t + \frac{A\tau}{2T_s}\cos(\omega_s-\omega_m)t + \frac{A\tau}{2T_s}\cos(2\omega_s+\omega_m)t$$

$$+\frac{A\tau}{2T_s}\cos(2\omega_s-\omega_m)t + \frac{A\tau}{2T_s}\cos(3\omega_s+\omega_m)t + \frac{A\tau}{2T_s}\cos(3\omega_s-\omega_m)t + \cdots \ldots\ldots\ldots..7.74$$

Now the frequency spectrum of vs(t) will be shown in the Fig.7.25

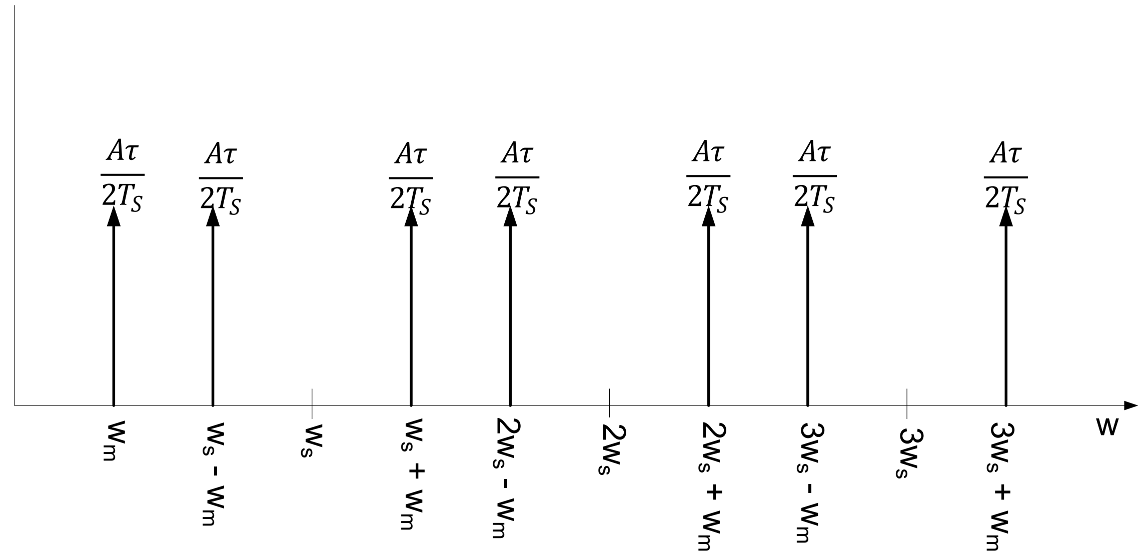


Figure 7.25 Frequency spectrum of $\mathbf{v_s(t)}$ with its radian frequency

If $\omega_s = 2\pi f_s \; and \; \omega_m = 2\pi f_m$ 7.75

Eq (7.24) can be modify to

$$v_s(t) = \frac{A\tau}{2T_s}\cos 2\pi f_m t + \frac{A\tau}{2T_s}\cos 2\pi(f_s + f_m)t + \frac{A\tau}{2T_s}\cos 2\pi(f_s - f_m)t$$

$$+\frac{A\tau}{2T_s}\cos 2\pi(2f_s - f_m)t + \frac{A\tau}{2T_s}\cos 2\pi(3f_s + f_m)t + \frac{A\tau}{2T_s}\cos(3f_s - f_m)t + \cdots \ldots \ldots 7.76$$

The frequency spectrum in (Hz) will now be

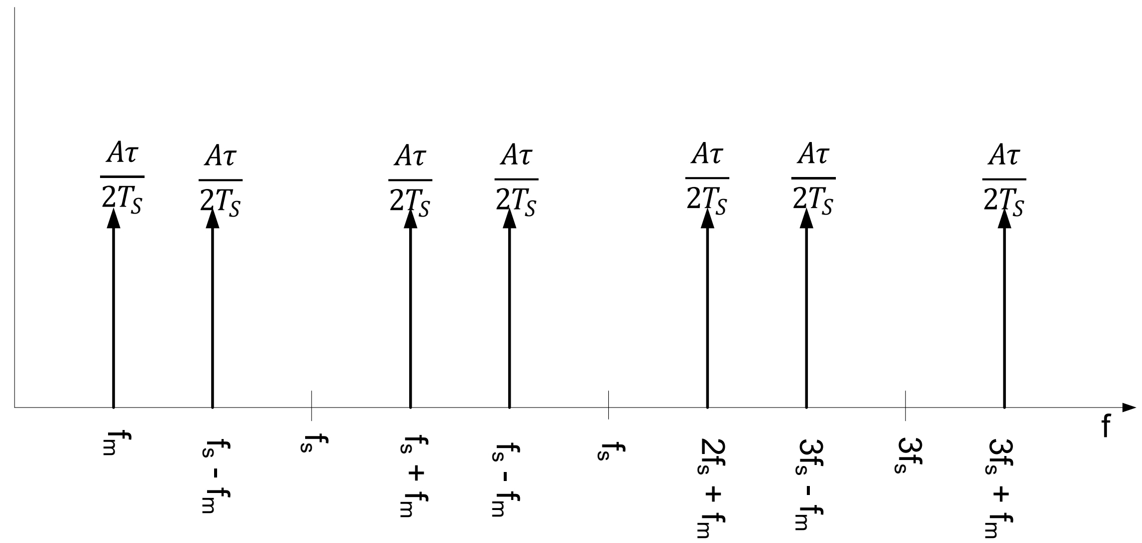


Figure 7.26 Frequency spectrum of $\mathbf{v_s(t)}$

From the analysis, the signal v_m (t) is recoverable from v_s (t) by passing v_s (t) through a low pass filter with cut off frequency greater than ω_m or f_m (side band frequency) but less than ω_s - ω_m, or f_s - f_m that is ω_m or f_m must be less than ω_s of f_s, . It means that ω_m must be less than or equal to ½ω_s. This is known as Nyquist sampling criterion or simply Nyquist criterion.

The Nyquist criterion states that the sampling frequency must be greater than or equal to twice the maximum frequency in the baseband signal. If this criterion is not observed, then there would be an over lapping known as aliasing. The sideband signal v_m(t) will become unrecoverable from v_s (t).

Assuming that $v_m(t) = A_1\ cos\omega_1 t + A_2\ cos\omega_2 t$ 7.77

$\omega_1 = 2\pi f_1\ and\ \omega_2 = 2\pi f_2$

$$v_s(t) = [A_1\ cos\omega_1 t\ +\ A_2\ cos\omega_2 t] \times \left[\frac{\tau}{2T_s} + \frac{\tau}{T_s}\cos\omega_s t + \frac{\tau}{T_s}\cos 2\omega_s t\right]$$

$$v_s(t) = \frac{\tau}{2T_s}[A_1\cos\omega_1 t +\ A_2\ cos\omega_2 t + 2A_1\cos\omega_1 t\cos\omega_s t + 2A_2 cos\omega_2 t\cos\omega_s t + 2A_1 cos\omega_1 t\cos 2\omega_s t + 2A_2\ cos\omega_2 t\cos 2\omega_s t]$$

$$v_s(t) = \frac{\tau}{2T_s}[A_1\cos 2\pi f_1 t + A_2\cos 2\pi f_2 t + A_1\cos 2\pi\,(f_s - f_1)t + A_1\cos 2\pi\,(f_s + f_1)t + A_2\cos 2\pi\,(f_s - f_2)t + A_2\cos 2\pi\,(f_s + f_2)t + A_1\cos 2\pi\,(2f_s - f_1)t + A_1\cos 2\pi\,(2f_s + f_1)t + A_2\cos 2\pi\,(2f_s - f_2)t + A_2\cos 2\pi\,(2f_s + f_2)t] \qquad 7.78$$

The frequency spectrum will now be as shown in Fig. 7.27 assuming $A_1 = A_2$:

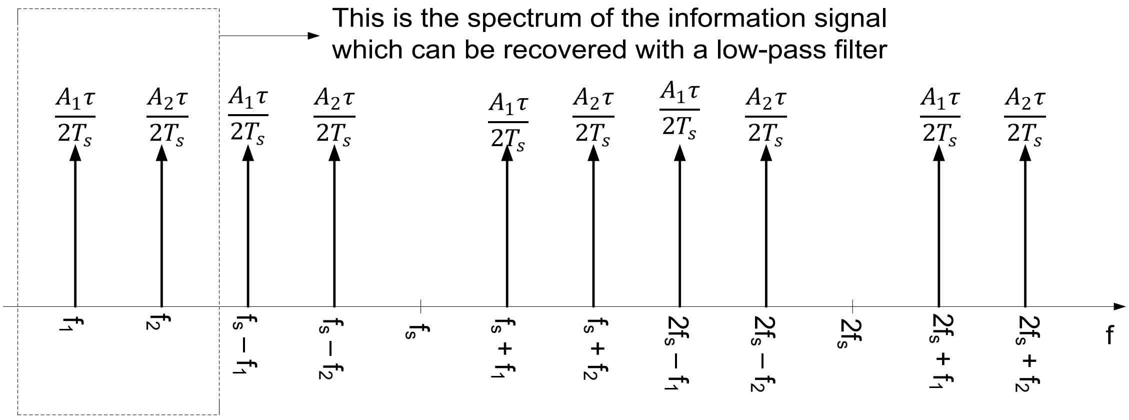


Figure 7.27 Frequency spectrum of v_s(t)

If Nyquist criterion is met, the spectrum will be as shown the Fig. 7.29 i. e, $f_s \geq 2f_{max}$, if $f_1 > f_2$, $f_{max} = f_1$ otherwise it is

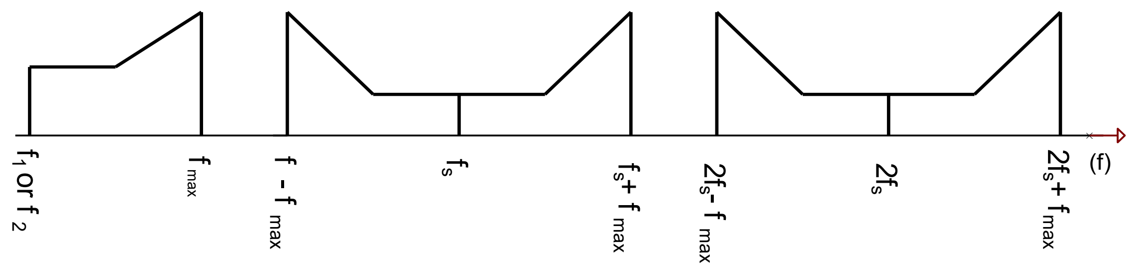


Figure 7.28

If $f_s < 2f_{max}$, then there will aliasing know as overlap

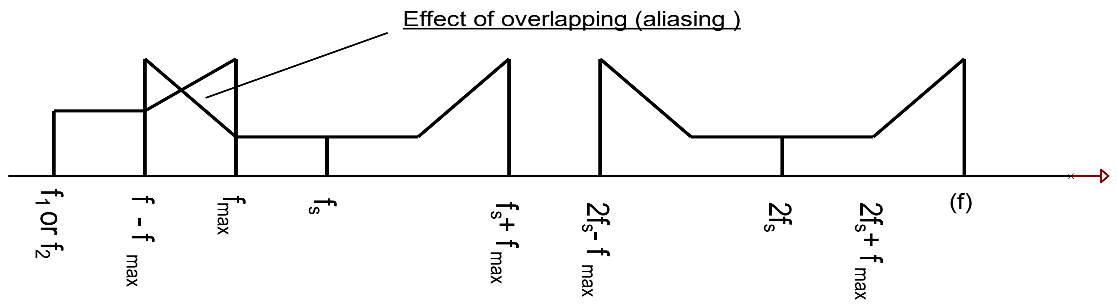


Figure 7.29

Aliasing can be a serious problem if the baseband includes any unwanted noise components higher than its highest frequency. This gets mixed up with the baseband on recovery.

To avoid this, a low passed filter is used to band limit the baseband signal before sampling and such a filter is called anti-aliasing.

Example 7.8

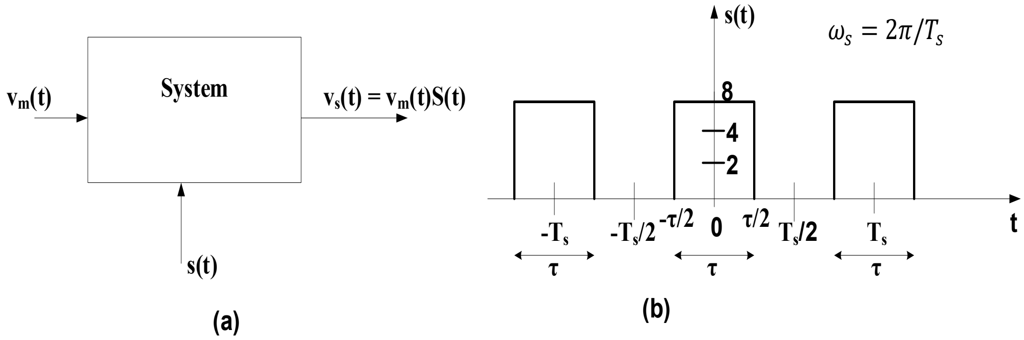


Figure 7.30

Fig. 7.30 (a) above shows a system. The signal s(t) is a train of pulse as in Fig. 7.30 (b). f_s is the frequency of s(t). The other signal is $v_m(t) = 6\cos 2\pi f_m t + 4\cos 12\pi f_m t$, $f_m = 18\text{ kHz}$

(i) Obtain the Fourier series of (t) from first principle
(ii) Suppose $\tau \ll T_s$ simply the expression of s(t) obtain in (i) above and sketch its frequency spectrum. Obtain an expression for the signal $v_s(t)$, using only the first two terms in (ii).
(iii) If $f_s = 18\text{ MHz}$, sketch the frequency spectrum of $v_s(t)$.
(iv) Discuss the recovery of $v_m(t)$. State problem(s) you envisaged (if any).
(v) Repeat (iv) with f_s = 200 kHz solution.

Recall that from Eq (7.58),

$$s(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}[a_n \cos n\omega_s t + b_n \sin n\omega_s t]$$

Also, from Eq (7.59), s(t) function is:

$$s(t) = \begin{cases} 0 & \frac{-T_s}{2} \le t \le \frac{-\tau}{2} \\ 8 & \frac{-\tau}{2} \le t \le 0 \\ 8 & 0 \le t \le \frac{\tau}{2} \\ 0 & \frac{\tau}{2} \le t \le \frac{T_s}{2} \end{cases}$$

Whose Fourier series can be found using Eq (7.60) as shown below.

$$a_n = \frac{1}{T_s}\int_{\frac{-T_s}{2}}^{\frac{T_s}{2}} s(t)\cos n\omega_s t\, dt$$

$$a_n = \frac{1}{T_s}\left[\int_{\frac{-\tau}{2}}^{0} 8\cos n\omega_s t\,\mathrm{dt} + \int_{0}^{\frac{\tau}{2}} 8\cos n\omega_s t\, dt\right]$$

$$a_n = \frac{8}{T_s}\left[\frac{\sin n\omega_s t}{n\omega_s}\Bigg|\begin{matrix}0\\ \frac{-\tau}{2}\end{matrix} + \frac{\sin n\omega_s t}{n\omega_s}\Bigg|\begin{matrix}\frac{\tau}{2}\\ 0\end{matrix}\right]$$

$$= \frac{8}{T_s}\left[\frac{\sin 0}{n\omega_s} - \frac{\sin n\omega_s\left(\frac{-\tau}{2}\right)}{n\omega_s} + \frac{\sin n\omega_s\left(\frac{\tau}{2}\right)}{n\omega_s} - \frac{\sin 0}{n\omega_s}\right]$$

$$= \frac{8}{T_s}\left[\frac{\sin n\omega_s\left(\frac{\tau}{2}\right)}{n\omega_s} + \frac{\sin n\omega_s\left(\frac{\tau}{2}\right)}{n\omega_s}\right]$$

$$= \frac{8}{T_s}\left[\frac{2sin\left(n\omega_s(\frac{\tau}{2})\right)}{n\omega_s}\right]$$

From Eqs (7.64) , (7.65) and (7.66) we have that

$$a_n = \frac{8}{T_s}\left[\frac{\left(\frac{\tau}{2}\right)}{\frac{\tau}{2}n\omega_s}.\, 2\sin n\omega_s(\frac{\tau}{2})\right]$$

$$a_n = \frac{8}{T_s}\times\frac{\tau\sin\left(n\omega_s\left(\frac{\tau}{2}\right)\right)}{\frac{\tau}{2}n\omega_s} = \frac{8\tau}{T_s}\,\mathrm{sinc}\left(n\omega_s\left(\frac{\tau}{2}\right)\right)$$

For n= 0, sinc 0 = 1

$$a_0 = \frac{8\tau}{T_s}$$

for n = 0, 1, 2, 3 ……..

$$S(t) = \frac{4\tau}{T_s} + \frac{8\tau}{T_s}\,\mathrm{sinc}\left(\omega_s\left(\frac{\tau}{2}\right)\right)\cos\omega_s t + \frac{8\tau}{T_s}\,\mathrm{sinc}\left(2\omega_s\left(\frac{\tau}{2}\right)\right)\cos 2\omega_s t$$
$$+ \frac{8\tau}{T_s}\,\mathrm{sinc}\left(3\omega_s\left(\frac{\tau}{2}\right)\right)\cos 3\omega_s t + \cdots$$

$if \quad \tau << T_s \, , \qquad \frac{\tau}{T_s} \to 0, \qquad \sin c(0) = 1,$

but $\omega_s = \frac{2\pi}{T_s},$ then we have $\text{sinc}\left(n\omega_s\left(\frac{\tau}{2}\right)\right) = \text{sinc}\left(\frac{2\pi}{T_s}\left(\frac{\tau}{2}\right)\right)$

Then, $\text{s(t)} = \frac{4\tau}{T_s} + \frac{8\tau}{T_s}\cos\omega_s t + \frac{8\tau}{T_s}\cos 2\omega_s t + \frac{8\tau}{T_s}\cos 3\omega_s t \ldots \ldots \ldots$

$s(t)$ spectrum will be as shown in the Fig.7.31 below

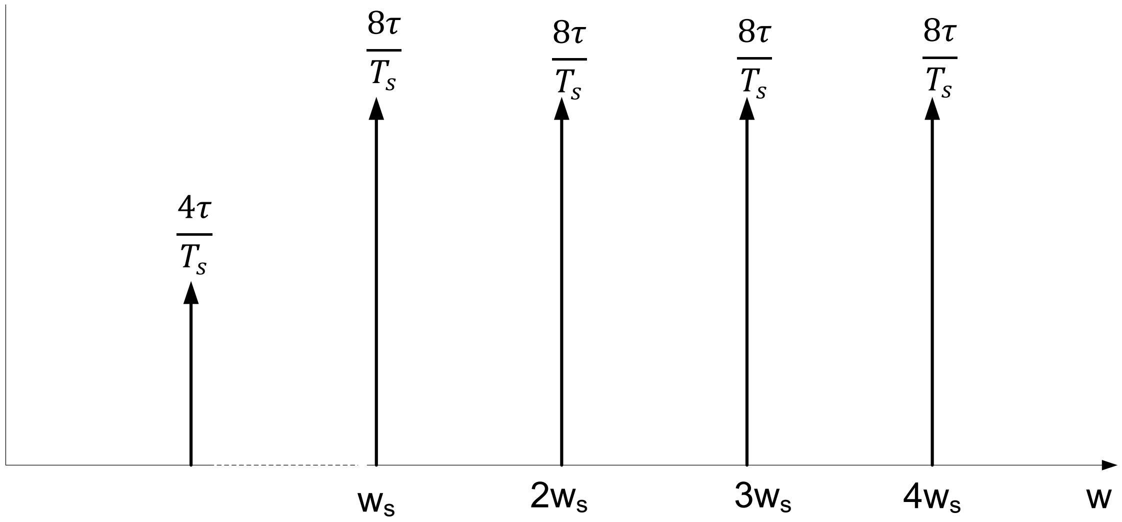


Figure 7.31

However, $v_s(t) = v_m(t) \times s(t)$

But $v_m(\text{t}) = 6\cos 2\pi f_m \text{t} + 4\cos 12\pi f_m \text{t}$

Therefore $v_s(t) = (6\cos 2\pi f_m \text{t} + 4\cos 12\pi f_m \text{t})\text{x}(\frac{4\tau}{T_s} + \frac{8\tau}{T_s}\cos\omega_s t + \frac{8\tau}{T_s}\cos 2\omega_s t)$

$= \frac{24\tau}{T_s}\cos 2\pi\,(f_m)t + \frac{16\tau}{T_s}\cos 2\pi\,(6f_m)t + \frac{48\tau}{T_s}\cos 2\pi\, f_s t\cos 2\pi\, f_m t$
$+ \frac{32\tau}{T_s}\cos 2\pi\, f_s t\cos 2\pi\,(6f_m)t + \frac{48\tau}{T_s}\cos 2\pi\,\,(2f_s)t\cos 2\pi f_m t\,\, +$
$\frac{32\tau}{T_s}\cos 2\pi\,(2f_s)t\cos 2\pi\,(6f_m)t + \cdots$

Recalled that $\cos\text{A}\cos\text{B} = \frac{1}{2}[\cos(\text{A}+\text{B}) + \cos(\text{A}-\text{B})]$

$\boldsymbol{v_s(t) = \frac{24\tau}{T_s}\cos 2\pi f_m t + \frac{16\tau}{T_s}\cos 2\pi\,(6f_m)t + \frac{24\tau}{T_s}\cos 2\pi\,(f_s - f_m)t +}$
$\boldsymbol{\frac{24\tau}{T_s}\cos 2\pi\,(f_s + f_m)t + \frac{16\tau}{T_s}\cos 2\pi\,\,(f_s - 6f_m)t + \frac{16\tau}{T_s}\cos 2\pi\,(f_s + 6f_m)t +}$

$$\frac{24\tau}{T_s}\cos 2\pi\,(2f_s - f_m)t + \frac{24\tau}{T_s}\cos 2\pi\,(2f_s + f_m)t + \frac{16\tau}{T_s}\cos 2\pi\,(2f_s - 6f_m)t + \frac{16\tau}{T_s}\cos 2\pi\,(2f_s + 6f_m)t + \cdots \ldots\ldots$$

From $f_m = 18\text{ kHz}$ & $f_s = 18\text{ MHz}$

f_m	0.018 MHz	$f_s + 6f_m$	18.108 MHz
$6f_m$	0.108 MHz	$2f_s - 6f_m$	35.892 MHz
$f_s - 6f_m$	17.892 MHz	$2f_s + 6f_m$	35.982 MHz
$f_s - f_m$	17.982 MHz	$2f_s + f_m$	36.018 MHz
$f_s + f_m$	18.018 MHz	$2f_s + 6f_m$	36.162 MHz

The spectrum of Vs(t) will be as shown in Fig. 7.32 below

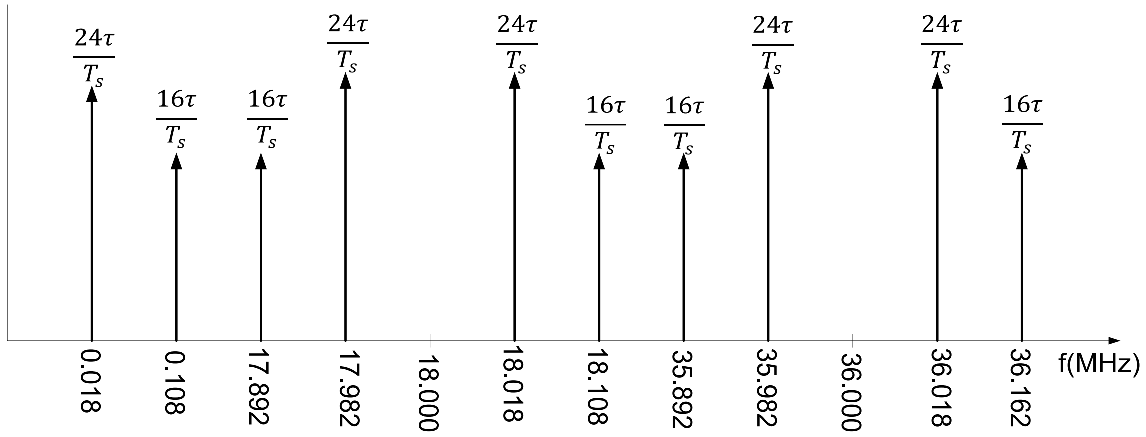


Figure 7.32

The signal $v_m(t)$ can be recover from $v_s(t)$ by passing $v_s(t)$ through a low pass filter with cut-off frequency greater than 108 kHz but less than 17 MHz

The problem envisage is that if the low-pass filter's cut-off frequency is not greater than 108 kHz and less than 17 MHz, or if the sampling frequency is not twice the maximum frequency of the base band signal $v_m(t)$, then the signal $v_m(t)$ will not be recovered because there will be an overlap known as aliasing for f_s = 200 kHz.

f_m	18 kHz
$6f_m$	108 kHz
$f_s - f_m$	182 kHz
$f_s + f_m$	218 kHz
$f_s + 6f_m$	308 kHz
$f_s - 6f_m$	92 kHz
$2f_s - f_m$	382 kHz
$2f_s + f_m$	418 kHz
$2f_s - 6f_m$	292 kHz
$2f_s + 6f_m$	508 kHz

Then there will be an effect of over lapping just as shown in the Fig. 7.33 below

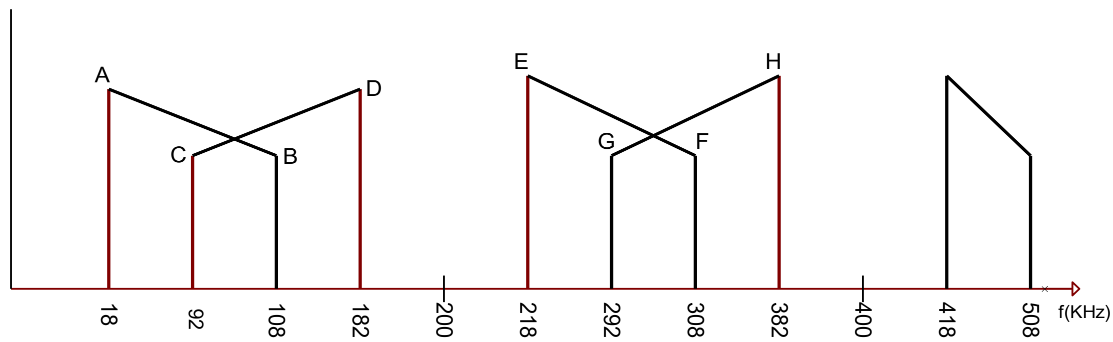


Figure 7.33

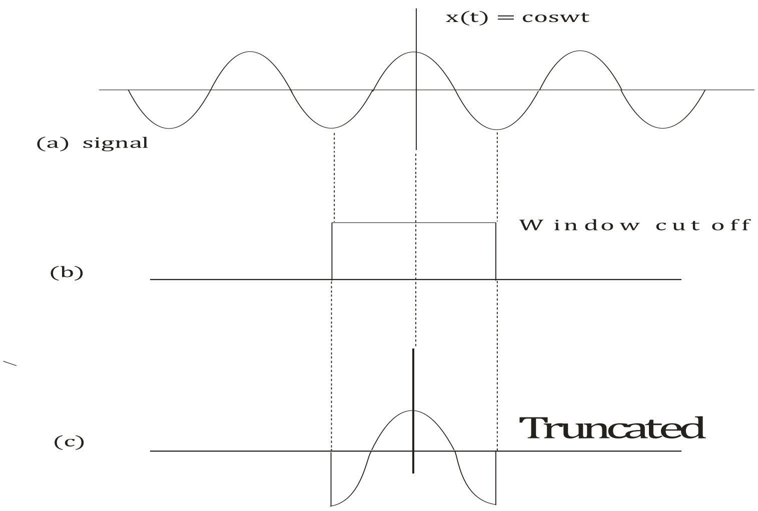
At points ABCD in Fig. 7.33 above, it can be seen that overlap occurred, i.e. AB overlap with CD and just as EF is with GH etc. Therefore, this is due to the fact that the Nyquist criterion was not met. Hence, the signal will not be recoverable due to aliasing.

7.12 Signal Truncation and Windowing

Truncated signals are signals which are zero outside certain time limits. Truncation occurs whenever a signal is switched on or off or is observed or recorded over a finite time interval. Truncation produces sudden discontinuities which lead to a spreading of spectral energy or broadening of the frequency spectrum.

Truncation at times are accidental and intentional. In pulse radars, a train of very short radio frequency pulses are transmitted into space, this is replaced by a target (air cut off for example) as echoes. The time taken for the echoes to return to the receiver gives the range of the target. If we record a signal during restricted time interval, assuming it to be zero elsewhere; truncation is the accidental in this case.

Truncation is achieved by multiplying a continuous signal by rectangular pulse otherwise known as observation window. The resultant effect of truncation is spreading or broadening of the frequency spectrum as stated earlier.



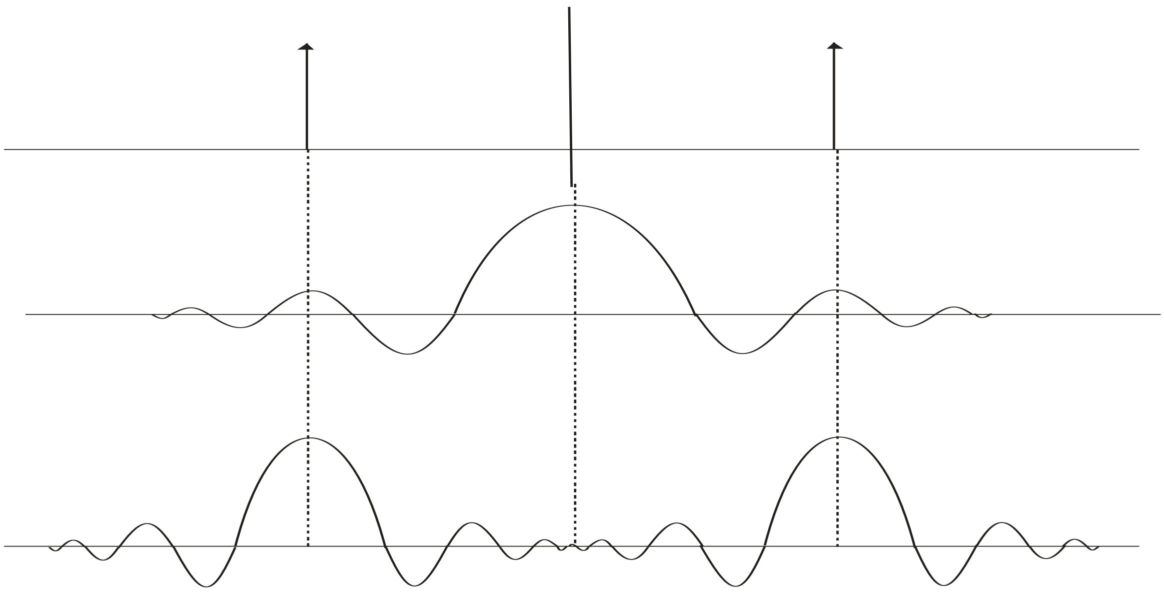


Figure 7.34

To reduce the spectral spreading, increasing the observations time (pulse duration) is not practicable. Other form of windows may be used to reduce spectral spreading. The signal process of truncating signal by multiplying it with an observation window is referred to as windowing. There are three conventional types of windowing

1. Rectangular Window
2. Triangular or Bartlett Window
3. Hanning Window

Rectangular Window

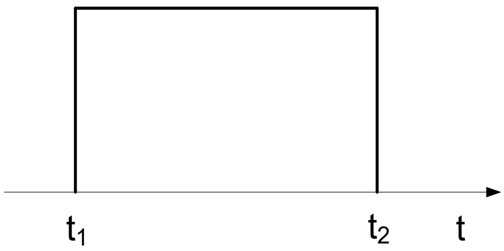


Figure 7.35

$$w(t) = \begin{cases} 1 & t_1 \le t \le t_2 \\ 0 & otherwise \end{cases} \qquad 7.79$$

Triangular or Bartlett Window

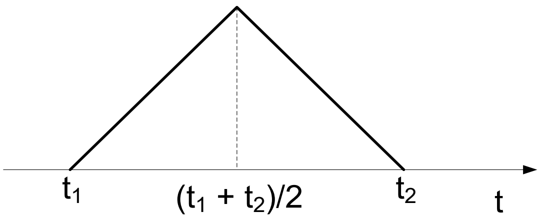


Figure 7.36

$$w(t) = \begin{cases} \frac{2}{t_2 - t_1}(t - t_1)\frac{t_1 + t_2}{2} \le t \le t_2 \\ \frac{2}{t_2 - t_1}(t - t_1)t_1 \le t \le \frac{t + t_2}{2} \\ 0 \qquad otherwise \end{cases} \qquad 7.80$$

Hanning Window

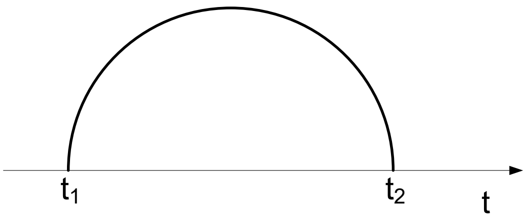


Figure 7.37

$$\omega(t) = \begin{cases} (1 + \cos \omega t) & t_1 \le t \le t_2 \\ 0 & otherwise \end{cases} \qquad 7.81$$

7.12.1 The Rectangular Window

When it is used, the spectral spreading is more pronounced. When triangular is used, the major 'ON' and 'OFF' discontinuities at the start and end of the waveform are avoided. The spectral spreading is less severe.

7.12.2 Hanning Window

This is half of the period of a raised cosine wave; its smooth transition further reduces spectral spreading. The signal being truncated and the window may be discreet in nature. This type of windowing techniques are often used in signal processing as well as in the design of discreet system, e.g. digital filters.

7.13 Mathematical Analysis of Observation Window

Let $x_{(t)}$ be a rectangular window and $y_{(t)}$ multiplied by $x_{(t)}$ such that we have $z_{(t)}$. If $y_{(t)}$ is a sinusoidal function, we can sketch the waveform of $x_{(t)}$ and obtain its Fourier transform from first principle and also sketch the Fourier transform of $x_{(t)}$. Also the Fourier transform of $y_{(t)}$ can equally be sketch and Fourier transform of $z_{(t)}$ can equally be obtained by convolution as well as its sketch.

Let $x_{(t)} = P_{\frac{\tau}{2}},\ y_{(t)} = A_m \cos \omega_{mt}$

The Fourier transform of $x_{(t)}$ is

$$F_{(\omega)} = \int_{\infty}^{\infty} x_{(t)} e^{-j\omega t} dt \qquad 7.82$$

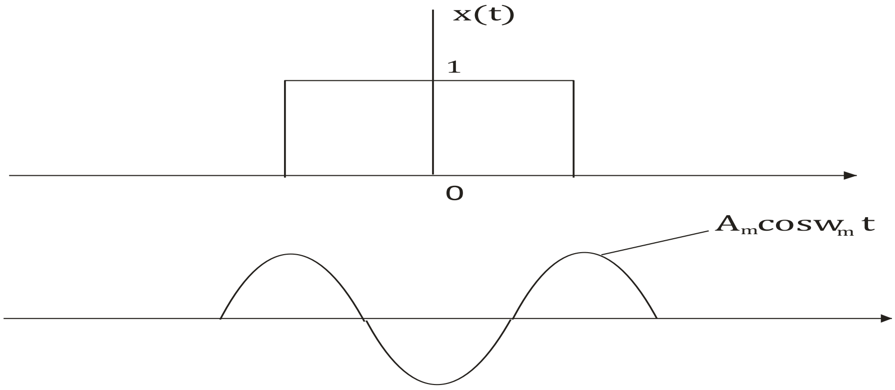


Figure 7.38

But the curve form of $x_{(t)}$ is

While the signal $y_{(t)}$ is

The mirror image of $y_{(t)}$ when view from the rectangular signal $x_{(t)}$ can be demonstrated as shown below

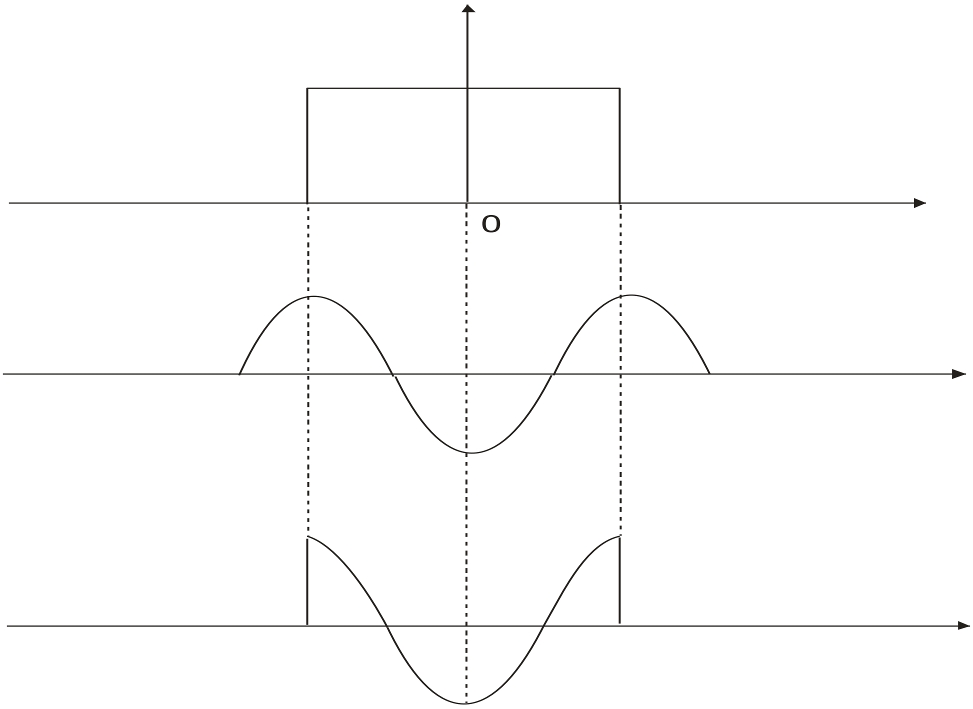


Figure 7.39

Now to obtain the Fourier transform of $x_{(t)}$, from Eq (7.39)

$$F(\omega) = \int_{-\infty}^{\infty} x_{(t)} e^{-j\omega t} dt$$

$$= \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} e^{-j\omega t} dt \quad 7.83$$

$$F(\omega) = \left[\frac{e^{-j\omega\tau}}{-j\omega}\right]_{-\frac{\tau}{2}}^{\frac{\tau}{2}} = \frac{e^{\frac{-j\omega\tau}{2}}}{-j\omega} - \left(\frac{e^{\frac{j\omega\tau}{2}}}{-j\omega}\right)$$

$$= \frac{e^{\frac{j\omega\tau}{2}}}{j\omega} - \frac{e^{\frac{-j\omega\tau}{2}}}{j\omega} \quad 7.84$$

$$= \frac{2}{\omega}\left[\frac{e^{\frac{j\omega\tau}{2}}}{2j} - \frac{e^{\frac{-j\omega\tau}{2}}}{2j}\right] \text{ (multiply Eq 7.84 numerator and denominator by 2)}$$

$$= \frac{2}{\omega}\sin\frac{\omega\tau}{2} \quad 7.85$$

$$= \frac{\tau\sin(\frac{\omega\tau}{2})}{(\frac{\omega\tau}{2})} \text{ (multiplying eq 7.85 numerator and denominator by } \frac{\tau}{2}\text{)}$$

$$= \tau \operatorname{sinc}\left(\frac{\omega\tau}{2}\right) \quad 7.86$$

Fourier sketch of x(t) and the frequency spectrum of y(t) is as shown below

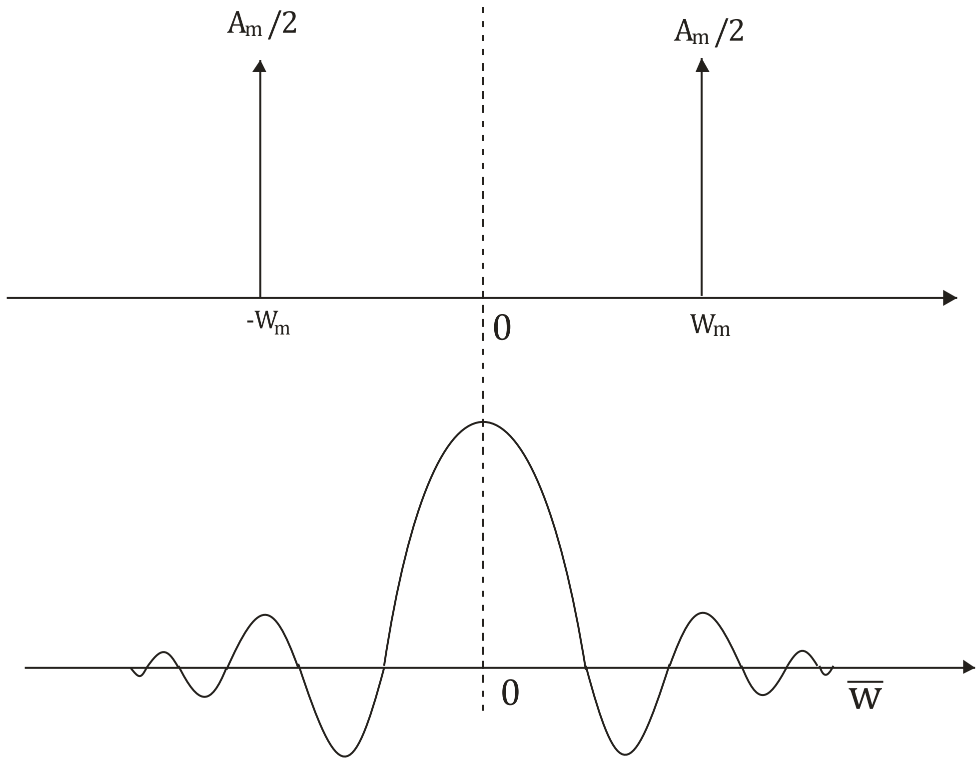


Figure 7.40

To convolute the signal above, we will have to redraw the spectrum such that we have the Fig. 7.41 and then will be an overlap at
$\overline{\omega} = \omega - \omega_m$, And at $\overline{\omega} = \omega + \omega_m$ as shown also

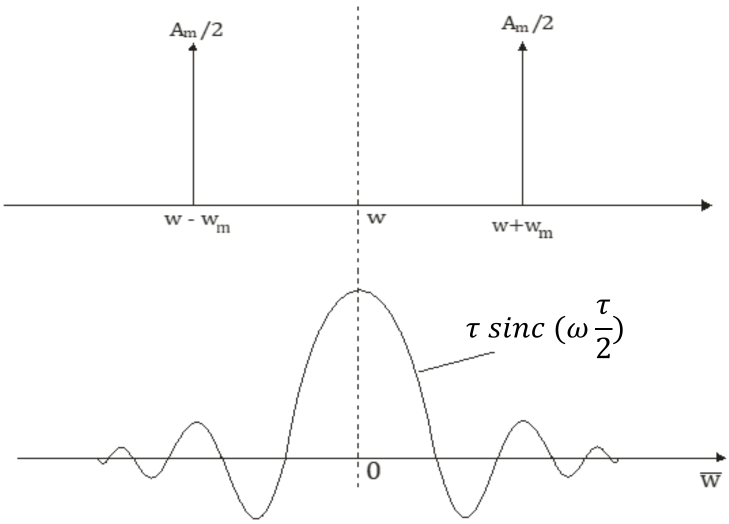


Figure 7.41

Such that $z(t) = \mathrm{x(t)} * \mathrm{y(t)}$

$$z(t) = \frac{\mathrm{A_m}}{2}\tau\mathrm{sinc}[(\omega - \omega_m)\frac{\tau}{2}] + \frac{\mathrm{A_m}}{2}\tau\mathrm{sinc}[(\omega + \omega_m)\frac{\tau}{2}]$$

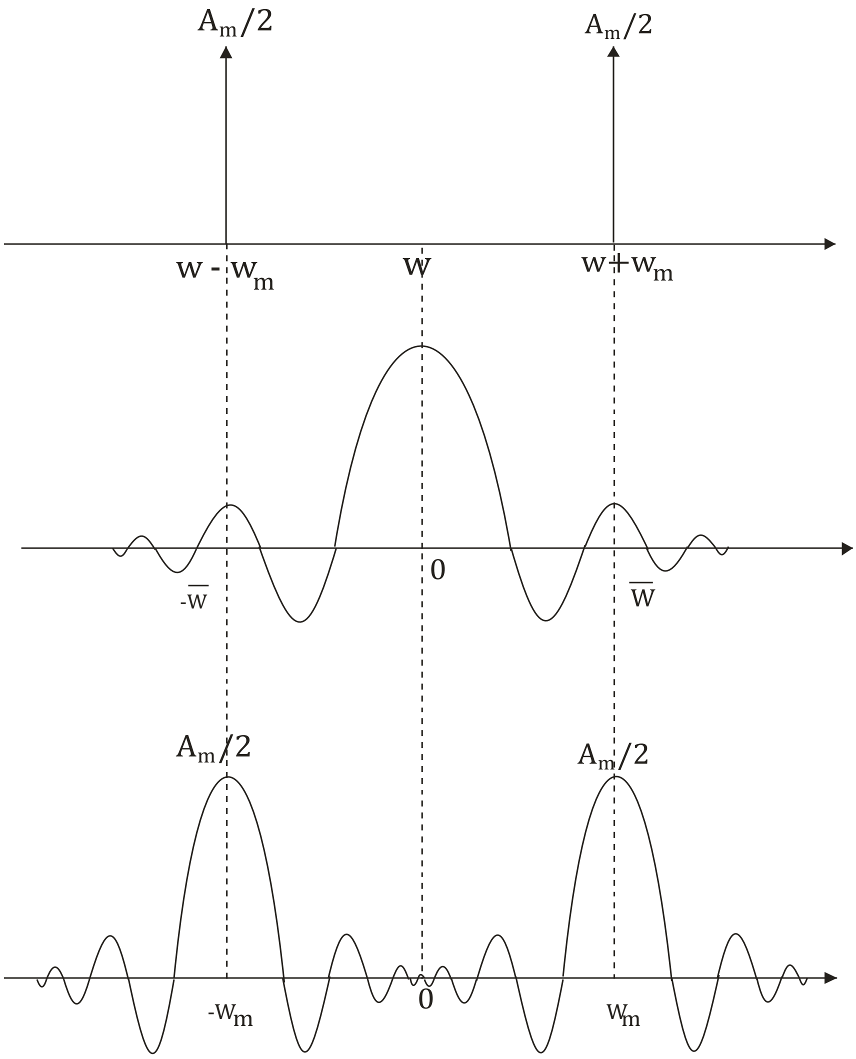


Figure 7.42

Example 7.9

$x_{(t)}$ is a rectangular window given as $x_{(t)} = ¼P_2(t)$, $y_{(t)} = 4\cos(\omega_m t)$ is multiplied with $x_{(t)}$ to give the output $z_{(t)}$.

i. Sketch the waveform of $x_{(t)}$.

ii. Obtain the Fourier transform of $x_{(t)}$ from first principle and sketch its transform.

iii. Sketch the Fourier transform of $y_{(t)}$.

iv. Obtain the Fourier transform of $z_{(t)}$ using convolution and sketch its transform.

Solution

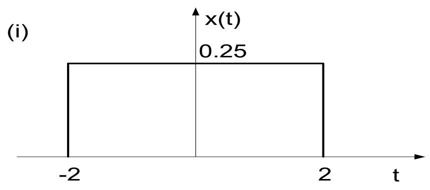


Figure 7.43

ii. $F(\omega) = \int_{-\infty}^{\infty} x_{(t)} e^{-j\omega t} dt$

$$F(\omega) = \int_{-2}^{2} 0.25\, e^{-j\omega t} dt = 0.25 \left[\frac{e^{-j\omega t}}{-j\omega}\right]_{-2}^{2}$$

$$F(\omega) = 0.25 \left[\frac{e^{-j2\omega}}{-j\omega} + \frac{e^{j2\omega}}{j\omega}\right] = \frac{0.25}{\omega}\left[\frac{e^{j2\omega}}{j} - \frac{e^{-j2\omega}}{j}\right]$$

$$F(\omega) = \frac{0.5}{\omega}\left[\frac{e^{j2\omega}}{2j} - \frac{e^{-j2\omega}}{2j}\right] = \frac{1}{2\omega}\sin 2\omega = \frac{1}{2}\left(\frac{\sin 2\omega}{\omega}\right) = \frac{1}{2}.2\left(\frac{\sin 2\omega}{2\omega}\right)$$

$$F(\omega) = \text{sinc}\,(2\omega)$$

The Fourier transform sketching is as shown in Fig. 7.44

iii.

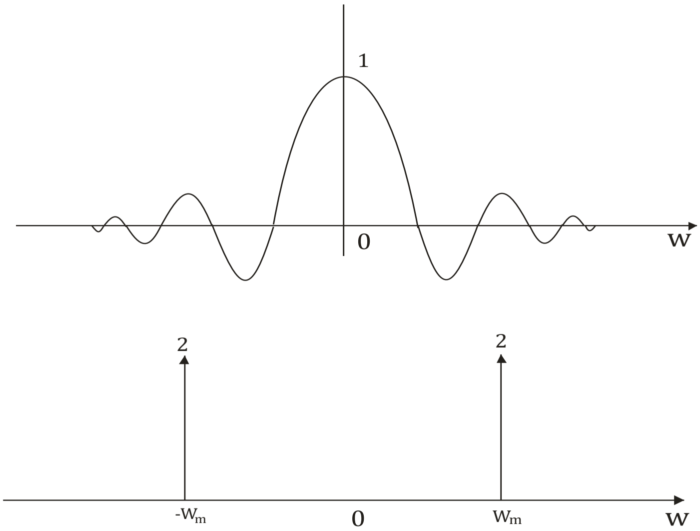


Figure 7.44

To convolute, let $\overline{\omega} = \omega + \omega_m$

$\overline{\omega} = \omega - \omega_m$

$z(t) = \text{x(t)} * \text{y}(t)$

$z(t) = 2\text{sinc } 2(\omega - \omega_m) + 2\text{sinc } 2(\omega + \omega_m)$

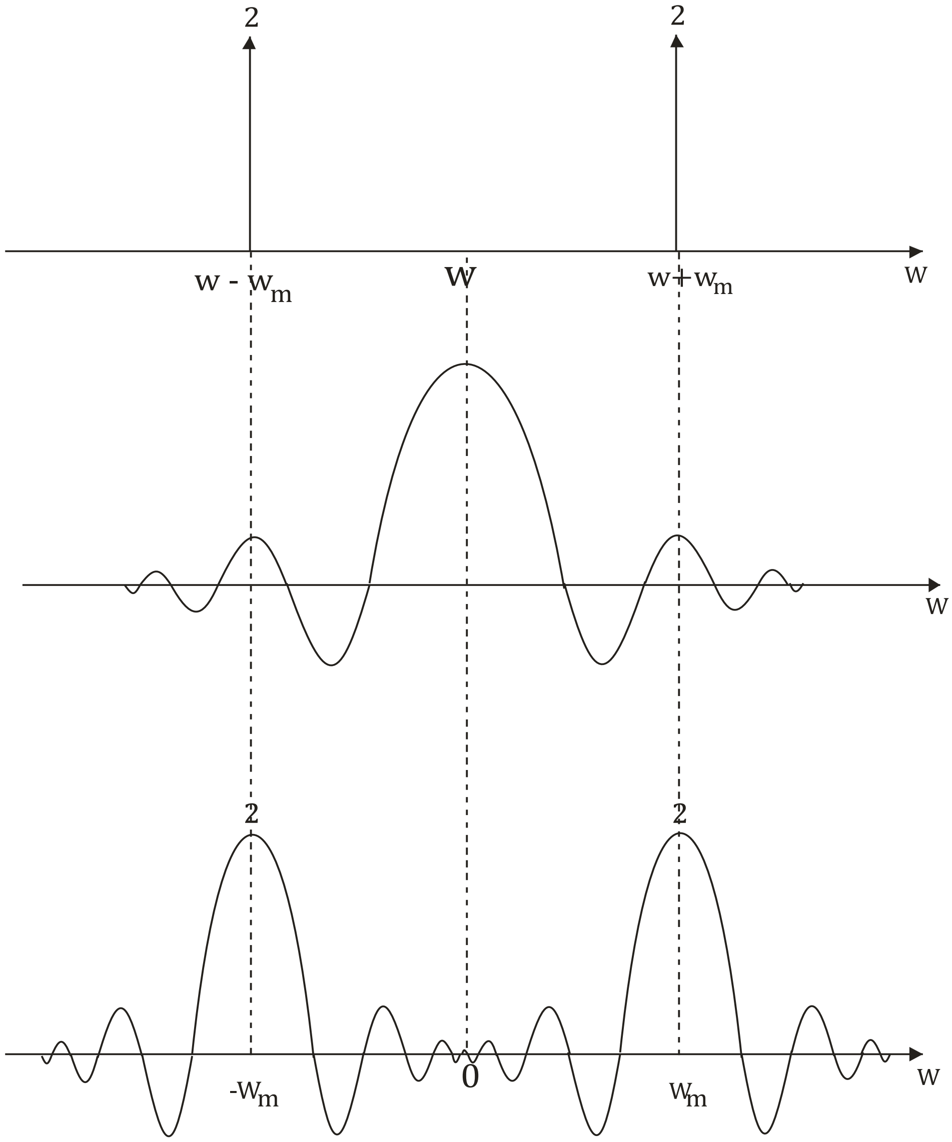


Figure 7.45

Miscellaneous solved examples

Example 7.10

A signal x(t) = 2cos 400πt + 6cos 40πt is ideally sampled at f_s = 500 Hz. if the sampled signal is passed through an ideal LPF with a cut-off frequency of 400 Hz. what frequency components will appear in the output.

Solution: A signal x(t) is given as

$$x(t) = A_1 \cos 2\pi f_1 t + A_2 \cos 2\pi f_2 t$$

So here $x\pi f_1 = 400\pi \Rightarrow f_1 = 200$ Hz

$\alpha\pi f_2 = 600\pi \Longrightarrow 320 \text{ Hz} = f_m$

Thus the maximum allowable frequency is from 0 to 320 Hz

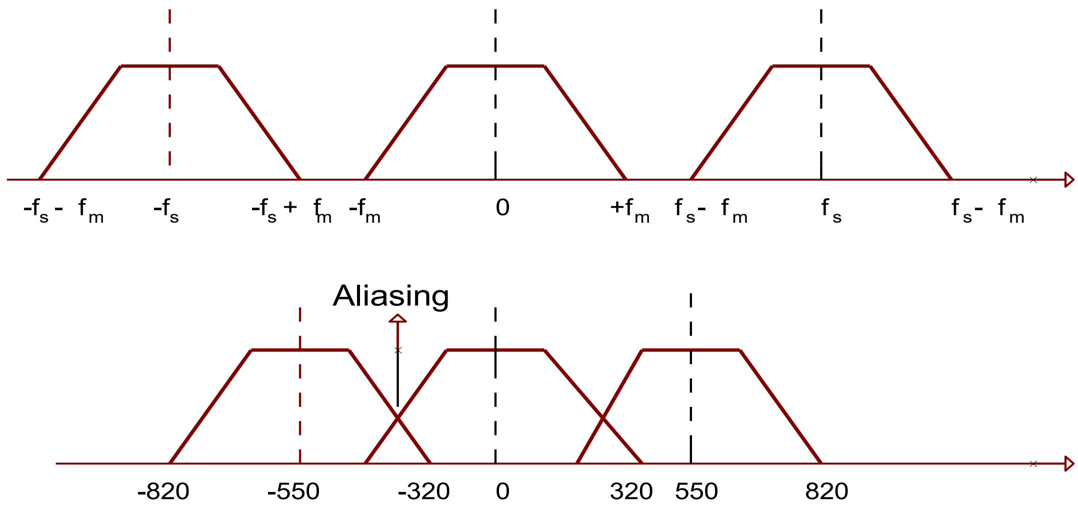


Figure 7.46

This signal is sampled at a sampling frequency of f_s = 500 Hz. but according to sampling theorem it should be sampled at

$f_s \leq 2f_m \therefore 2 \times 320 \rightarrow f_s \geq 640$ Hz

Thus the spectrum in general is shown as:

as there cut-off frequency of LPF is at 400 Hz all the sampled frequencies spear but as there is an aliasing error, the reconstruction is not possible.

Eampled 7.11

The signal g(t) =10 cos 20πt cos 200πt is sampled at the rate of 250 samples per seconds

i. Determine the spectrum of the resulting signal
ii. Specify the cut-off frequency of the ideal reconstruction filter so as to recover g(t) from its sampled version.
iii. What is the Nyquist rate for g(t)?

Solution: g(t) = 10 cos 20πt cost 20πt

= 5(2 cos 20πt cos 200πt)

Using 2cos A cos B = cos (A +B) + cos (A-B)

→ g(t) t = [cos (20πt+ 200πt) + cos (20πt– 200πt)]

= 5[cos (220πt) + cos (180πt)]

→ f_1 = 90 Hz f_2 = 110 Hz

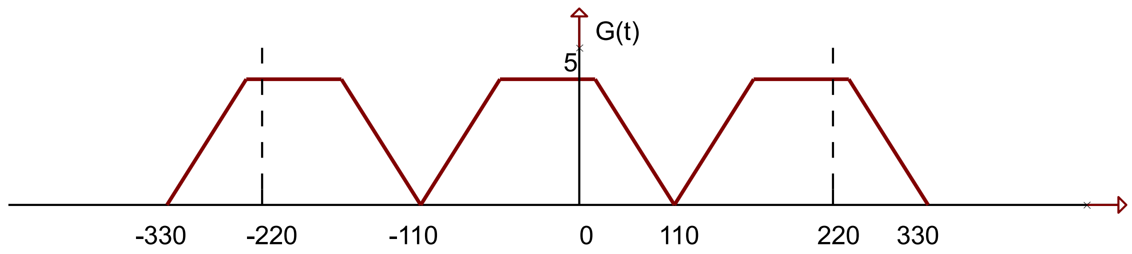


Figure 7.47

Max frequency component $f_2 = f_{max}$ = 110 Hz
Thus f_s should be ≥ 220 Hz. thus the spectrum becomes, the cut-off frequency of the ideal reconstruction filter should be more than 220 Hz

Example 7.12

The signal g(t) = 10 cos 60 πt $\cos^2$ 160 πt is sampled at the rate of 400 samples per seconds. Determine the range of permissible cut-off frequencies for the ideal reconstruction filter that may be used to recover g(t) from its sampled version.

Solution: $g(t) = 10 \cos 60\pi t \cos^2 160\pi t$

$$= 10 \cos 60\pi t \, \frac{(1 + \cos 320\pi t)}{2}$$

$g(t) = 5 \cos 60\pi t + 2.5 [\cos 400\pi t + \cos 260\pi t)$

f_1 = 30 Hz, f_2 = 200 Hz

f_3 = 130 Hz

f_m = 200 Hz

$f_s \geq 2 \times 200 \rightarrow f_s \geq 400$ Hz

7.14 Chapter Review Problems

7.1 Define and explain the analog modulation system.

7.2 What is pulse modulation? Explain its advantages over CW modulation. Discuss the application of pulse modulation.

7.3 Enumerate the types of pulse modulation. Describe PDM system in detail.

7.4 State and explain sampling theorem in time domain.
Prove that if a signal whose highest frequency is W Hz has been sampled at rate of 2 W samples per second, the impulse train through an ideal is low pass filter whose cut off frequency W Hz.
Discuss a typical communication system using pulse amplitude modulation with special reference to its bandwidth and signal two noise ratio requirements.

7.5 Explain the similarity between PMM and PDM. Discuss the SNR characteristic of PAM and then compare it PDM.

Write short notes on :

i. Advantages of pulse modulation.
ii. PPM and PDM
iii. PAM.
iv. SNR characteristic of PAM

7.6 Discuss clearly the principle of PAM/AM transmission and reception

$(AMIE\ W\ S - 1981)$

7.7 Explain with the diagrams, how low PPM signals are generated and the modulated signal is recovered from PPM waveform. Estimate the channel bandwidth requirement.

$(AMIE\ W\ 1980)$

7.8 Write short notes on the following

i. Detection of PDM signals.
ii. Pulse time modulation comparison of pulse modulation systems
iii. SNR improvements in PPM. $(AMIEW\ 1980)$
iv. Demodulation of PPM signals. $(Grad.ITETE$ Dec.1983)
v. Slicer.

7.9 Explain with diagrams, how PDM signals are generated and the modulated signal is recovered from PDM wave. Show the spectrum of $(Rookee\ University, 1984.85)$.

7.10 Explain with suitable circuit diagram the generation of PPM signal and explain how these signals are demodulated.

7.11 Explain with suitable diagrams how a PPM signal can be converted to PAM signals. Draw the waveform of the various stages. $(Banaras\ University$ 1982)

7.12 Carefully plot the spectrum of flat-top sampled PAM which has a 1 kHz sine modulating signal, a sampling frequency of 8 kHz and a pulse width of 3125 µs, up to 6th harmonies of the sampling frequency. Draw it with a dotted line the amplitude curve.

$(Communication\ fugg, Rookee\ University\ 1981 - 82)$

7.13

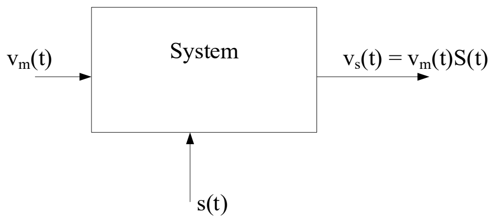
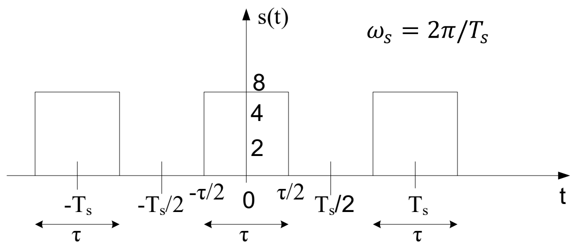


Fig Q1 (a)



(b)

Fig Q1(a) above shows a system. The signal s(t) is a train of pulses as in Fig Q1 (b). f_s is the frequency of s(t). The other signal is $v_m(t) = 6\cos 2\pi f_m t + 4\cos 12\pi f_m t$. If f_m = 18 kHz,

i. What form of signal processing takes place in the system of Fig (a).
ii. Obtain the Fourier series of s(t) from first principle.

iii. Suppose τ<<T_s, simplify the expression of s(t) obtained in (ii) above and sketch its frequency spectrum.

iv. Obtain an expression for the signal $v_s(t)$, using only the first two terms in (iii).

v. If f_s = 18MHz, sketch the frequency spectrum of $v_s(t)$. Discuss the recovery of $v_m(t)$ from $v_s(t)$. State problem(s) you envisaged (if any)

vi. Repeat (v) with f_s = 200 kHz.

(University of Ibadan, TEL525- Signal Processing 2002/2003 BSc degree Examination)

7.14

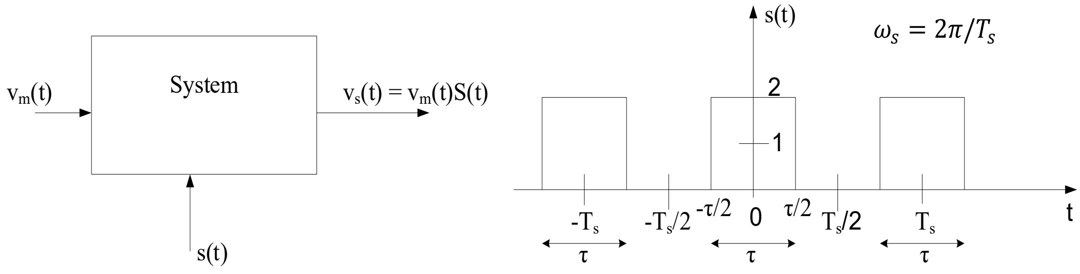


Fig Q2 (a) (b)

(a) (i) What is sampling?

(ii) State the sampling theorem.

(b) Fig Q2 (a) shows a sampler. The sampling signal s(t) is a train of pulses as shown in Fig Q2 (b). f_s is the frequency of s(t). The baseband signal is

$$v_m(t) = \cos 2\pi f_m t + \cos 10\pi f_m t \quad and \quad f_m = 50 \text{ kHz}.$$

(i) Obtain the Fourier series of s(t) from first principle.

(ii) Suppose $\pi << T_{s,}$ simplify the expression of s(t) obtained in (i) above . Sketch the frequency spectrum of s(t).

(iii) Obtained an expression for the sample signal $v_s(t)$ using only the first two terms in (ii).

(iv) If $f_s = 800$ kHz, sketch the frequency spectrum of $v_s(t)$. Discuss the recovery of the baseband signal from the sampled signal. State problem(s) you envisaged if any.

(v) Repeat (iv) with f_s= 400 kHz.

(University of Ibadan, TEL525- Signal Processing 2010/2011 BSc degree Examination)

7.15

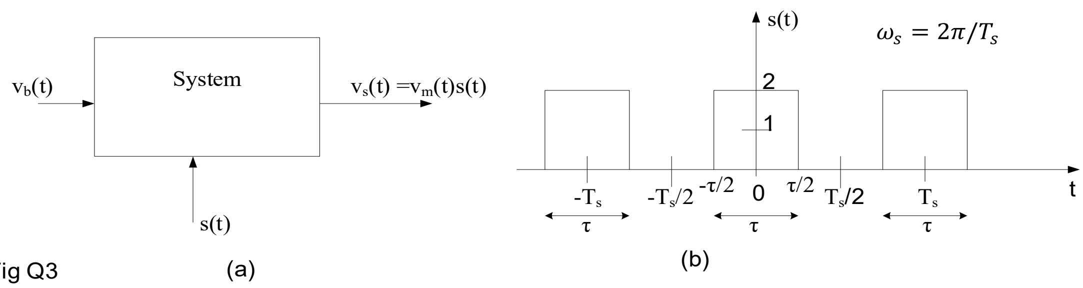


Fig Q3 (a) (b)

(a) (i) Explain briefly what you understand by "Sampling".

(ii) State the sampling theorem.

(b) Fig. Q3 (a) shows a sampler. The sampling signal s(t) is a train of pulses as shown in Fig. Q3(b). The baseband signal is $v_b(t)$= 1 + cos$\omega_1 t$ + cos$4\omega_1 t$

(i) Obtain the Fourier series of s(t) from first principle.

(ii) Suppose $\tau << T_{s,}$ simplify the expression of s(t) obtained in (i) above. Sketch the frequency spectrum .

(iii) Obtained an expression for the sample signal $v_s(t)$ using only the first two terms in (ii).

(iv) If $\omega_s = 12\omega_1$, sketch the frequency spectrum of $v_s(t)$. Discuss the recovery of the baseband signal from the sampled signal. State problem(s) you envisaged if any.

(v) Repeat (iv) with ω_s= $6\omega_{1.}$.

(University of Ibadan, TEL525- Signal Processing 2012/2013 BSc degree Examination)

7.16

(a) (i) What is windowing?

(ii) What are the effect(s) of windowing and how can the effect(s) be reduced?

(b) x(t) is a rectangular window given as $x(t) = \frac{1}{4}P_2\ (t)$

$y(t) = 4cos(\omega_m t)$ is multiplied with $x(t)$ to give the output $z(t)$.

(i) Sketch the waveform of $x(t)$

(ii) Obtain the Fourier transform of $x(t)$ using convolution and sketch and sketch its transform.

(University of Ibadan, TEL525- Signal Processing 2012/2013 BSc degree Examination)

7.17

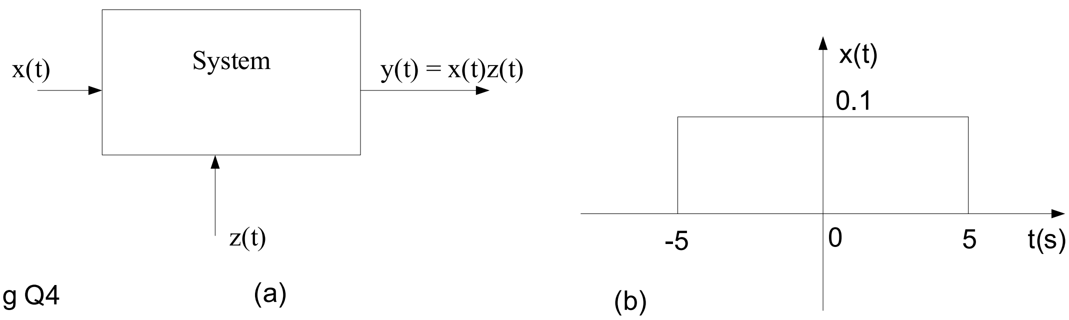


Fig Q4 (a) (b)

(a) The waveform of the signal z(t) of Figure Q4(a) is shown in Figure Q4(b). The signal x(t) is given as x(t) = 4cos ($\omega_m\ t$).

(i) Determine the Fourier Transport z(t)

(ii) Sketch the Fourier transform of z(t).

(iii) Hence, determine the Fourier Transform of y(t) using convolution.

(iv) Sketch the Fourier transform of y(t) and the waveform of x(t), z(t) and y(t).

(v) What form of signal processing takes place in the system of Figure Q4(a)?

(vi) Deduce the effect of that form of signal processing on x(t). How can this effect be reduced?

(University of Ibadan, TEL525- Signal Processing 2011/2012 BSc degree Examination)

7.18

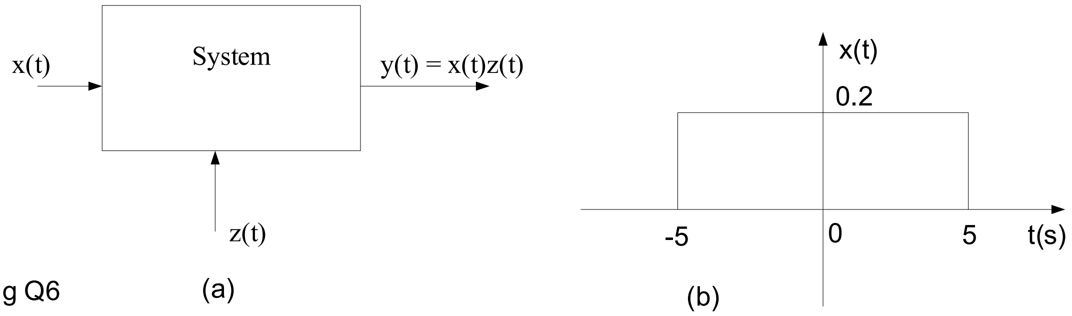


Fig Q6 (a) (b)

(a) The waveform of the signal z(t) of Figure Q1(a) is shown in Figure Q1(b). The signal x(t) is given as x(t) = 6cos $(\omega_m t)$.

(i) Determine the Fourier Transport z(t)

(ii) Sketch the Fourier transform of z(t) and z(t).

(iii) Hence, determine the Fourier Transform of y(t) using convolution.

(iv) Sketch the Fourier transform of y(t) and the waveform of x(t), z(t) and y(t).

(v) What form of signal processing takes place in the system of Figure Q1 (a)

(vi) Deduce the effect of that form of signal processing on x(t). How can this effect be reduced?

(University of Ibadan, TEL525- Signal Processing 2012/2013 BSc degree Examination)

7.19

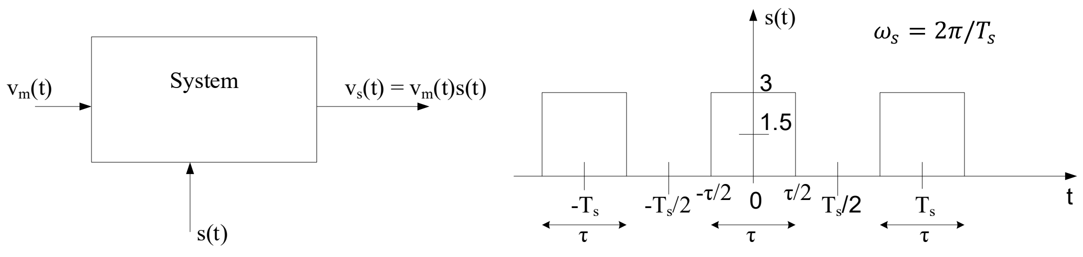


Fig Q7 (a) (b)

Fig Q7 (a) shows a system. The sampling signal s(t) is a train of pulses as shown in Fig Q7(b). f_s is the frequency of s(t). The baseband signal is $v_b(t)$=1+ cos$\omega_1 t$+ cos$4\omega_1 t$

(i) What form of signal processing takes place in the system of Figure Q7(a)

(ii) Obtain the Fourier series of s(t) from first principle.

(iii) Suppose $\tau << T_{s,}$ simplify the expression of s(t) obtained in (ii) above. Sketch the frequency spectrum of s(t)

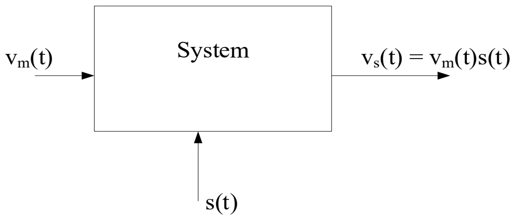
(iv) Obtained an expression for the sample signal $v_s(t)$ using only the first two terms in (iii).

(v) If $\omega_s = 12\omega_1$, sketch the frequency spectrum of $v_s(t)$. Discuss the recovery of $v_m(t)$ from $v_s(t)$. State problem(s) you envisaged (if any).

(v) Repeat (iv) with ω_s= $6\omega_1$.

(University of Ibadan, TEL525- Signal Processing 2009/2010 BSc degree Examination)

7.20



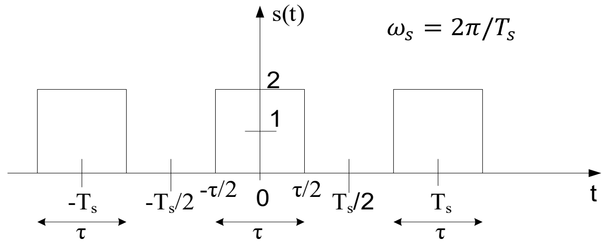


Fig Q8 (a) (b)

Fig Q8 (a) shows a sampler. The sampling signal s(t) is a train of pulses as shown in fig Q8(b).

The baseband signal is $v_m(t)$=2cos$2\pi f_m t$+ 3cos$14\pi f_m t$, $f_m = 42$ kHz.

(i) Obtain the Fourier series of s(t) from first principle.

(ii) Suppose $\tau << T_{s,}$ simplify the expression of s(t) obtained in (i) above . Sketch the frequency spectrum of s(t) .

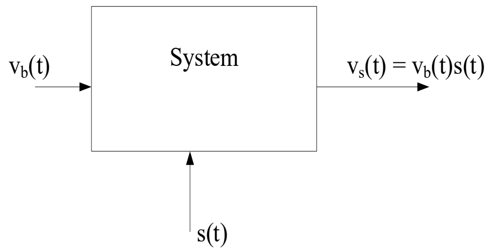
(iii) Obtained an expression for the sample signal $v_s(t)$ using only the first two terms in (ii).

(iv) If $f_s = 0.8$ MHz, sketch the frequency spectrum of $v_s(t)$. Discuss the recovery of the baseband signal from the sampled signal. State problem(s) encountered (if any).

(v) Repeat (iv) with f_s= 0.05 MHz.

(University of Ibadan, TEL525- Signal Processing 2008/2009 BSc degree Examination)

7.21



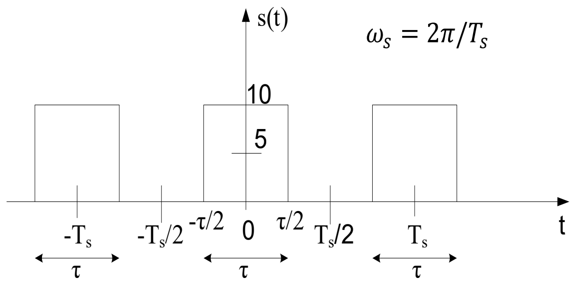


Fig Q9 (a) (b)

Fig Q (a) shows a sampler. The sampling signal s(t) is a train of pulses as shown in Fig Q(b). The baseband signal is v_b =1+ cos$\omega_1 t$+ cos$6\omega_1 t$.

(i) Obtain the Fourier series of s(t) from first principle.

(ii) Suppose $\tau << T_{s,}$ simplify the expression of s(t) obtained in (i) above . Sketch the frequency spectrum of s(t).

(iii) Obtained an expression for the sample signal $v_s(t)$ using only the first two terms in (ii).

(iv) If $\omega_s = 18\omega_1$, sketch the frequency spectrum of $v_s(t)$. Discuss the recovery of baseband signal from the signal. State problem(s) you envisaged (if any).

(v) Repeat (iv) with ω_s= $8\omega_1$.

(University of Ibadan, TEL525- Signal Processing 2007/2008 BSc degree Examination)

7.22 x(t) is a rectangular window given as $x(t)$= $\frac{1}{6}P_3$(t)

y(t) = $4\cos\,(\omega_m t)$ is multiplied with x(t) to give the output z(t).

(i) Sketch the waveform of x(t) and obtain its Fourier transform from first principle. Sketch the Fourier transform of x(t).

(ii) Sketch the waveform of y(t).

(iii) Obtain the Fourier transform of z(t) using convolution. Sketch the Fourier transform of z (t).

(University of Ibadan, TEL525- Signal Processing 2010/2011 BSc degree Examination)

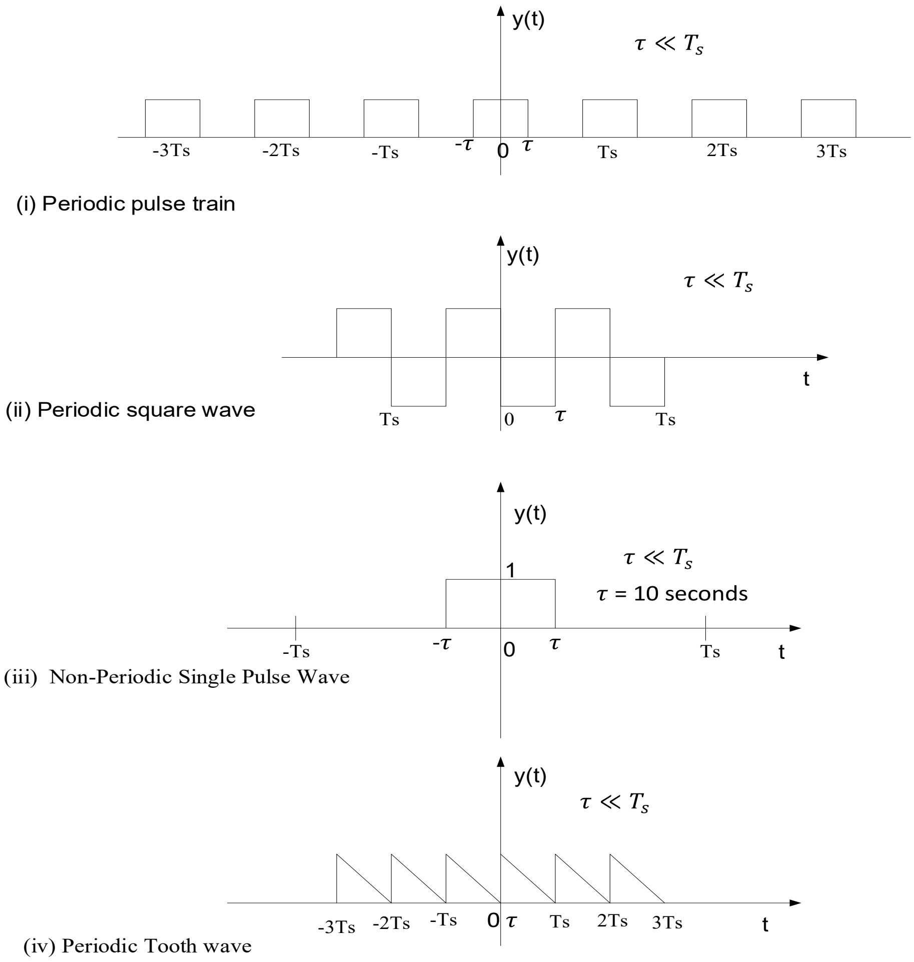


Fig. 2: Available Signals

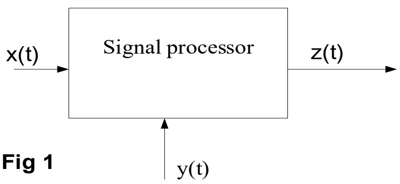


Fig 1

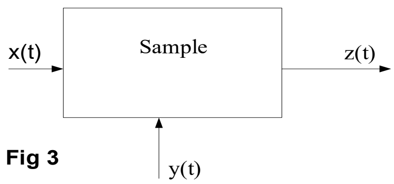


Fig 3

7.23. The sampler of Fig 3 is designed to sample the baseband signal x(t) which is given as

x(t) = 3 + 2cos $(2\pi x 10^3 t)$ + 7cos $(10\pi x 10^3 t)$ Select an appropriate signal from the list of available signals in Fig 2 as the second input y(t) to the sampler.

(i) Obtain the Fourier series of the selected signal y(t) from first principle.

(ii) Obtained an expression for the sample signal $z(t)$ using only the first two terms in (i) above.

(iii) Supposing T_s = 4x 10^{-5} s. Does this choice of T_s satisfy the sampling Theorem? Sketch the frequency spectrum of $z(t)$. Discuss the recovery of the baseband x(t) from the output z(t). State problem(s) encountered (if any).

(iv) Repeat (iii) with T_s= 6 x 10^{-5} s.

(University of Ibadan, TEL525- Signal Processing 2002/2003 B.Sc. degree Examination)

7.24 x(t) is a rectangular window given as $x(t)$= $\frac{1}{10}P_5$(t)

y(t) = 2 sin$\omega_m t$ is multiplied with x(t) to give the output z(t).

(i) Sketch the waveform of x(t) and obtain its Fourier transform from first principle. Sketch the Fourier transform of x(t).

(ii) Obtain the Fourier transform of z(t) using convolution. Sketch the Fourier transform of z (t).

(University of Ibadan, TEL525- Signal Processing 2004/2004 B.Sc. degree Examination)

7.25 x(t) is a rectangular window given as $x(t)$= $\frac{1}{20}P_{10}$(t).

y(t) = 2sin$\omega_m t$ is multiplied with x(t) to give the output y(t).

(i) Sketch the waveform of x(t) and obtain its Fourier transform from first principle. Sketch the Fourier transform.

(iii) Sketch the Fourier transform of x(t).

(ii) Obtain the Fourier transform of y(t) using convolution. Sketch the Fourier transform of y(t).

(University of Ibadan, TEL525- Signal Processing 2001/2002 BSc degree Examination)

7.26

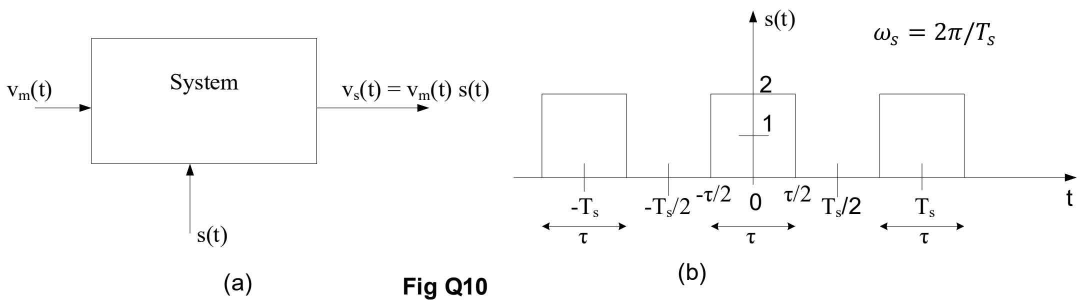


Fig Q10

Fig Q10(a) shows a sampler. The sampling signal s(t) is a train of pulses as shown in fig Q10(b).

The baseband signal is $v_m(t)$ = 4cos$\omega_1 t$ + cos3$\omega_1 t$.

(i) Obtain the Fourier series of s(t) from first principle.

(ii) Hence obtain the Fourier series of s(t) for $\tau << T_{s.}$ Sketch the frequency spectrum of s(t) .

(iii) Obtained an expression for the sampled signal $v_s(t)$ using only the first two terms in (ii).

(iv) If $\omega_s = 10\omega_1$, sketch the frequency spectrum of $v_s(t)$. Discuss the recovery of the baseband signal from the sampled signal. State problem(s) encountered (if any).

(v) Repeat (iv) with ω_s= $5\omega_1$.

(University of Ibadan, TEL525- Signal Processing 2000/2001 BSc degree Examination)

7.27

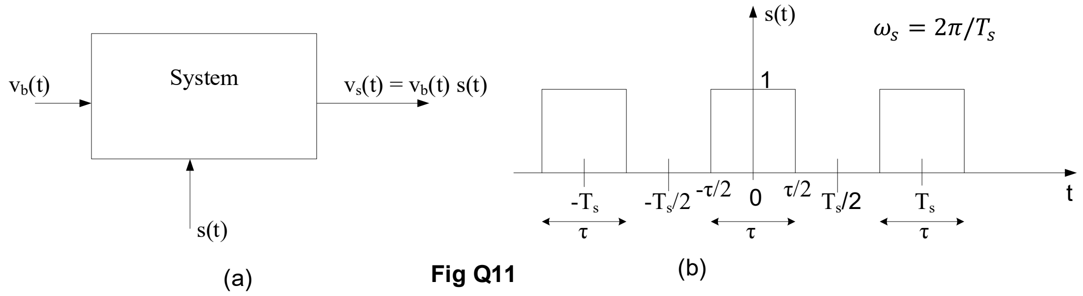


Fig Q11

Showing Fig Q11(a). The baseband signal is $v_b(t) = cos\omega_1 t + \cos 3\omega_1 t$

(i) Obtain the Fourier series of s(t) from first principle.

(ii) Suppose $\tau << T_{s.}$ simplify the expression of s(t) obtained in (i) above. Sketch the frequency spectrum of s(t).

(iii) Obtained an expression for the signal $v_s(t)$ using only the first two terms in (ii).

(iv) If $\omega_s = 12\omega_1$, sketch the frequency spectrum of $v_s(t)$. Discuss the recovery of the baseband signal from the sampled signal. State problem(s) encountered (if any).

(v) Repeat (iv) with $\omega_s = 4\omega_1$.

(University of Ibadan, TEL525- Signal Processing 2002/2003 BSc degree Examination)

7.28

x(t) is a rectangular window given as $x(t) = \frac{1}{2}P_t\ (t)$

y(t) $= 4cos(\omega_m t)$ is multiplied with x(t) to give the output z(t).

(i) Sketch the waveform of x(t) and obtain its Fourier transform from first principle. Sketch the Fourier transform from first principle. Sketch the Fourier transform of x(t)

(ii) Sketch the waveform of x(t).

(iii) Obtain the Fourier transform of z(t) using convolution. Sketch the Fourier transform of z(t).

(University of Ibadan, TEL525- Signal Processing 2007/2008 BSc degree Examination)

CHAPTER 8

DIGITAL MODULATION

8.0 Introduction

Apart from telegraphy, there was hardly any pulse-type communication until just before the World War II. Then the advent of television and radar quickly changed the picture. Now, a lot of telecommunication is in digital (pulse) from, and the proportion is digital communication, increasingly at the expense of analog communication, is caused by two inter working factors. The first is the fact that a lot of information to be transmitted is in digital from (i.e. 0's and 1's) to start with, and so sending it in that form is clearly the simplest technique. The second factor has been the advent of large-scale integration, which has permitted the use of complex coding systems that take the best advantage of channel capacities. Accordingly, it is very important to have a good working knowledge of the fundamentals of digital and data communication.

All the modulation types have been analogue representations of the message. Pulse-code modulation (usually abbreviated as PCM) is distinctly different in concept. It is digital modulation in which the message is represented by a coded group of digital (discontinuous-amplitude) pulse. Delta modulation (normally abbreviated as DM is variation of PCM).

In an analog-modulation discussed in the previous chapters, the modulation parameter varies continuously and can take on any value corresponding to the range of the message. When the modulated wave is mixed with noise, there is no way for the receiver to discern the exact transmitted value. Suppose, however, that only a few discrete (i.e. discontinuous) values are allowed for the modulated parameter and if the separation between these value is large compared to the noise disturbances, it will be a simpler matter to detect at the receiver precisely which specific values are intended. Thus, the effects of random noise can be virtually eliminated, which is the whole idea of digital modulation.

8.1 Analog Pulse Modulation Methods

We know that in analog modulation system, some parameter of a sinusoidal carrier is varied according to the instantaneous value of the modulating signal. In pulse modulation methods, the carrier is no longer a continuous signal but consists of a pulse train. Some parameter of which is varied according to the instantaneous value of the modulating signal. There are two types of pulse modulation systems as under:

a. Pulse Amplitude Modulation (PAM)
b. Pulse Time Modulation (PTM)

In pulse amplitude modulation (PAM), the amplitude of the pulses of the carrier pulse train is varied in accordance with the modulating signal where as in pulse time modulation (PTM), the timing of the pulse of the carrier pulse train is varied. There are two types of PTM as under:

a. Pulse Width Modulation (PWM)
b. Pulse Position Modulation (PPM)

In pulse width modulation, the width of the pulses of the carrier pulse train is varied in accordance with the modulating signal whereas in pulse position modulation (PPM), the position of the pulses of the carrier pulse train is varied. Figure 8.1 shows three types of pulse analog modulation methods.

According to the sampling theorem, if a modulating signal is band limited to F_m i.e. (the maximum frequency component in the signal is f_m), the sampling frequency must be at least $2f_m$ and hence the frequency of the carrier pulse train must also be at least $2f_m$. At this point, it may be noted that all the above pulse modulation methods (i.e. PAM, PWM and PPM) are called analog pulse modulation methods because the modulating signal is analog in nature in PAM, PWM and PPM.

Note: (i) Pulse Width Modulation is also known as Pulse Duration Modulation (PDM).
(ii) If a signal is said to be band limited to F_m, then it means that the maximum frequency component in this signal is F_m.

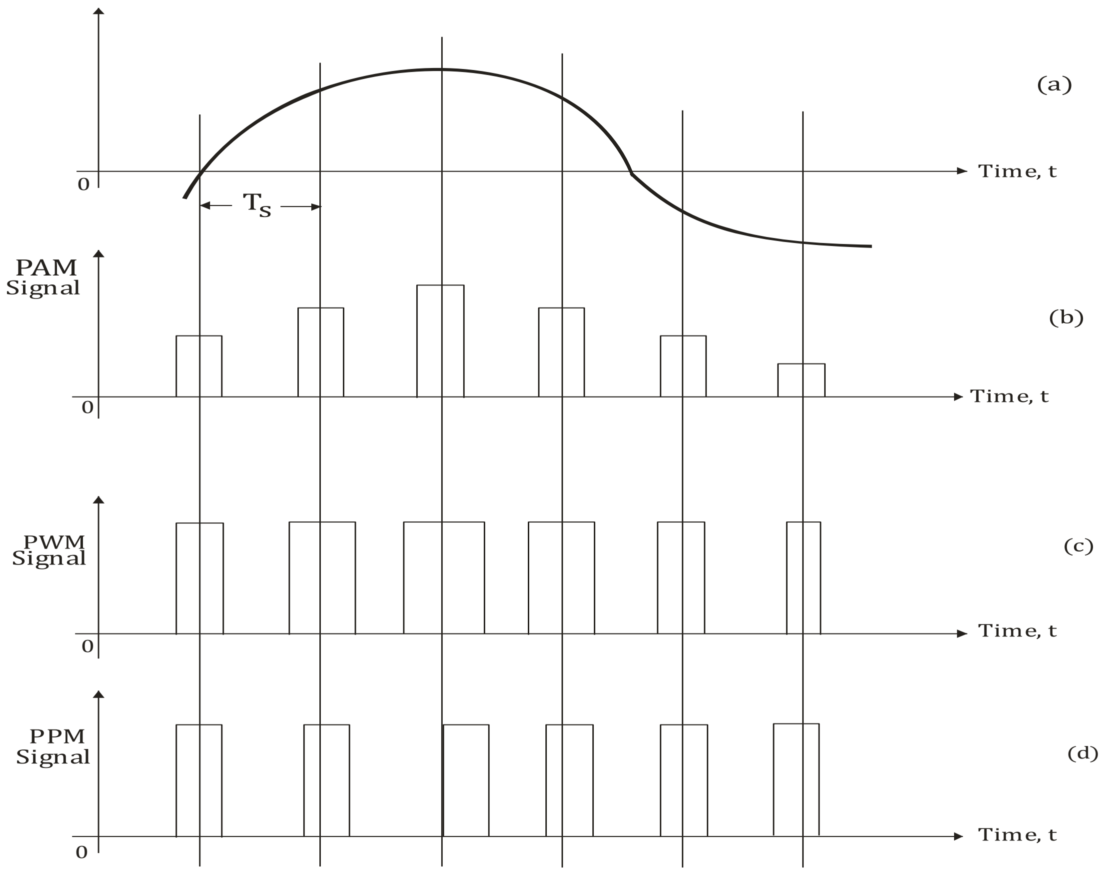


Figure 8.1 Pulse Amplitude Modulation (PAM)

Pulse amplitude modulation may be defined as that type of modulation in which the amplitudes of regularly spaced rectangular pulses vary according to instantaneous value of the modulating or message signal. In fact, the pulse in a PAM signal may be of flat top type or natural type or ideal type. Actually all the sampling methods which we shall discuss later are basically pulse amplitude modulation methods.

Out of these three pulse amplitude modulation methods, the Flat Top PAM is most popular and is widely used. The reason for using Flat Top PAM is that during the transmission, the noise interferes with the top of the transmitted pulses and this noise can be easily removed if the PAM pulse has flat top.

However, in case of natural samples PAM signal, the pulse has varying top in accordance with the signal variation. Now, when such type of pulse is received at the receiver, it is always contaminated by noise. Then it becomes quite difficult to determine the shape of the top of the pulse and thus amplitude detection of the pulse is not exact. Due to this, errors are introduced in the received signal.

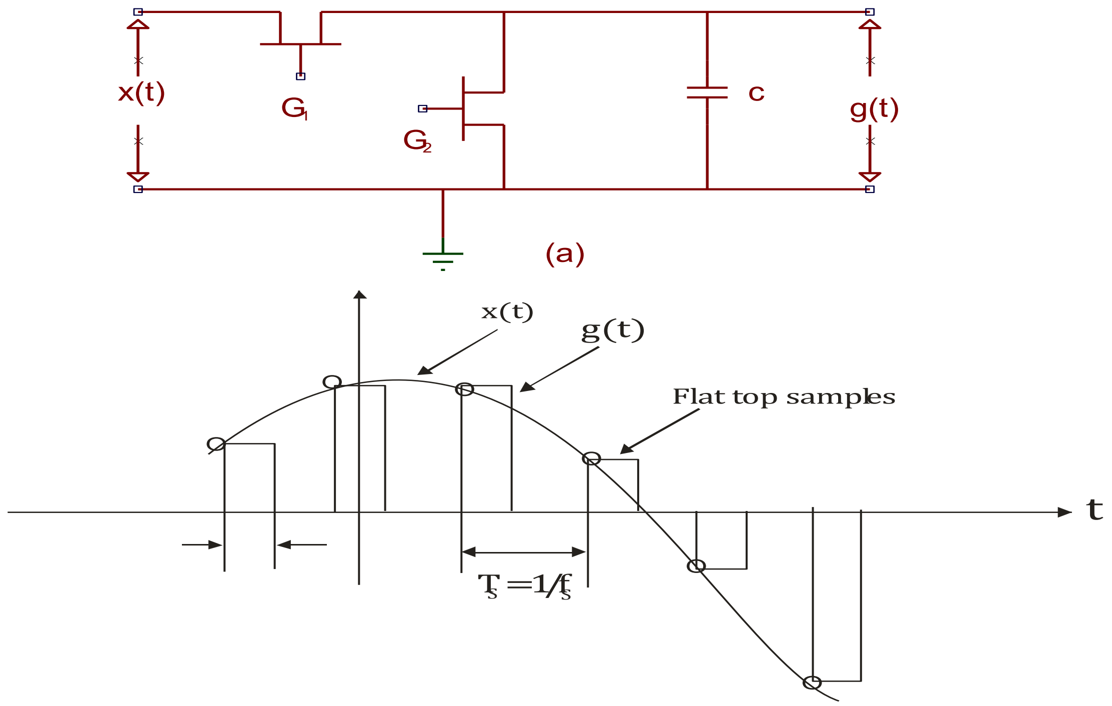


Figure 8.2 (a) Sample and hold circuit generating flat top sampled PAM (b) Waveforms of tap sampled PAM

Therefore, Flat Top Sample PAM is widely used.

Principle of Operation: A sample and hold circuit to produce flat top sampled PAM. The working principle of this circuit is quite easy. The sample and hold (S/H) circuit consist of two field effect transistors (FET) switches and a capacitor. The sampling switch is closed for a short duration by a short pulse applied to the gate G1 of the transistor. During this period, the capacitor 'C' is quickly charged up to a voltage equal to the instantaneous sample value of the incoming signal x(t). Now, the sampling switch is opened and the capacitor 'C' holds the charge. The discharge switch is then closed by a pulse applied to gate G2 of the other transistor. Due to this, the capacitor 'C' is discharge to zero volts. The discharges the switch is then opened and thus capacitor has no voltage. Hence, the output of the sample and hold circuit consists of a sequence of flat top samples as shown in Fig. 8.3

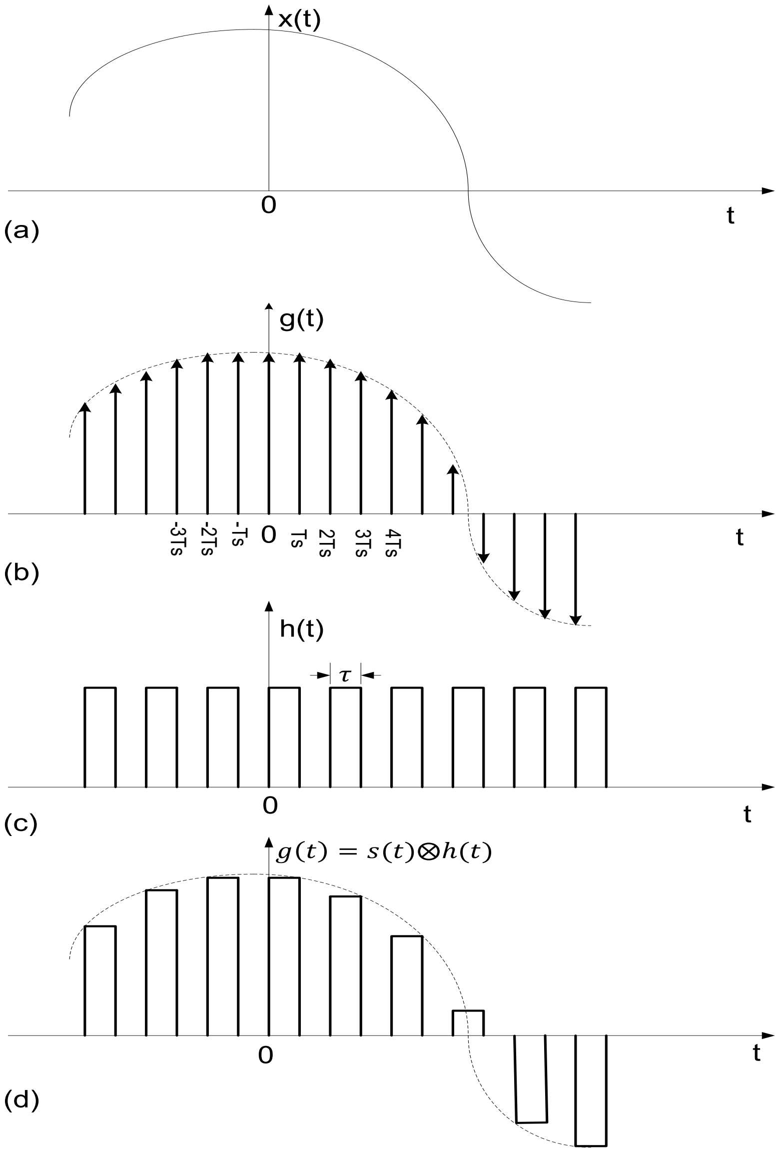


Figure 8.3: Baseband signal x(t) instantaneous sample single s(t) (c) Constant pulse width function h(t) (d) Flat top sampled PAM signal g(t) obtained through convolution of h(t) and s(t)

8.2 Transmission Bandwidth in Pulse Amplitude Modulation (PAM)

In pulse amplitude modulated (PAM) signal the pulse duration (τ) is considered to be very small in comparison to time period (i.e. sampling period) T_s between any two samples i.e.

$\tau << T_s$ 8.1

Now, if the maximum frequency in the modulating signal x(t) is F_m, then according to sampling theorem, the sample frequency F_s must be equal to or higher than the Nyguist rate i.e.

Or $f_s \geq 2f_m$

0r $\frac{1}{T_s} \geq 2f_m$ $(\because f_s = \frac{1}{T_s})$ 8.2

Or $T_s \leq \frac{1}{2f_m}$

But according to equation 8.1, we have $\tau << T_s$

Hence $\tau << T_s \leq \frac{1}{2f_m}$ 8.3

Now, if the 'ON' and 'OFF' time of the pulse amplitude modulated (PAM) pulse is same as show in Fig. 8.4 (a) then maximum frequency of the PAM pulse will be:

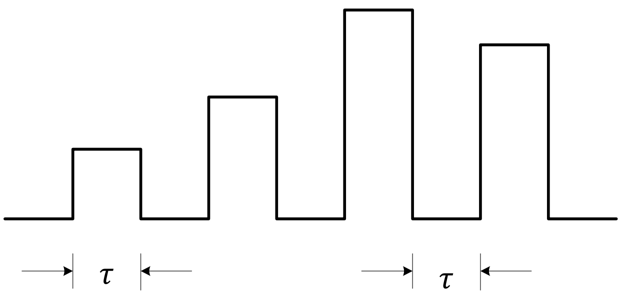


Figure 8.4 (a) Illustration of Maximum Frequency in PAM Signal

$f_{max} = \frac{1}{\tau+\tau} = \frac{1}{2\tau}$ 8.4

Therefore, the bandwidth required for the transmission of a PAM signal would be equal to the maximum frequency f_{max} given by the Eq (8.4).

Thus, we have transmission bandwidth

$BW \geq f_{max}$ 8.5

But $f_{max} = \frac{1}{2\tau}$

Hence $BW \geq \frac{1}{2\tau}$

Again, since $\tau << \frac{1}{2f_m}$

Therefore $BW \geq \frac{1}{2\tau} >> f_m$ 8.6

$BW >> f_m$ 8.7

8.3 Demodulation of PM Signals

As discussed earlier, demodulation is the reverse process of modulation in which the modulation signal is received back from a modulated signal. For pulse-amplitude modulated (PAM) signals. The demodulation is done using a holding circuit. Fig. 8.4 (b) shows the block diagram of a PAM modulation.

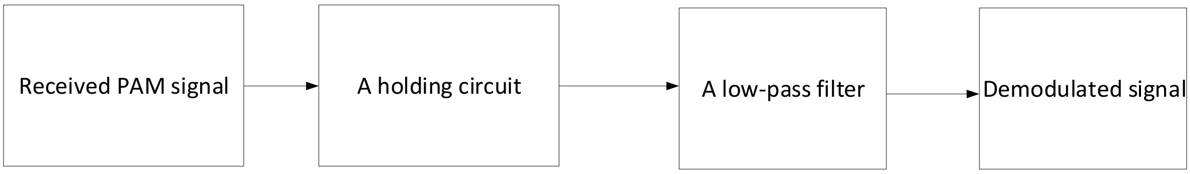


Fig 8.5 (b) A block diagram of PAM Demodulator

In this method, the received PAM signal is allowed to pass through a holding circuit and a low pass filter (LPF) as shown in above figure. Now, Fig. 8.6(a) illustrates a very simple holding circuit. Here the switch 'S' is closed after the arrival of the pulse and it is opened at the end of the pulse. In this way, the capacitor 'C' is changed to the pulse amplitude value and it holds the value during the internal between the two pulses. Hence, the sample values are held as shown in Fig. 8.6 (b). After this holding circuit output is smoothed in low pass filter as shown in Fig. 8.6 (c).

It may be observed that some kind of distortion is introduced due to the holding be observed that some kind of distortion is introduced due to the holding circuit. In fact the circuit of Fig. 8.6b is known as zero-order holding circuit. This zero-order holding circuit considers only the previous sample to decide the value between the two pulses. With these, we bear in mind that the first order hold circuit considers the previous two samples whereas a second order holding circuit consider the previous three samples and so on. However, as the order of the holding circuit increases, the distortion decreases at the cost of the circuit complexity. In fact, the amount of permissible distortion decides the order of the holding circuit.

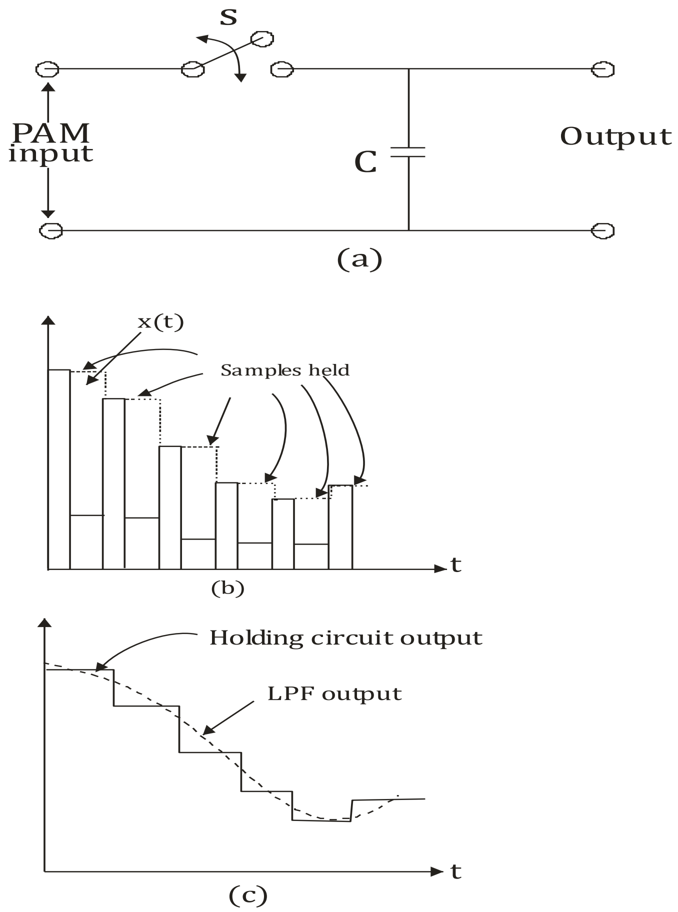


Fig 8.6 (a) A zero-oder holding circuit (b) the output of holding circuit (c) the output of a law pass filter (LPF)

8.4 Transmission of Pulse Amplitude Modulation (PAM)

If the PAM signals are to be transmitted directly i.e. over a pair of wires, than no further signal processing is necessary. However, if they are to be transmitted through the space using an antenna, they must first be amplitude or frequency or phase modulated by a high frequency carrier and only then they can be transmitted. Thus, the overall system will be then known as PAM-AM or PAM-FM or PAM-PM respectively. At the receiving end, AM or FM or PM detection is first employed to get the PAM signal and then to message signal is recovered from it.

Example 8.1

For a PAM transmission of voice signal having maximum frequency equal to f_m = 6 kHz, calculated the transmission bandwidth. It is given that the sampling frequency f_s = 16 kHz and the pulse during $\tau = 0.2T_s$.

Solution: From the fact that the sampling period T_s is expressed as:

$$T_s = \frac{1}{f_s} = \frac{1}{16 \times 10^3}\ s$$
$$T_s = 6.25 \times 10^{-5}\ s$$
$$\mathbf{T_s = 6.25\ \mu s}$$
Also, τ is given as
$$\tau = 0.2T_s$$
$$\tau = 0.2 \times 6.25$$
$$\boldsymbol{\tau = 1.25\ \mu s}$$
But the transmission bandwidth for PAM signal is expressed as
$$BW \geq \frac{1}{2\tau}$$
$$BW \geq \frac{1}{2 \times 1.25 \times 10^{-6}} \geq \frac{10^6}{2 \times 1.25}$$
$$\mathbf{BW \geq 40\ kHz}$$

8.4.1 Drawback of Pulse Amplitude Modulation (PAM) Signal

a. The bandwidth required for the transmission of a PAM signal is very large in comparison to the maximum frequency present in the modulating signal.

b. Since the amplitude of the PAM pulse varies in accordance with the modulating signal therefore the interference of noise is maximum is a PAM signal. This noise cannot be removed easily.

c. Since the amplitude of the PAM signal varies, therefore, the also varies the peak power required by the transmitter with modulating signal.

8.5 Pulse Time Modulation

As discussed earlier, there are two types of pulse time modulation i.e. pulse width modulation and pulse position modulation. In both Pulse Width Modulation (PWM) and Pulse Position Modulation (PPM), some time-parameter of the pulse is modulated. In PWM, width of the pulses is varied whereas in PPM position of the pulse is varied. However, in both the methods amplitude of the pulses is kept constant.

8.5.1 Production of PWM and PPM Signals

The figure 8.6 shows the block diagram to generate both PWM and PPM with their corresponding waveform. It consist of both sampling and modulation operation.

The saw-tooth generator generates the saw-tooth signal of frequency f_s (i.e. period T_s). The saw-tooth, also known as sampling signal is applied to the inverting input of the comparator.

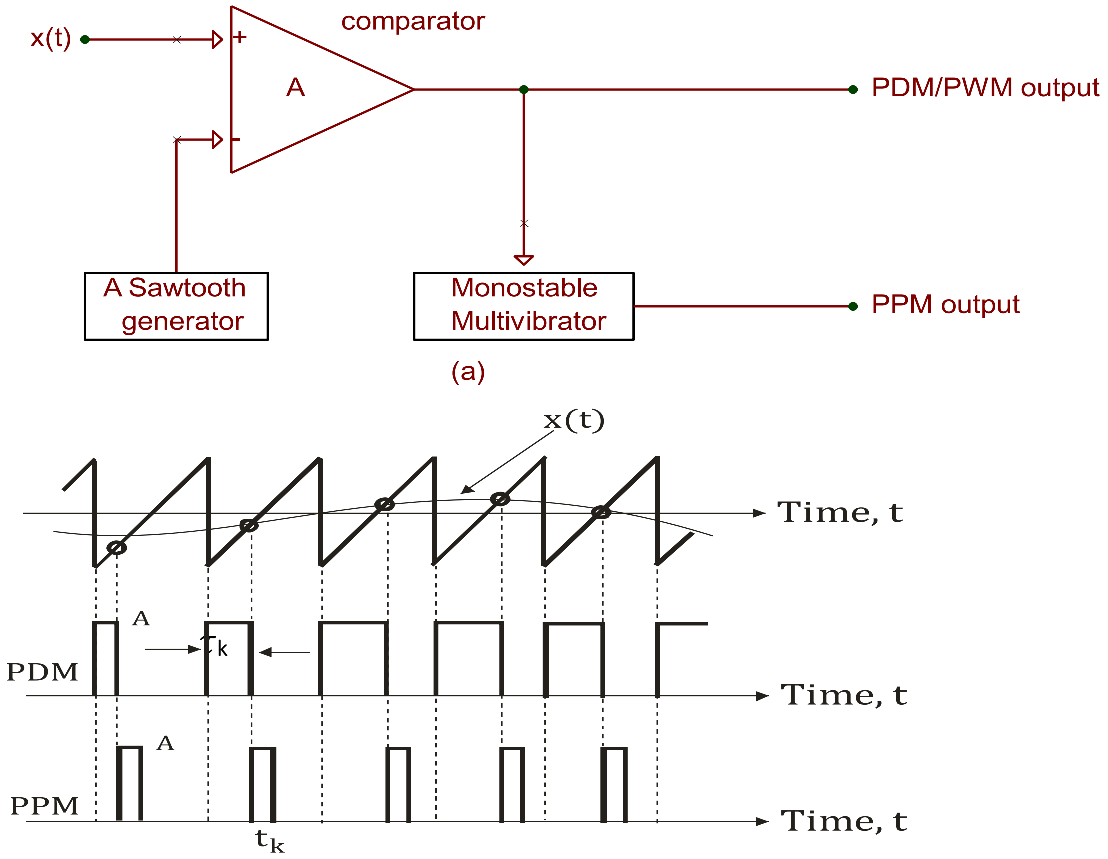


Figure 8.7 Generation of PPM and PDM waveform (a) Block diagram (b) waveforms

The modulating signal x(t) is made to apply to the non-inverting input of the comparator. The output of the comparator is high only when instantaneous value of x(t) is higher than that of saw-tooth waveform. Hence, the leading edge of PDM signal occurs at the fixed time period say kTs whereas the trailing edge of output of comparator depends on the amplitude of signal x(t). Now, when saw-tooth waveform voltage is greater than voltage ofx(t)at that instant, the output of comparator remains zero. The trailing edge of the output of comparator (PDM) is modulated by the signal x(t)If the saw-tooth waveform is reversed, then trailing edge will be fixed and leading edge will be modulated. If saw-tooth waveform is replaced by a triangular waveform, then both leading trailing edges will be modulated.

The pulse duration modulation (PDM) or PWM signal is nothing but output of the comparator. The amplitude of this PDM or PWM signal will be positive saturation of the comparator, which is shown as 'A' in the waveform. The amplitude is same for all the pulses.

To generate pulse position modulation PPM, PDM signal is used as the trigger input to one monostable multivibrator. The monostable output remains zero until it is triggered. The monostable is triggered on the falling (trailing) edge of PDM. The output of monostable then switches to positive saturation level 'A'. This voltage remains high for a fixed period then goes low. The width of the pulse can be determined by monostable. The pulse is thus delayed from sampling time kTs depending on the amplitude of signal x(t)at kTs.

As can be seen from the waveform, both PPM and PDM posses DC value. The amplitude of all the pulses is same. Therefore, non-linear amplitude distortion as well as noise interference does not affect the detection at the receiver. However both PPM and PDM needs a sharp rise time and fall time for pulses in order to preserve the message information. Rise time should be very less than T_s i.e.

$$t_r \ll T_s$$

And transmission bandwidth must be $BW \geq \frac{1}{2t_r}$

However, it should be noted that the transmission bandwidth of PPM and PDM is higher the PAM. The power requirement of PPM is less than that of PDM because of short duration pulses. But it may be further reduced by transmitting only edges rather than pulse.

Thus, transmission bandwidth of PDM and PPM:

$$BW \geq \frac{1}{2t_r} \qquad 8.8$$

8.6 Demodulation of PWM

In this circuit, the transistor T_1 as has shown in figure 8.8 acts as an inverter. During the time-interval from A to B when the PWM signal is high, the input to the transistor T_2 is low. Thus, during this time-interval, the transistor T_2 is cut-off and the capacitor C is charged through an RC combination. Also during the time interval from B to C, when the PWM signals is low the input to the transistor T_2 is high and so it gets saturated. The capacitor C then discharges very rapidly through transistor T_2. The collector voltage of transistor T_2 during the interval from B to C is then low. Hence, the waveform whose envelope is the modulating signals. When this is passed though a second order OP-AMP low-pass filter, the desired demodulated signal is obtain.

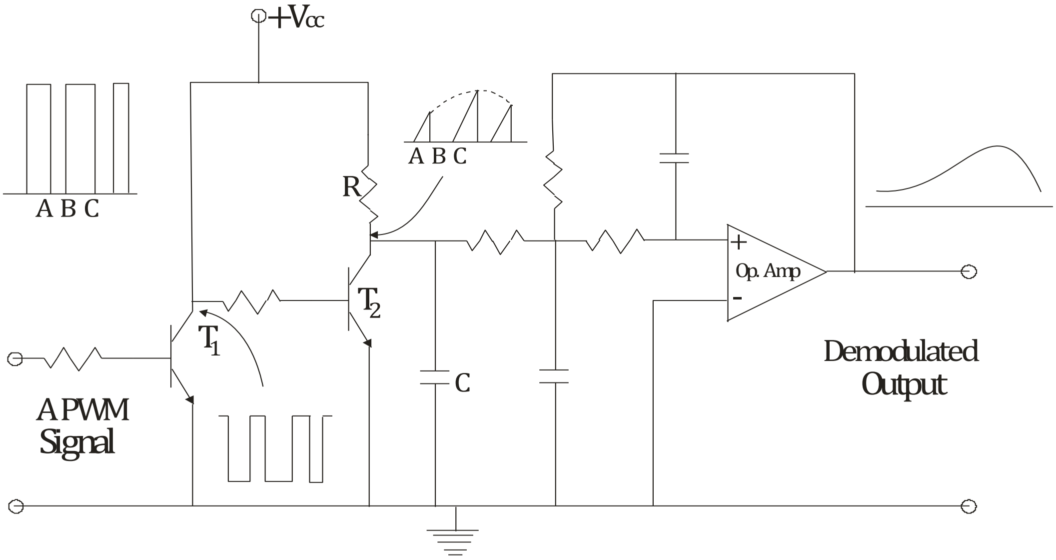


Fig 8.8: A PWM Demodulator circuit

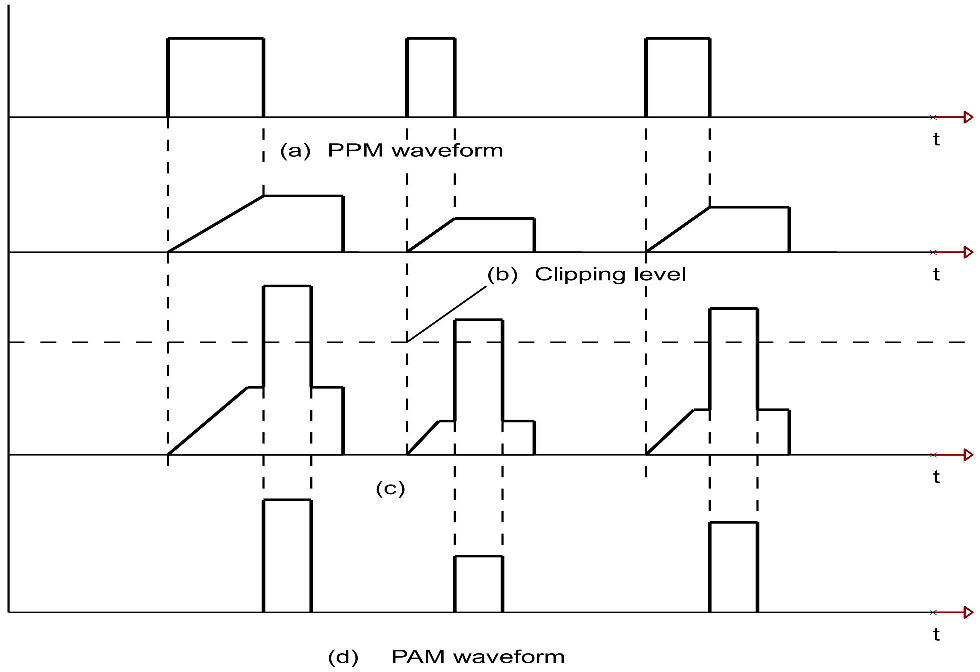


Fig 8.9 Demodulator of PWM signals (a) PWM waveform (b) Ramp waveform with Portich, (c) Ramp waveform with locally generated pulse on porch (d) PAM waveform

8.7 Demodulation of PPM Signals

In Fig. 8.9 shows a PPM demodulator circuit. This circuit makes use of the fact that the gaps between the pulses of a PPM signal contain the information regarding the modulating signal. During the gap from A to B between the pulses, the transistor is cut-off and the capacitor C gets charged through RC combination. During the pulses duration from B to C, the capacitor is discharged through the transistor and so, the collector voltage becomes low. Hence, the waveform at the collector is approximately a saw-tooth waveform whose envelope is the modulating signal. Now, when this is passed through a second order OP-AMP low pass filter, the desired demodulated output is obtained.

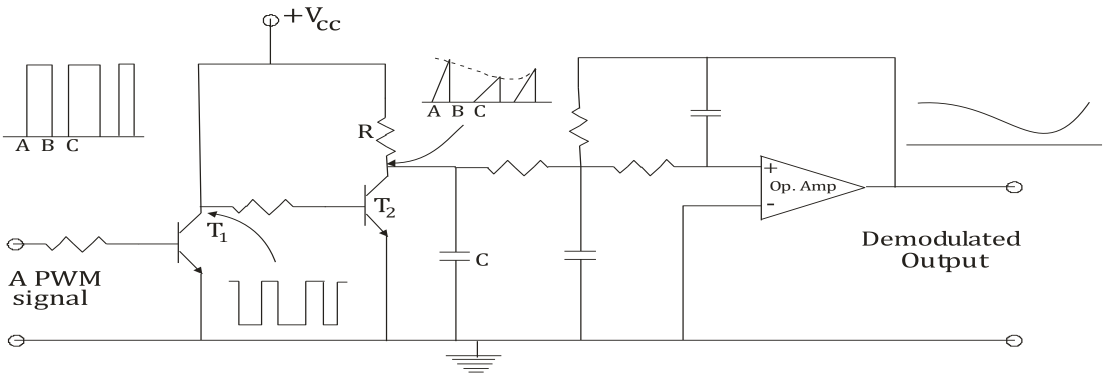


Fig 8.10 PPM Demodulator Circuit

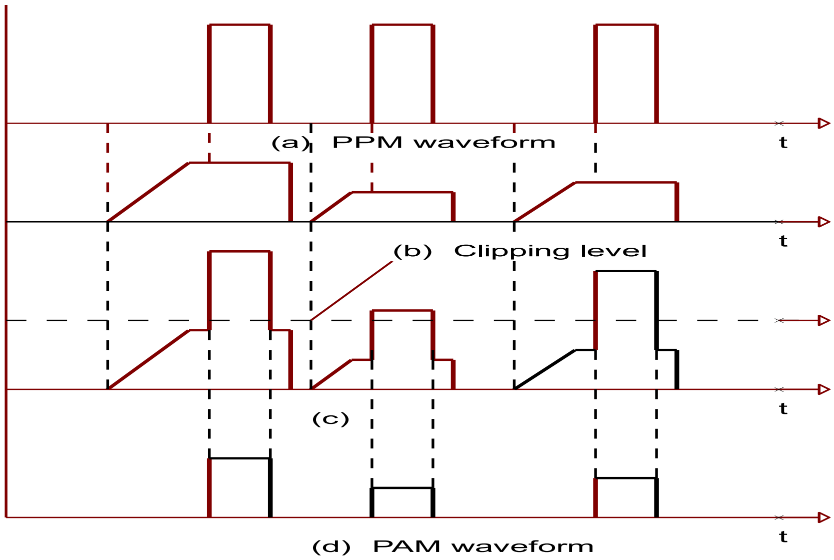
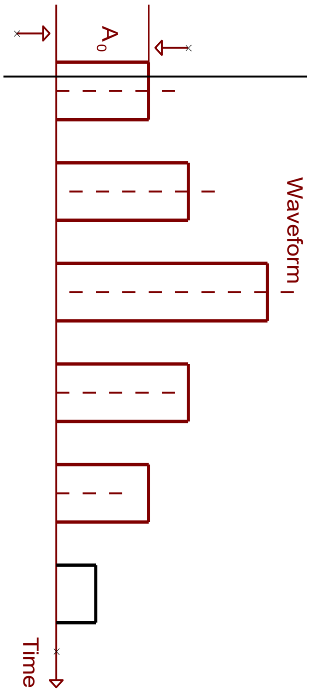
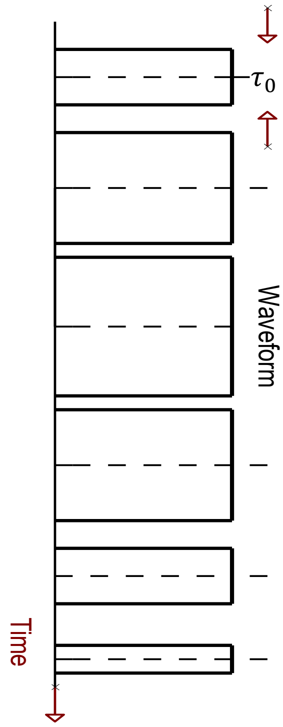
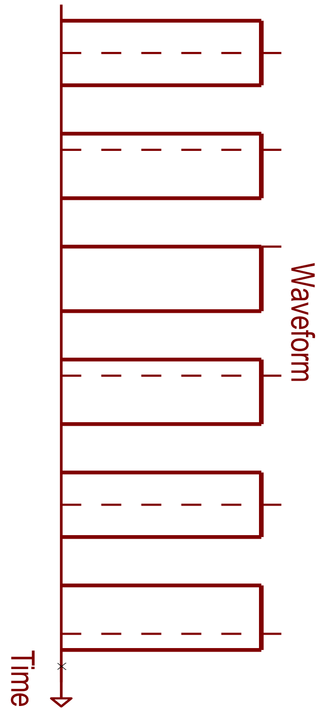


Fig 8.11 Demodulator of PWM signals (a) PPM waveform (b) Ramp waveform with Portich, (c) Ramp waveform with locally generated pulse on porch (d) PAM waveform

Comparison of various pulse analog modulation methods

In this section, let us compare PAM, PWM and PPM in the form of a table 8.1

Table 8.1 Comparison of PAM PPM and PDM

S/N	Pulse Amplitude Modulation (PAM)	Pulse width/Duration Modulation (PWM) or(PDM)	Pulse position modulation (PPM)
1			
2	Amplitude of the pulse is proportional to amplitude of modulating signal	Width of the pulse is proportional to amplitude of modulating signal	The relative position of the pulse is proportional to the amplitude of modulating signal
3	The bandwidth of the transmission channel depends on width of the pulse	Bandwidth of transmission channels depends on rise time of the pulse	Bandwidth of transmission channels depends on rising time of the pulse
4	The instantaneous power of the transference varies	The instantaneous power of the transmitter varies.	The instantaneous power of the transmitter remains constant.
5	Noise interference is high	Noise interference is minimum.	Noise interference is minimum.
6	System is complex.	Simple to implement.	Simple to implement.
7	Similar to amplitude modulation.	Similar to frequency modulation.	Similar to phase modulation.

8.8 Pulse Digital Modulation Techniques

Digital modulation techniques include PCM, DM and DPCM. This section describes each of them and also recovering approximate analog message signal from them.

8.8.1 Pulse Code Modulation (PCM)

The most common technique to analog signal to digital (digitization) is called pulse code modulation (PCM). A PCM encoder has three processes as shown in Fig. 8.10

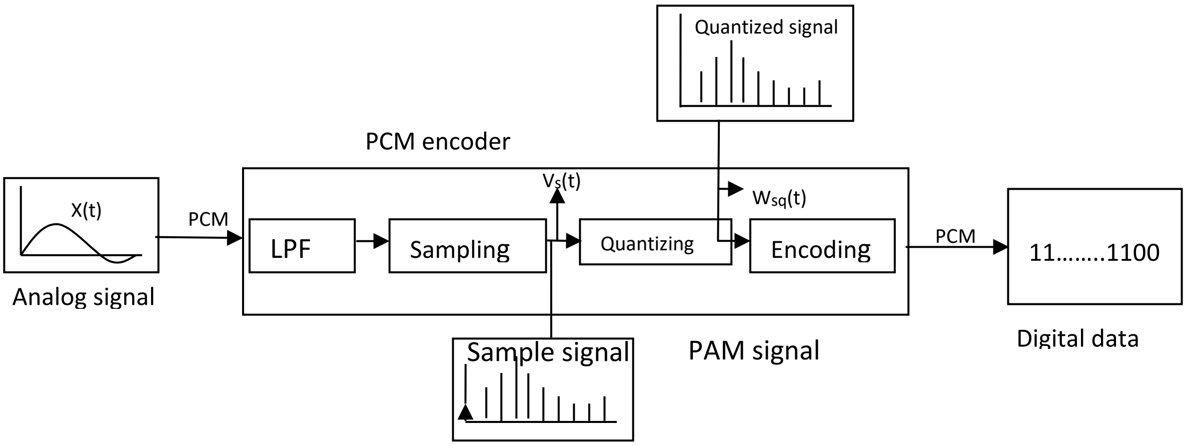


Fig 8.12 Component of PCM

The processes are as follows

1. The analog signal is sampled.
2. The sampled signal is quantized.
3. The quantized values are encoded as stream of bits.

In PCM, the total amplitude range which the signal may occupy is divided into a number of standard levels, as shown in Fig. 8.12. Since these levels are transmitted in a binary code. The actual number of levels is a power of 2; 16 levels are shown here for simplicity, but practical systems use as many as 128. By a process called quantization, the level actually sent at any sampling time is the nearest standard (or quantum) level. As shown in Fig. 8.13, should the signal amplitude be 6.9 V, at any time, it is not sent as a 6.9 V pulse, as it might have been in PAM, nor as a 6.9 µs wide pulse as in PWM, but simply as the digit 7 V is the standard amplitude nearest to 6.9 V.

Furthermore, the digit 7 is sent at that instant of time as a series of pulse corresponding to the number 7. Since there are 16 level (2^4), 4 binary places are required; the number becomes 0111, and could be sent as a 0PPP, where P = Pulse and 0 = no-pulse. Actually, it is often sent as a binary number back-to-front; i.e. as 1110, or PPP0, to make demodulation easier.

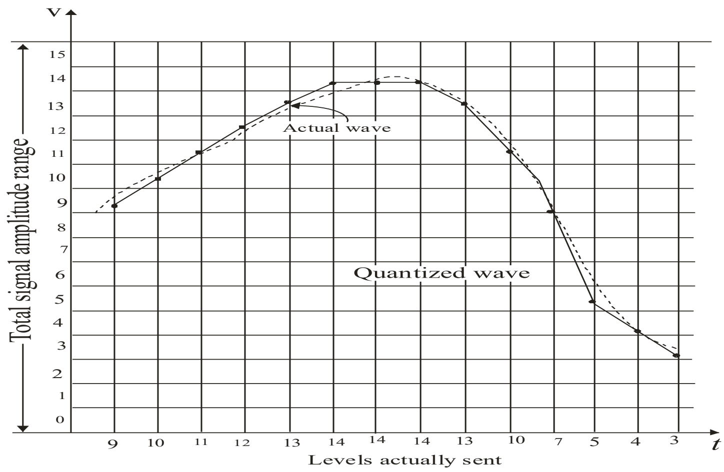


Fig 8.13 Quantization of signal for pulse code modulation

As shown in Fig. 8.13, the signal is continuously sampled, quantized, coded as sent, as each sample amplitude is converted to the nearest standard amplitude and into the corresponding back-to-front binary numbers. Provided sufficient quantizing levels are used, the result cannot be distinguished from that of analog transmission.

A supervisory or signaling bit is generally added to each code group representing a quantized sample. Hence each group of pulses denoting a sample, here called a word, is expressed by means of n+1 bits, where 2^n is the chosen number of standard levels. However, for all the modulation types covered so far, whether pulse or continuous wave (CW) have been analog representation of the message. Pulse-code modulation (usually abbreviated as PCM) is distinctly different in concept coded group of digital (discontinuous-amplitude) pulse. Delta modulation (normally abbreviated as DM) is variation of PCM.

In an analog modulation disused in the previous chapters, the modulation parameter varies continuously and can take on any value corresponding to the range of the message. When the modulated wave is mixed with noise, there is no way for the receiver to discern the exact transmitted value. Supposed, however, that only a few discrete (i.e. discontinuous) value is allowed for the modulated parameter and if the separation between these values is large compared to detect at the receiver precisely which specific value are intended. Thus, the effects of random noise can be virtually eliminated which is the whole idea of digital modulation.

As already stated, PCM is an abbreviation of Pulse Code Modulation. In the pulse modulation system discussed in this chapter, the actual value of the sample was transmitted by operating on a standard pules. In PCM, the continuous time function is sampled in the usual manner, and the magnitude of each sample is then rounded off to the nearest value of these permitted to be transmitted since only a finite number of values are permitted, the sampled value may be transmitted by a code group. The process of rounding off is called quantizing levels.

Although many codes are possible, the digital code is most popular because of its simplicity in detection and instrumentation. In a digital code, only two levels are transmitted usually 1 and 0 corresponding to the carrier ON and OFF respectively. If we want to transmit the sampled value in 1-volt step from zero to 30volts, then a five digit 001 to 111 will be required, representation 2^0 and 2^3-1 respectively. When the pulse on presents, the code of bit 1 is use to indicate it and should to pulse appears between t_1 and t_5, then this corresponds to $2^0+2^1+2^2+2^3+2^4 = 31$, should a pulse appear at t_4 and t_1 as illustrated, then the corresponding value of the code group would be $2^3+2^0 = 9$.

With a five-digit code represented in figure 8.13 all integers from 0 through 31 are possible, thus, yielding a total of 32 levels. A six-digit code group can be used or 64 levels, 7 digits for 128 levels, or in general an n digit code allows 2^n values to be transmitted.

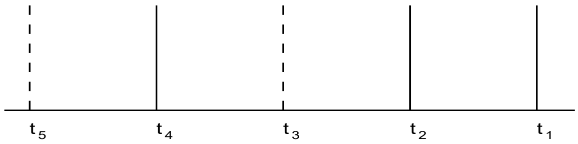


Fig 8.14. Five-digit binary code group

Modulating wave and its associated digital code is shown in Fig. 8.15. In practice signal would most likely appear with both positive values with an average of zero. To ease the coding problem, a bias is often added so that signal will only assumed positive values as shown in Fig. 8.15.

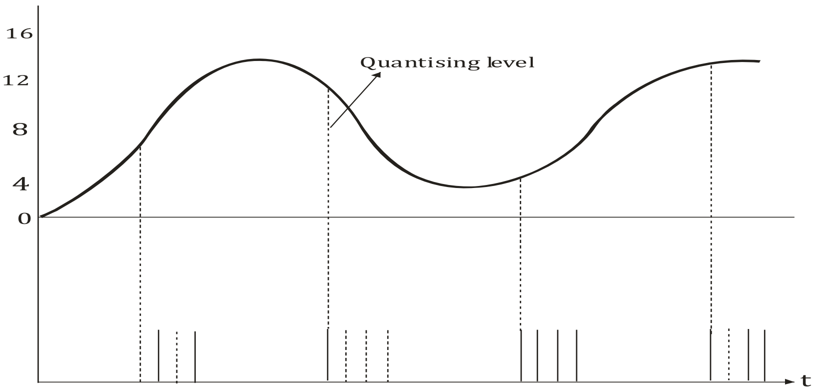


Figure 8.15 Quantizing PCM

However, for all the modulation types covered so far, whether pulse or continuous wave (CW) have been analog representation of the message. Pulse-code modulation (usually abbreviated as PCM) is distinctly different in concept coded group of digital (discontinuous-amplitude) pulse. Delta modulation (normally abbreviated as DM) is variation of PCM.

In an analog modulation discussed in the presiding chapters, the modulated parameter varies continuously and can take on any value corresponding to the range of the message. When the modulated wave is mixed with noise, there is no way for the receiver to discern the exact transmitted value. Supposed, however that only a few discrete (i.e. discontinuous) values are allowed for the modulated parameter and if the separation between these values is large composed to the noise disturbances, it will be a simple meter to detect at the receiver precisely which specific values are intended. Thus the effects of random noise can be virtually eliminated which is the whole idea of digital modulation.

As already stated, PCM is an abbreviation of pulse code modulation. In the pulse modulation system discussed in this chapter, the actual value of the sample was transmitted by operating on a standard pulse. In the PCM, the continuous time function is sampled in the usual manner, and the magnitude of each sample is then rounded off to the nearest value of those permitted to be transmitted since only a finite number of values are permitted, the sampled value may be transmitted by a code group. The process of rounding off is called quantizing, while the possible levels permitted are the quantizing levels.

Although many codes are possible, the digital code is most popular because of its simplicity in detection and instrumentation. In a digital code, only two levels are transmitted usually 1 and 0 corresponding to the carrier ON and OFF respectively if we want to transmit the sampled values in 1-volt step.

Most typical signal for example speech has a large peak-to-r.m.s factor. Compression is used to reduce the large peak and thus increase the lower values. This will help in using fewer number of quantizing levels for a given accuracy and will also reduce channel bandwidth. The receiver then expands the signal in the inverse manner in which it was compressed and restores the average value.

Eye Patterns

We can get a good qualitative indication of the performances of a PCM system by examining the bit (abbreviation for binary digit) stream on a CRO (an abbreviation of cathode ray oscilloscope). The time base of CRO is set so that triggers at the bit rate and yields a sweep lasting 1 time-slot duration. In the ideal case of no noise and no bandwidth restriction, the bit-stream waveform would appear as at the left in Fig. 8.16(a) and the CRO pattern at the right.

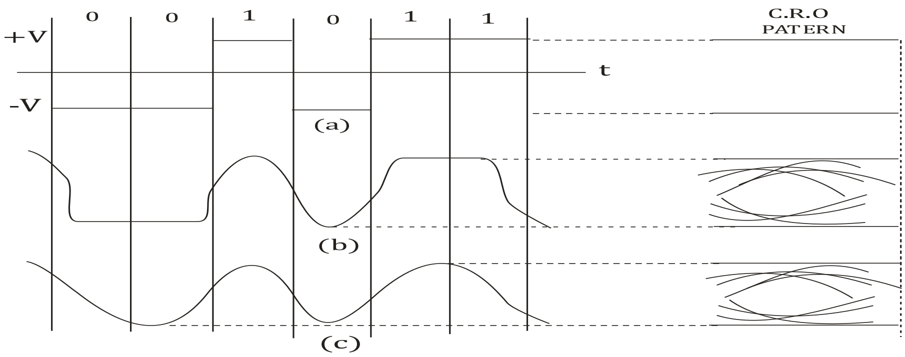


Fig 8.16 Eye Patterns in PDM

The CRO pattern would consist of two horizontal lines. In Fig. 8.16(b) the bit-train waveform illustrates the effects limited of bandwidth and CRO depicts eye and hence the pattern is called an eye-pattern. A longer bit train would add more traces to the CRO pattern, generally filling in the periphery and leaving an opening an eye, in the center of the figure. In Fig. 8.16(c), the eye is closed somewhat indicating further limiting bandwidth. The addition of noise to the bit stream would close the eye still further. Thus eye pattern of PCM system gives information regarding noise and bandwidth.

8.8.2 Quantizing and Coding

As shown in Fig. 8.12, the continuous signal $V_{(t)}$ is first passed through low pass filter having bandwidth W and sampled to give v(t). The sampled signal is then quantized to the nearest predetermined value with quantize level x. The resulting sampled and quantized signal is encoded in equal time (by virtue of sampling) and quantizing. Finally $X_{sq}(t)$ is operated on by an encoder with ON/OFF quantized samples to appropriate digital code words, one code word for each sample, and generates the corresponding baseband PCM signal as a digital waveform. Thus PCM generator shown in Fig. 8.10 is nothing but an analog to digital converter (normally abbreviated as A/D).

The parameters on the encoded signal depends upon the number of quantum level x, which is the ration of $|V_{in}|/V_{in(max)}$ for each code word must uniquely represent one of the possible quantized sample. Let y be the ration of $V_{out}/V_{out(max)}$, and the parameters chosen such that

$$y = \frac{\log(1+\mu x)}{\log(1+\mu)} \qquad \begin{array}{|c|} 0 < x < \frac{1}{A} \\ -\frac{1}{A} < x < 0 \end{array} \qquad 8.9$$

For this text, we assume $0 < x < \frac{1}{A}$ and $\mu = 100$ throughout

$$x = \frac{V_{in}}{V_{in_{max}}} \qquad 8.10$$

$$y = \frac{V_{out}}{V_{out_{max}}} \qquad 8.11$$

Step size $= \frac{V_{p-p}}{2^n}$ 8.12

No of levels $= 2^n$ 8.13

Error or max noise $= \frac{V_{max}}{2^n}$ 8.14

Signal − to − noise ratio $\left(\frac{S}{N}\right) = 2^{2n}$ 8.15

Normalize quantization level $= V_{max} - \frac{V_{max}}{2^n}$ 8.16

At the final step of PCM generations, the baseband signal may modulate RF carrier for transmission purposes.

8.8.3 Companding

Changes in amplitude often occur more frequently in the lower amplitudes than in the higher ones when the instantaneous amplitude of the analog signal is not uniformed. To achieve non-uniform, quantization, companding and expanding process is used. The signal is companded at the sender before conversion; it is expanded at the receiver after conversion. Companding means reducing the instantaneous voltage amplitude for large values; expanding is the opposite process. Companding gives greater weight to strong signals and less weight to weak ones. If the steps are uniform in size, small variation in amplitude signal power signal-to-quantizing ration then large variation in amplitude signal is seen. To overcome this with a handicap of a fixed number of levels, it is advantageous to taper the step size so that the overall transmission is distortion free.

Thus before application to quantizer, the signal is passed through a nonlinear network which has an input out characteristic as shown in Fig. 8.17. As a result of which a given signal change at low amplitude will carry the quantizer through more steps than will be the case at large amplitudes, a signal transmitted through a network with the characteristic shown in Fig. 8.17 will have the extremities of its waveform compressed, the compression being more pronounced with increased amplitude. Hence the network is called a compressor. The inverse process is performed by an expander. The combination of compressor and an expander is called a compander which then perform the operation of companding.

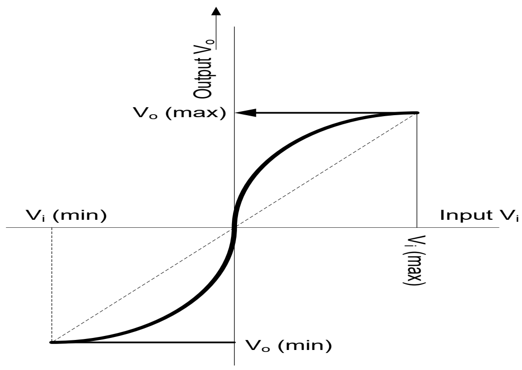


Fig 8.17 Input-Out Characteristic Providing Compression

In Fig. 8.17, dotted line shows no compression while continuous line depicts compression close together at low signal amplitudes and further apart at large amplitudes. Such variation of step size gives an improvement in signal to noise for small signal although strong smalls will be diminished. Tapering of steps is very useful in speech signals.

Although it is possible to design a quantizer with tapered steps, it is more practical to achieve an equivalent effect by distorting the signal before applying it to the transmitter quantizer. An inverse distortion has to be introduced.

If for example we have the max and minimum voltage signal to be +20 V and -20 V i.e. baseband signal of ±20 V for companding process, let n-bit be such that n=3 for $0 < x < \frac{1}{A}$, $\mu = 100$, from Eq (8.9) to (8.6).

From Eq. 8.12

$$\text{Step size} = \frac{40}{2^3} = \frac{40}{8} = \mathbf{5\ V}$$

From Eq (8.14)

$$\text{Error or max noise} \quad = \frac{20}{2^3} = \mathbf{2.5\ V}$$

From Eq (8.16)

$$\text{Normalize quantization level} = 20 - \frac{20}{2^3} = \mathbf{17.5\ V}$$

Table 8.2

	Range	**Normalize Quantization (Voltage Level)**	**Code**
7	15 - 20	17.5	111
6	10 - 15	12.5	110
5	5 - 10	7.5	101
4	0 - 5	2.5	100
3	(-5) - 0	-2.5	011
2	(-10) - (-5)	-7.5	010
1	(-15) - (-10)	-12.5	001
0	(-20) - (-15)	-17.5	000

Example 8.2

Let the sinusoidal signal in Fig. 8.18(a) be sampled at time t = 0, 5, 10,……40 s. The maximum amplitude of the signal is 16 V. We are to

i. Draw the PAM signal using single line

ii. Draw the PCM single using three bit per sample. Prepare a table to show the possible voltage range, quantization levels and corresponding PCM codes. Determine the signal-to-quantizing noise ration in dB.

For $0 < x < \frac{1}{A}$

$$\text{Step size} = \frac{32}{2^3} = \mathbf{4\ V}$$

Error or Maximum Noise $= \frac{16}{2^3} = \mathbf{2\ V}$
Normalize quantization level $= 16 - \frac{16}{8} = \mathbf{14\ V}$

Table 8.3

	Range	**Normalize Quantization (Voltage Level)**	**Code**
7	12 – 16	14	111
6	8 – 12	10	110
5	4 – 8	6	101
4	0 – 4	2	100
3	(-4) - 0	-2	011
2	(-8) - (-4)	-6	010
1	(-12) - (-8)	-10	001
0	(-16) - (-12)	-14	000

In companding process,
From Eq (8.9)
when $\mu = 100$
For $t = 0$ s, V_{in} = 2V, i.e. when the PAM pulse touches the sinusoidal signs, done through sampling process, $x = \frac{|V_{in}|}{V_{in(max)}} = \frac{2}{16} = \mathbf{0.125}$

$$y = \frac{\log(1+\mu x)}{\log(1+\mu)} = \frac{\log(1 + 1/8 \times 100)}{\log(1+100)} = \frac{\log 13.5}{\log 101} = \mathbf{0.564}$$

$$y = \frac{V_{out}}{V_{out(max)}}$$

$V_{out} = y \times V_{out(max)} = 0.56 \times 16 = \mathbf{9.02\ V}$
Checking from table 8.3, the range where it fell is between 8-12 volts and the corresponding code is (110) i.e. **V_{out} 9.02 V, code (110).**

For $t = 5$ s, V_{in} = 0.5V, again when the sampled signal is 0.5V from the graph of fig 8.18 (a)

$$x = \frac{0.5}{16} = \mathbf{0.03125}$$

$$y = \frac{\log(1 + 0.03125 \times 100)}{\log(1+100)} = \frac{\log 4.125}{\log 101} = \mathbf{0.3070}$$

$V_{out} = y \times V_{out(max)} = 0.307 \times 16 = \mathbf{4.9128\ V}$
Checking from table 8.3, the corresponding range for $y = 4.913$V is between 4 and 8, with a code of **(101)**
For $t = 10$ s, V_i = 10 V

$$x = \frac{|V_{in}|}{V_{in(max)}} = \frac{10}{16} = \mathbf{0.625}$$

$\mathrm{y} = \frac{\log(1+\mu x)}{\log(1+\mu)} = \frac{\log(1+0.625\times100)}{\log 101} = \mathbf{0.899}$

$\mathrm{V_{out}} = \mathrm{y} \times \mathrm{V_{out(max)}} = 0.899 \times 16 = \mathbf{14.39\ V}$

The same way, $\mathbf{V_{out}} = (\mathbf{14.39})$**, code (111)**

For $t = 15$ s, V_i = 14.5 V

$$x = \frac{|\mathrm{V_{in}}|}{\mathrm{V_{in(max)}}} = \frac{14.5}{16} = \mathbf{0.90625}$$

$$\mathrm{y} = \frac{\log(1+\mu x)}{\log(1+\mu)} = \frac{\log(1+0.90625\times100)}{\log 101} = \mathbf{0.9789}$$

$\mathrm{V_{out}} = \mathrm{y} \times \mathrm{V_{out(max)}} = 0.9789 \times 16 = \mathbf{15.66\ V}$

The same way, $\mathbf{V_{out}}(\mathbf{15.66V})$**, code (111)**

For $t = 20$ s, V_i = -0.5 V

$$x = \frac{|\mathrm{V_{in}}|}{\mathrm{V_{in(max)}}} = \frac{0.5}{16} = \mathbf{0.03125}$$

$$\mathrm{y} = \frac{\log(1+\mu x)}{\log(1+\mu)} = \frac{\log(1+0.03125\times100)}{\log 101} = \mathbf{0.3070\ V}$$

$\mathrm{V_{out}} = \mathrm{y} \times \mathrm{V_{out(max)}} = 0.3070 \times 16 = \mathbf{-4.9\ V}$

The same way, $\mathbf{V_{out}}(\mathbf{-4.9V})$**, code (010)** looking at the range in the table 8.3

For $t = 25$ s, V_i = -1 V

$$x = \frac{|\mathrm{V_{in}}|}{\mathrm{V_{in(max)}}} = \frac{1}{16} = \mathbf{0.0625}$$

$$\mathrm{y} = \frac{\log(1+\mu x)}{\log(1+\mu)} = \frac{\log(1+0.0625\times100)}{\log 101} = \mathbf{0.4292}$$

$\mathrm{V_{out}} = \mathrm{y} \times \mathrm{V_{out(max)}} = 0.4292 \times 16 = \mathbf{-6.87\ V}$

The same way, $\mathbf{V_{out}}(\mathbf{-6.87V})$**, code (010)**

For $t = 30$ s, V_{in} = -10 V

$$x = \frac{|\mathrm{V_{in}}|}{\mathrm{V_{in(max)}}} = \frac{10}{16} = \mathbf{-0.625}$$

$$\mathrm{y} = \frac{\log(1+\mu x)}{\log(1+\mu)} = \frac{\log(1-0.625\times100)}{\log 101} = \mathbf{0.7853}$$

$$\mathrm{V_{out}} = \mathrm{y} \times \mathrm{V_{out(max)}} = 0.7853 \times 16 = \mathbf{12.5651\ V}$$

The same way, $\mathbf{V_{out}}(\mathbf{-12.565\ V})$**, code (000)**

For $t = 35$ s, V_{in} = -16 V

$$x = \frac{|\mathrm{V_{in}}|}{\mathrm{V_{in(max)}}} = \frac{16}{16} = \mathbf{1}$$

$$\mathrm{y} = \frac{\log(1+\mu x)}{\log(1+\mu)} = \frac{\log(1+1\times100)}{\log 101} = \mathbf{1}$$

$$\mathrm{V_{out}} = \mathrm{y} \times \mathrm{V_{out(max)}} = 1 \times 16 = \mathbf{-16\ V}$$

The same way, $\mathbf{V_{out}}(\mathbf{-16\ V})$**, code (000)**

For $t = 40$ s, V_{in} = -0.5 V

$$x = \frac{|V_{in}|}{V_{in(max)}} = \frac{0.5}{16} = \mathbf{0.03125}$$
$$y = \frac{\log(1 + \mu x)}{\log(1 + \mu)} = \frac{\log(1 + 0.03125 \times 100)}{\log 101} = \mathbf{0.3070\ V}$$
$$V_{out} = y \times V_{out(max)} = 0.3070 \times 16 = \mathbf{-4.9\ V}$$
The same way, $\mathbf{V_{out}}(\mathbf{-4.9V})$**, code (010)** looking at the range in the table 8.3

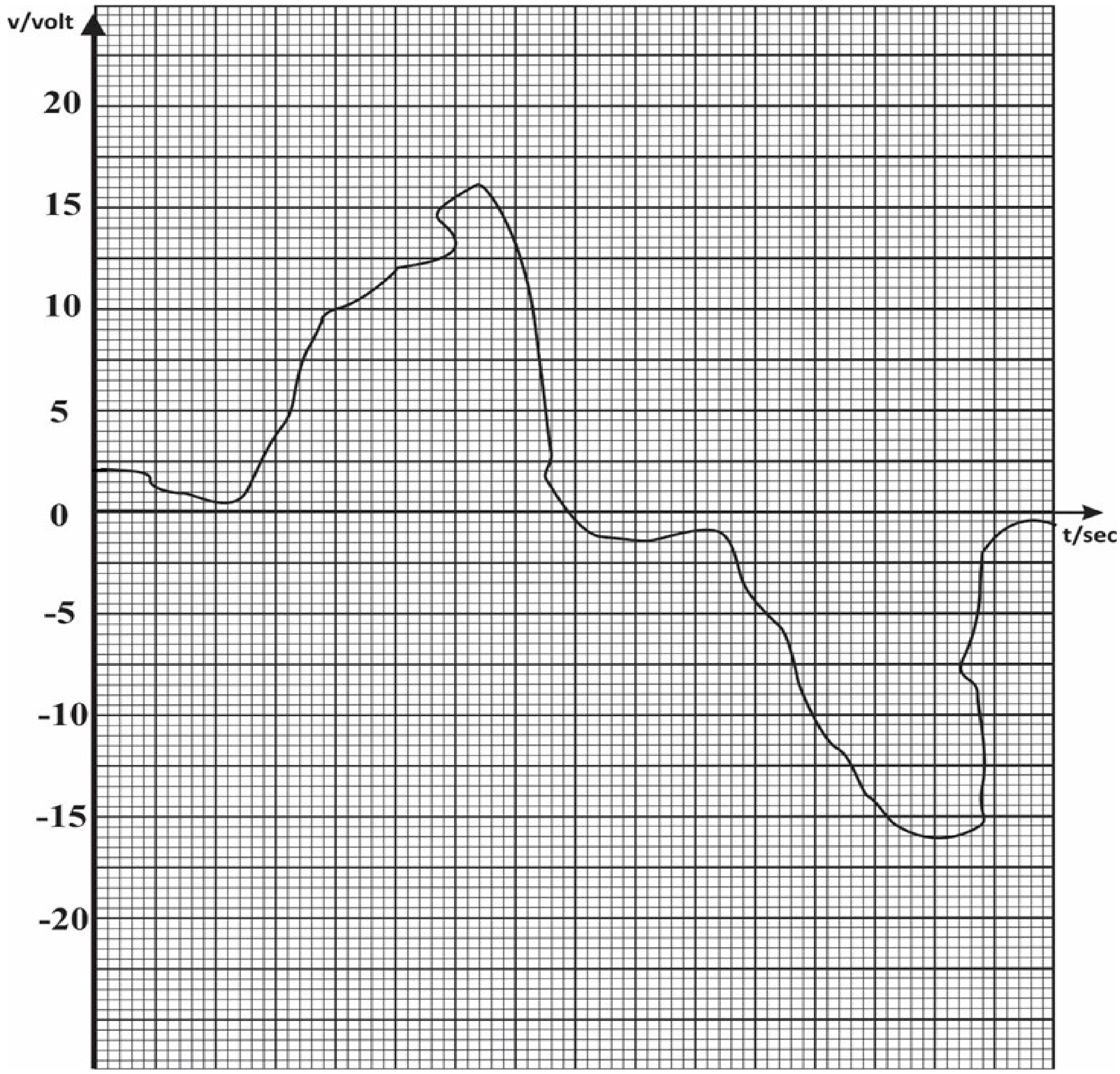


Figure 8.18(a) a sinusoidal signal

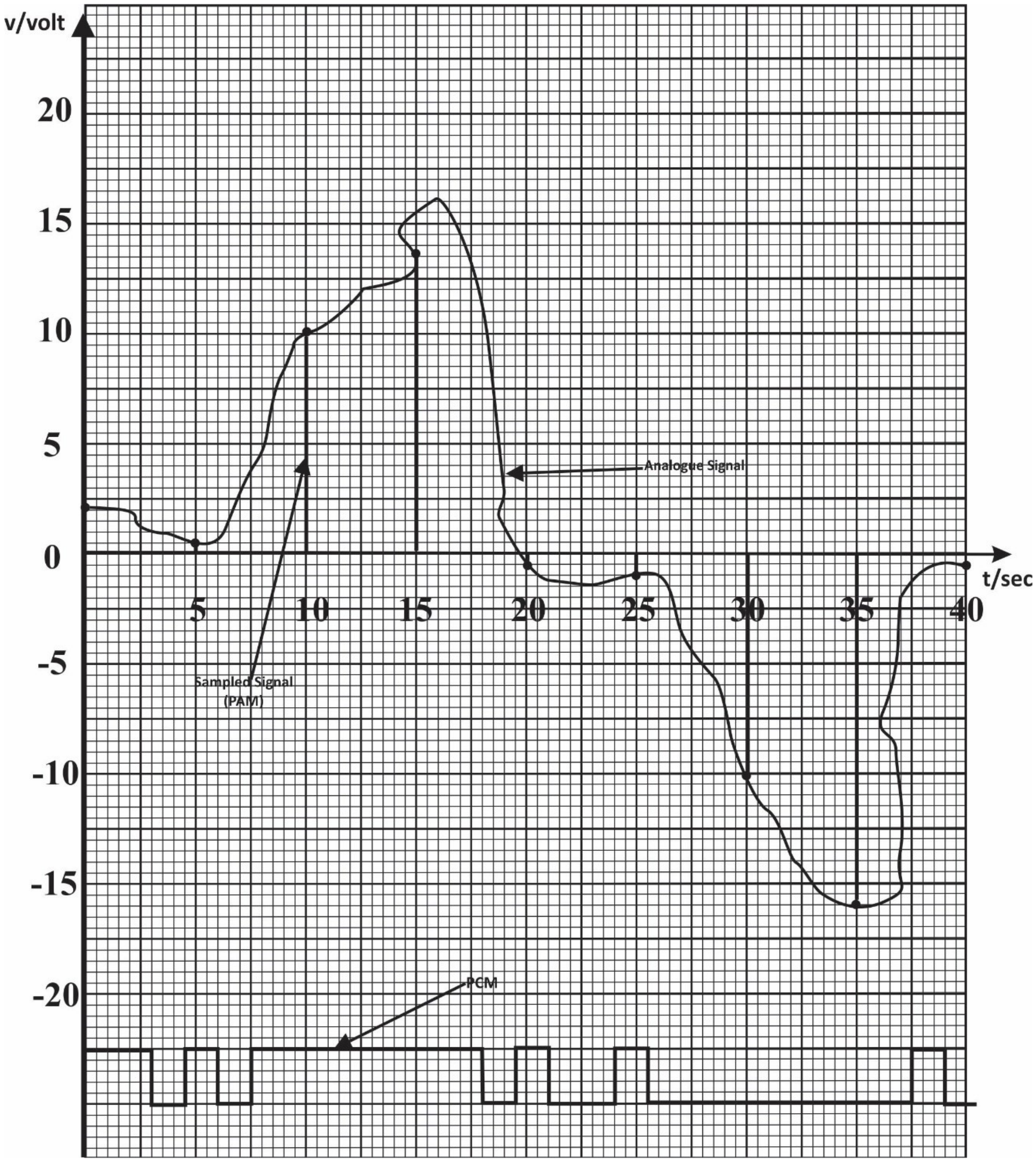


Figure 8.18(b) a sinusoidal signal, PAM and PCM produced

At the receiving end, expanded can be use to produce the baseband signal when the transmitted code is decoded. For example, if a code of 111 is received, with a normalize quantization of 1.875 V, step size of 0.25 and it is expected that the maximum signal voltage will be 2 V, then prepare a table just as seen in table 8.3, trace the receive code 111 to the normalize quantization level and pick the output voltage to be 1.875 V. Then substitute the value of 1.875 V to be V_{out}, hence since $V_{out(max)}$ is 2 V, then solve for y using Eq (8.1), to solve for x, such that $\log(1+\mu x) = y(\log(1+\mu)$. But your baseband signal is gotten from $x = \frac{|V_{in}|}{V_{in(max)}}$, hence $V_{in} = xV_{in(max)}$ then plot Vin against time. This will produce the transmitted signal.

8.9 Delta Modulation (DM)

Delta modulation is an offspring of PCM that has advantage of greatly simplified equipment. DM is the simplest known method for converting an analog signal to digital form. In exchange for these equipment savings, DM generally requires a large transmission bandwidth then PCM. For voice signal, recent developments have brought the bandwidth requirement to a point where DM is less than PCM.

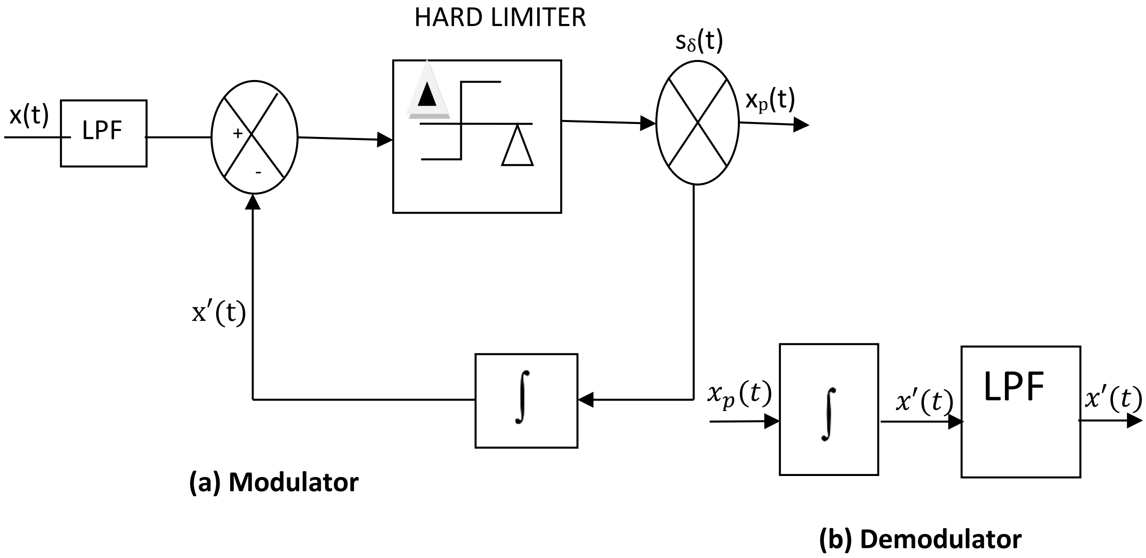


Figure 8.19 Delta Modulation System

Fig. 8.19(a) is a functional block diagram of a delta modulation system.
The message $x(t)$is compressed with a stepwise approximation $x(t)$ by subtraction; the difference being passed through a hard limiter whose output equal $\pm\Delta$ depending on the sign of $x(t) - x'(t)$, this, in turn, modulates the ideal sampling wave $s(t)$ to produce.

$$x_{p(t)} = \sum_k \Delta \text{sign}\,[x(KT_s - x'(KT_s)]\delta(t_o KT_s) \qquad 8.17$$

An impulse waveform which $x_{(t)}$ is generated by integration. Since there are only two possible impulse weight $x_p(t)$, the signal actually transmitted in a binary waveform. The modulator shown in Fig. 8.19(a) consist of an integrator and low pass filter, yielding $x(t)$ plus quantization noise.

To understand these operations, Fig. 8.13 shows typical waveform $x(t), x'(t)$, and $x_p(t)$. To start with $x'(t) < x(t)$ so the first impulse has weight $+E_0$. When feedback and integrated, that impulse produce a stepwise change in $x(t)$ or height + Eo. This process continues through the start-up interval until $x'(t)$ exceeds and causes a negative impulse. If $x(t)$; then remains constant $x(t)$ exhibit a hunting behaviour known as idling noise. When $x(t)$ is changing, $x(t)$ follows it in stepwise fashion unless the rate of change is too great, illustrated at the right of the figure. This slope-overload phenomenon is a basic limitation of delta modulation.

Even with no slope overload distortion, the DM signal, like the PCM system, gives rise to quantization noise is due to the digital approximation of the continuously varying function. When the message signal is the speech signal (for which DM is preferred), the quantization noise will be heard as granular noise, granular noise can be minimize by increasing the sampling rate so that the message signal approximately is better.

Thus we see that the maximum slope is equal to $2\pi f_m A_m$. If the step size is E_0 and the sampling rate is F_s, so that sampling intervals $T_s = \frac{1}{f_s}$ and the corresponding slope is

$$\frac{E_o}{T_s} = E_o f_s \qquad 8.18$$

The slope overload distortion can be avoided, if the following inequality holds

$$E_o f_s \geq 2\pi f_m A_m$$

So that $A_m < \frac{E_o f_s}{2\pi F_m}$

Or $$E_o \geq \frac{2\pi f_m A_m}{f_s} \qquad 8.18a$$

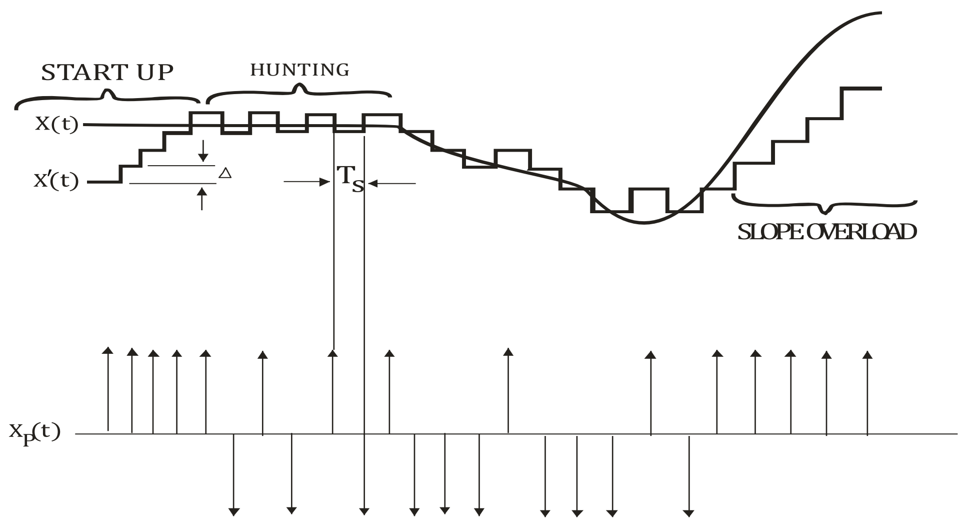


Figure 8.20 Delta Modulation Waveform

Except for overload $x'(t)$reasonably approximately $x(t)$ - especially if E_o and T_s are small and low-pass filtering at the demodulator further improves the approximation. But observe that $x(t)$is not the transmitted signal. Rather, the transmitted signal is a binary representation of $x_p(t)$ and the binary digits merely indicate the polarity of the difference between $\mathrm{x(t)}$, $x_p(\mathrm{t})$ and $\mathrm{x'(t)}$ as t = KT_s; that is why it is given the name Delta Modulation.

To study the performance of DM, we first derive a condition for preventing slope overload with lone modulation, $x(t) = \mathrm{A_m cos2\pi f_m t}$. The maximum message slope then is

$$[\frac{dx(t)}{\mathrm{dt}}]_{\max} = 2\pi \mathrm{f_m A_m} < 2\pi\omega \qquad 8.18b$$

Where the upper bound comes from our message conventions $A_m < 1$ *and* $f_m \leq \omega$. Now the maximum slope of $x'(t)$ is $E_o/T_s = E_o f_s$ so a sufficient condition for no slope overload is $\mathrm{E_o f_s} \to 2\omega$ and therefore $\mathrm{f_s} \gg 2\omega$ if $\mathrm{E_o} \ll 1$ the latter bring required to make $x'(t)$ a good approximation to $x(t)$.

Eq (8.18a) is overly conservative unless the message spectrum is flat over w. More typical message fall off well below $f = \omega$ and condition can be relaxed. Thus if there is some frequency $f_o < \omega$ such that

$$\mathrm{G_o(f)} \leq (\frac{\mathrm{f_o}}{\mathrm{f}})^2 \mathrm{G_o(f_o)}$$

$$\mathrm{f_o} \leq |\mathrm{f}| \leq \omega \qquad 8.19$$

$$\text{Then } \left|\frac{\mathrm{dx(t)}}{\mathrm{dt}}\right| \leq 2\pi \mathrm{f_o} \text{ and instead of Eq (8.10),} \qquad \mathrm{f_s} \leq \frac{2\pi \mathrm{f_o}}{\mathrm{E_0}} \qquad 8.20$$

Voice signals generally satisfies equation (8.19) with $\mathrm{f_o} \cong 800$ Hz as composed to $\omega \cong 4$ kHz, so Eq (8.20) represents significant reduction of the sampling frequency and the transmission bandwidth $\mathrm{B_T} \geq \mathrm{f}/2$.

Now as far as quantizing noise, we write

$$\mathrm{x_{(t)}} = x'(t) + \varepsilon(t) \qquad 8.21a$$

Where from Eq (8.13a) $|\varepsilon_{(t)}| = |x'(t)| - x_{(t)} \leq \mathrm{E_o}$ 8.21b

in absence of slope overload. Assuming that $\varepsilon(\mathrm{t})$ has a uniform distribution, the mean square error equals $\varepsilon^2 = \frac{\mathrm{E_o}}{3}$. We cannot, however, take ε^2 as the output quantization noise No because the LPF in Eq (8.21b) operates on the stepwise signal $x'(t)$, rather than reconstructing quantized sample value as in PCM. To determine M_D, let us make the reasonable assumption that $G_E(f)$ is essentially constant, over

$|\mathrm{f}| \leq \omega$ and $\mathrm{G_\varepsilon(o)} \cong \varepsilon^2 \mathrm{T_s} = \frac{\varepsilon^2}{\mathrm{f_s}}$ Therefore

$$\mathrm{N_D} = \int_{-\omega}^{\omega} \mathrm{G_\varepsilon(f)df} \cong \frac{\omega\varepsilon^2}{\mathrm{f_s}} = \mathrm{E_0^2}\frac{\omega}{3\mathrm{f_s}} \qquad 8.22a$$

and

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)_{\mathrm{D}} = \left(\frac{3\mathrm{f}}{\mathrm{E_0^2}\omega}\right)\overline{\mathrm{x^2}}$$

Now, using Eq (8.22) to eliminate E_o and inserting the bandwidth ratio

$\beta = \frac{\mathrm{B_T}}{\omega} \geq \frac{\mathrm{f_s}}{2\omega}$, we have

$$\left(\frac{S}{N}\right)_{Dmax} = \frac{6}{\pi^2}\left(\frac{\omega}{f_o}\right)^2 \beta^3\overline{x^2} \qquad 8.22b$$

This includes the case of Eq (8.18) by letting $f_o = \omega$.

From Eq (8.22), we see that the performance of DM falls somewhere between PCM and PPM. Since the wideband noise reduction goes at β^3. Like PCM, the transmitted signal is digital so regenerative repeater are allowed, and the terminal equipment is much less complex than PCM.

Drawbacks of delta modulation, as already stated DM suffers from the following drawbacks and we shall see how they can be overcome in practice.

Quantization noise, which can be reduced by increasing the PRF (abbreviation of pulse repetition frequency) normally designated as f_g.

Slope overload distortion, which decreases with the increase in step size (tht is the pulse amplitude E_o) and increasing f_s.

Idling noise, which can be reduce by reducing E_o.

Thus we see that increase in f_s required increase in transmission bandwidth. However, the requirement of E_o are conflicting as far as overload distortion and noise are concerned; and the only solution to this problem is to adjust E_o.

8.10 Adoptive-Delta Modulation

A delta-modulation which adjust its step size E_o to overcome its drawbacks (discussed in the preceding article) is called adoptive delta modulation normally abbreviated as ADM. It adopts itself to the changing signal conditions. Even then the transmitted pulses are of constant amplitudes, here only the polarity changes, exactly as in the case of simple delta modulation.

A block diagram of an ADM transmitter is shown in Fig. 8.21(a). Compared to DM adoptive delta modulation transmitter has a variable gain amplifier put between transmitter output (the pulse modulator output) and the feedback integrator network.

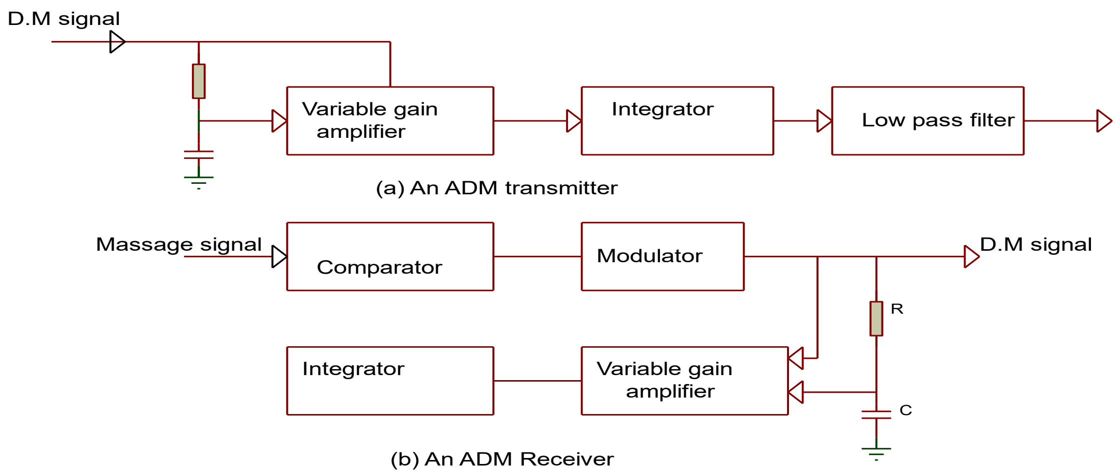


Fig 8.21 Adoptive Delta Modulation System

The gain of this variable gain amplifier varies with the magnitude of the control input voltage. The control input voltage is obtained from the transmitted pulse sequence. The receiver block diagram shown in Fig. 8.21b and the transmitter in Fig. 8.21a. The gain, of the amplifier, follows the magnitude of the control input voltage.

As we know, the message signal is a constant or varying by a very small quantity, the DM system goes hunting and thus generates alternative positive and negative pulses. The alternating sequence will give to an almost zero at the capacitor C, so that gain control voltage will be minimum. Therefore, the amplifier gain will be minimum and the amplitude of the pulses applied to the integrator, will be minimum resulting in reduction in quantizing or idling noise. When the input signal amplitude is increasing or decreasing rapidly, the modulator output-pulses sequence will produce a large voltage across C (the polarity is now immaterial), which in turn causes the variable to give a large gain and thus feed higher amplitude pulses to the integrator in the feedback path. Therefore the output of the integrator will rise (or fall) more rapidly and will be closer to the message signal than what it would have been with the constant amplitude voltage.

The size of the steps applied to the feedback integrator is thus variable, varying according to the message which is controlled entirely by the transmitted pulse sequence. The same sequence is variable at the receiver so that the pulses going to the receiver integrator can be varied in exactly the same way as at the transmitter.

8.11 PCM Reception

The Fig. 8.22 shows the portion of a PCM receiver following carrier demodulation if any. The analog PCM waveform contaminated by random noise n(t) is operated on by an A/D converter that regenerates the digital code (+error). From these code words, the decoder determines the quantized sample values (with error of course) and generates $x_{sq(t)}$ which is processed by an LPF to yield the output analog signal $x'(t)$.

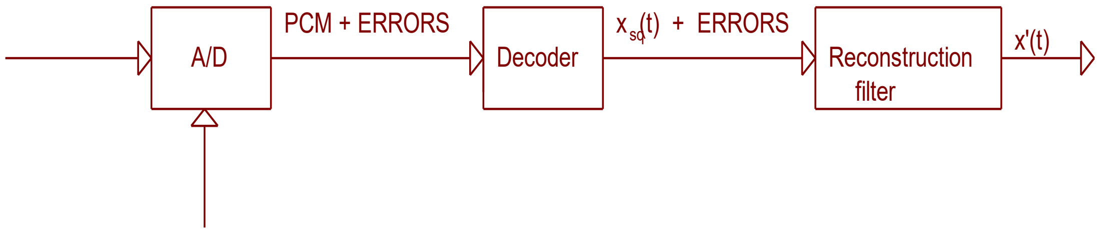


Fig 8.22 Portion of PCM Receiver

If the signal to noise ratio at the A/D converter is only modestly large, the error probability is sufficiently small that one can ignore the effects of the random noise. Despite this condition, $x'_s(t)$, not $x_s(t)$; i.e. reconstruction is based on the quantized samples rather than exact values. Furthermore, there is no way of obtaining exact values at the receiver; that information was discarded at the transmitter in the quantizing process. Thus perfect message reconstruction is impossible in PCM even when random noise has negligible influence.

As a result, the quantization effect is a basic limitation of coded system; just random noise is a limitation of conventional analog system.

8.12 Bandwidth Requirements

Since several digits are required for each message bandwidth, an estimate the bandwidth will be much greater than the message bandwidth. An estimate of the bandwidth is obtained as follows:

Quantized samples occur at a rate of $f_s \geq 2\omega$ samples per second, so that must be bit rate r = nf_s digits per second. When the PCM signal is transmitted by baseband pulses, the required minimum bandwidth termed as Nyquist bandwidth, is $1/2nf_s$ which is somewhat larger than nf_n. So we have the required channel bandwidth as

$$B = \frac{1}{2}(bit\ rate) = \frac{nf_s}{2} \geq v\omega \qquad 8.23$$

Let v be the number of digit (i.e. pulses) in the code words we require $\mu^v \geq Q$ for unique encoding. Therefore, the parameters should be so chosen such that

$$\mu^v = Q \qquad 8.24a$$

$$v = \log_\mu Q \qquad 8.24b$$

Where μ is the amplitude of the pulses, Q is the quantize level.

Therefore, the baseband PCM bandwidth is thus a minimum of equation 8.24b times the message bandwidth.

Furthermore, the pulse in a PCM system usually occurs as a uniform, rate for minimum bandwidth. If there are to be M messages multiplexed, each code group having n pulses, and sampling occurring at rate of $2f_m$, where f_m is the maximum frequency of each message, it then follows that the time between pulses is $\frac{1}{2nmf_m}$, and the minimum transmission bandwidth is given by

$$\omega = nmf_m \qquad 8.25$$

This shows a very important result, that the channel bandwidth is proportional to the number of pulses per code grouped thus is related to the accuracy with which the signal may be received.

It will be proved subsequently that $S_o/N_o \cong 2^{2n}$ and the fact that the channel bandwidth is proportional to n, we conclude that the output signal-to-noise ratio increases exponentially with bandwidth. This is a much greater improvement than observed for the other pulse system discussed in the preceding chapter.

8.13 SNR Characteristic

The signal-to-noise ratio (SNR) at the output of the decoder will be determined here for the case of large signal-to-noise ratios existing in a channel. This being the case, it is then possible to assume that the noise on the channel never (or at least very seldom) causes a pulse to be lost or misinterpreted by the receiver. The only source of noise to be considered is that due to the original quantization of the signal, for error does result in this process thereby causing the receiver output to have an error or noise component. Thus, any given required sample may be in error by as much as a half quantizing level (K/2).

Consider a PCM/AM system show in Fig. 8.23. The message is sampled at the Nyquist rate ($1/2f_m$), encoded, and transmitted.

Note that the samples to be encoded do not have any error associated with them. Quantizing error is introduced at the encoder. The samples received are not, then the same samples that were add to the encoder. The receiver samples may be separated into two example signals, one due to quantization error and one due to the true value of the samples as shown in Fig. 8.24

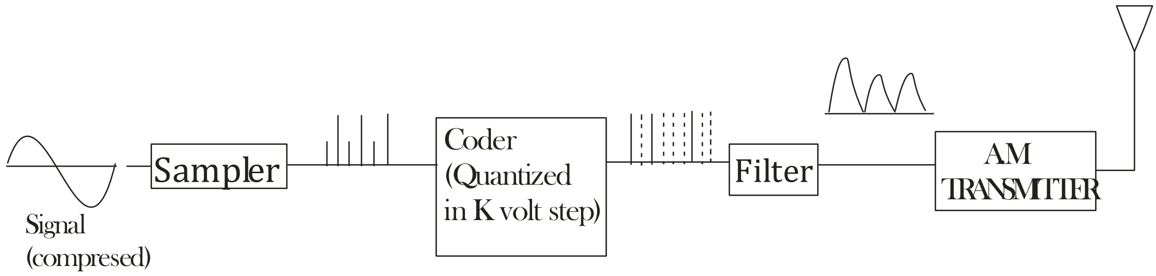


Fig 8.23a PCM Transmitter

These two waves when applied to a low-passed-filter produce two output waves. One is the desired message (identical with the input message) and the other a noise wave due to quantization.

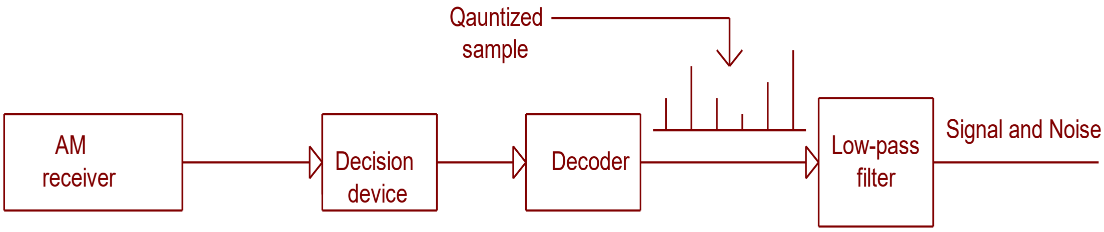


Figure 8.23b PCM Receiver

The output signal power is computed from the desired part of the output, which is also equal to the power of the input message. An approximate result may be obtained by approximating the message or output signals as show in Fig. 8.24.

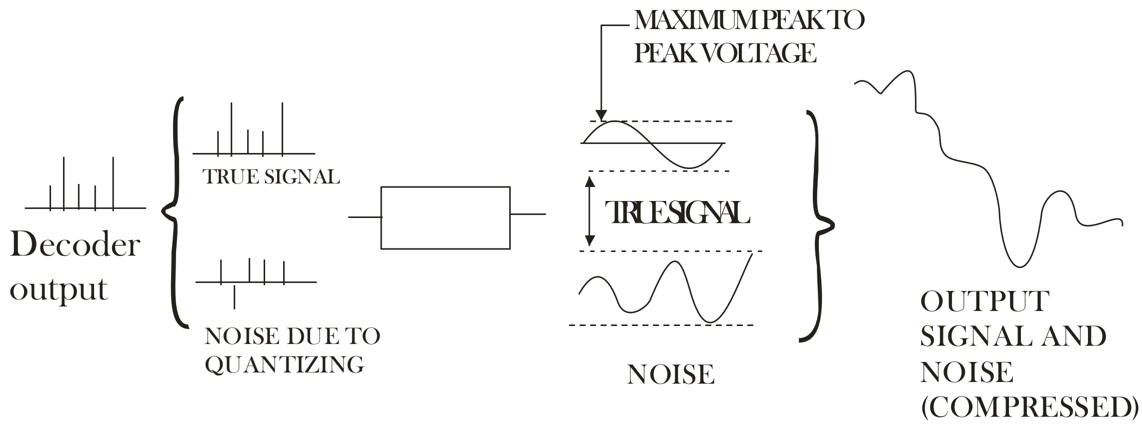


Figure 8.23c True Signal and Noise in PCM System

From this figure we see that the maximum peak-to-peak value of the signal is $(2^n-1)K/2$. Assuming the signal is properly compressed, on the 2^n see we would expect to have one each of the level ±k, ±2k, ………… $\pm(2^n-1)K/2$. The means square value over this 2^n second interval is

$$S = \frac{1}{2n}\left[1^2 + (-K^2) + (2K^2) + \cdots + \frac{(2^n-1)^2K^2}{4} + \frac{(-1^2)(2^n-1)^2K^2}{4}\right]$$

or

$$S_n = \frac{K^2}{2^n}\sum_{n=1}^{2^{n-1}} 2x^2$$

$$= \frac{K^2}{12}(2^n-1) \qquad 8.26$$

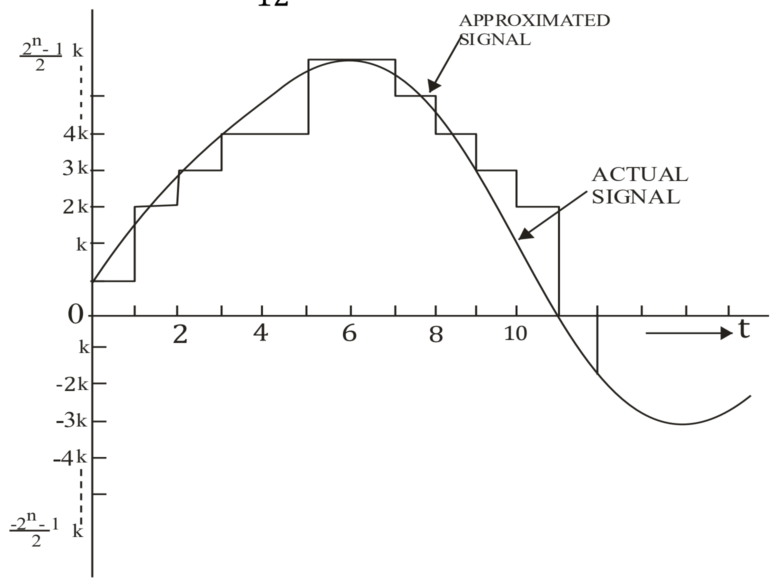


Figure 8.24 Approximated Signal

The output noise will have a maximum value of K/2. Approximating this wave in the same manner as the signal loads to a noise power of

$$N_o = \frac{K^2}{12} \qquad 8.27$$

The output signal to noise is then given by

$$\left(\frac{S_o}{N_o}\right) = 2^{2n} - 1 \qquad 8.28$$

For large signal to noise ratios in the channel (10dB or more) observing that $\left(\frac{S_o}{N_o}\right) \cong 2^{2n}$ and recalling that channel bandwidth is proportional to n, we conclude that the output signal to noise ratio increases exponentially with bandwidth. This is a much greater improvement than observed for the other pulse systems.

8.14 Advantages of PCM

Despite the inherent complexity of mechanization here are a number of marked advantages in the use of PCM.

1. PCM affords a considerable increase in the output signal to noise ratio at the receiver.
2. The code group is usually so synchronized that the time of occurrence of each pulse is known at the receiver. The only decision which needs to be made by the receiver is whether a pulse is present or not; nothing need be known concerning its amplitude width etc. This is an advantage which other pulse system do not possess.
3. PCM permits repeating or amplifying the encoded signal without significant distortion being introduced again owing primarily to the nature of the coded signal.
4. As already stated that the output signal-to-noise ratio increase exponentially with bandwidth. This is a much greater improvement than observed for other pulse system.
5. A PCM system designed for analog message transmission is readily adapted to other input signals particularly digital data, hereby promoting flexibility and system utilization.
6. By virtue of the regeneration capacity PCM, is distinctly advantageous for system having many repeater stations. Thus with respect to long distance telephone, this has been greatest assets.

Thus PCM should be given due consideration for applications involving, TDM, minimum power, a diversity of message types, (i.e. analog and digital), or many repeater stations. Because most of these factors are present in long-hand telephone transmission, PCM appears to be the great hope of the future of telephony. However, in more routine applications, the lost of the hardware for coded modulation is usually much more compared to that for analog modulation unless considerate advantage is made in integrated circuit.

8.15 Advantages of DM

The major disadvantage is that, (i.e. PCM and DM) almost noise free regeneration ability. Thus they are superior to FM system particularly when the signal is to be transmitted over long distance and the use of repeater become essential because of the distance involved. This is due to the fact that PCM and DM signal repeaters can be regenerative, while the analog signal repeaters cannot be. Let us now discuss regenerative repetition.

A repeater of a digital signal, can receive the signal like the receiver at the destination, identify the digits and generate these digits afresh. If the signal-to-noise ratio at the receiver is not too low, the identification of the received signal digits is almost noise free. Thus the regenerated digits will be almost identical with those at the first transmitter. Therefore in the successive repeater stages the regenerated signal is almost identical with that of the first transmitter.

In the case of the analog signal repeater, the signal is received and amplified to the required signal power level. The noise power will keep in accumulating at each repeater, so that the total noise, at the final receiver will be n times as large as on the first repeater (n being the number of repeater) for comparable signal power. If the signal to noise ratio, at the first repeater is S/Ns it will drop to S/N at the final receiver. The original signal to noise ratio S/N can be restored by increasing the signal. This improvement will be much larger that what is in the case of regenerative repeater.

8.16 Comparison of PCM and DM

After having discussed the advantages of PCM system and superiority of PCM and DM system over other system we will now compare PCM and DM system:

1. The main advantage of DM over PCM is the extreme simplicity of the transmitter and receiver circuits.
2. A DM system, with the same bit transmission rate gives better performance than a PCM system.
3. For speech transmission, DM system gives better performance at lower bit rate, (i.e. 40Kleps and below). At higher bit rates the PCM system gives better performances.
4. DM generally requires large transmission bandwidth than PCM.

8.17 PSK, FSK and DPSK

When it becomes necessary to superimpose a binary PCM waveform on a carrier, than PM and FM are commonly utilized. Because of the special (two level) nature of the carrier modulating signal, phase modulation is termed as Phase-Shift Keying (normally abbreviated as PSK), and frequency modulation is called Frequency Shift Keying normally abbreviated as FSK. Let us first discuss PSK.

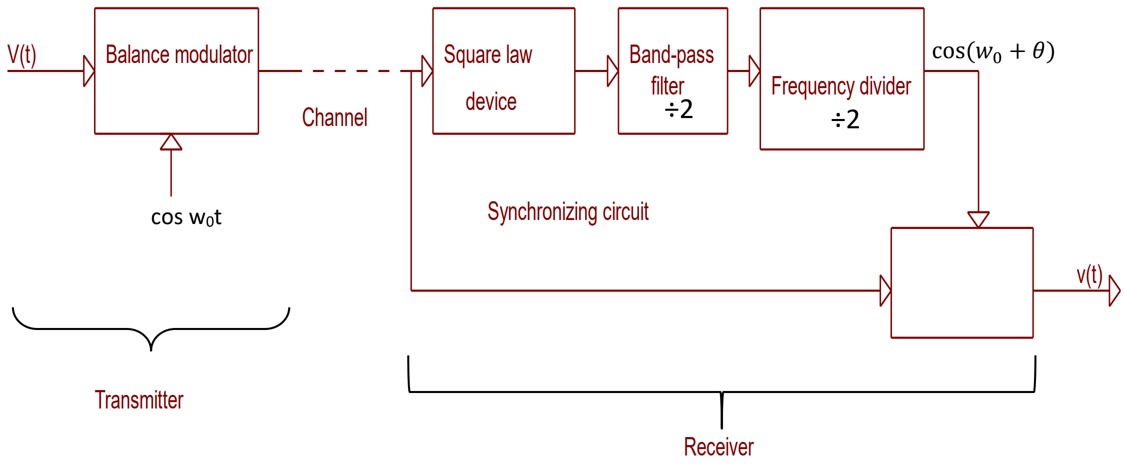


Figure 8.25 A Binary PSK System

i PSK: For simple phase shift keying, the un-shifted carrier $\mathrm{Vcos}\omega_0 t$ is 180°, or π radians $[\mathrm{Vcos}(\omega_0 t + \pi) = -\mathrm{Vcos}\omega_0 t]$, is transmitted to indicate 0 condition. The modulating circuitry is quite simple for this, since it is only necessary to provide two switches, an inverter. Demodulation is achieved by subtracting the received carrier from a derived synchronous reference carrier of constant pulse.

Consider that a binary signal $V_{(t)}$, which takes on the values V(t) = ±V, is to be the modulating waveform in PSK system. The PSK waveform is

$$v_{p(t)} = A\cos[\omega_o t + \emptyset] \qquad 8.29a$$

Where A is a fixed amplitude, $\emptyset = 0$ for $\mathrm{V(t)} = +\mathrm{V}$ and $\emptyset = \pi$ for $\mathrm{V(t)} = -\mathrm{V}$

Alternatively the Eq (8.29) can be written as

$$v_{\mathrm{p(t)}} = \frac{\mathrm{V(t)}}{\mathrm{V}} \mathrm{Acos}\omega_o t \qquad 8.29b$$

So that $v_{\mathrm{p(t)}} = \mathrm{Acos}\omega_o t$ for $v_{(\mathrm{t})} = +\mathrm{V}$, and

$v_{\mathrm{p(t)}} = -\mathrm{Acos}\omega_o t$ for $v_{(\mathrm{t})} = -\mathrm{V}$

The waveform of Eq (8.29) can be generated by applying waveform $v_{(t)}$ and the carrier $\cos\omega_o t$ to a balance modulator as shown in Fig. 8.25. The received signal will have the form

$$v_{\mathrm{p(t)}} = \frac{v_{(\mathrm{t})}}{\mathrm{V}} \mathrm{Acos}(\omega_o t + \theta)$$

Where θ is a phase angle which depends on the effective length of path between transmitter and receiver, detection, (i.e. demodulation) is performed simultaneously; hence we require the waveform $\cos(\omega_o t + \theta)$ at the receiver. A synchronizing circuit which can extract the waveform $\cos(\omega_o t + \theta)$ from the received signal itself is shown in Fig. 8.25.

The output of the square law device is $\mathrm{A}^2\cos^2(\omega_o t + \theta)$ since $\frac{\mathrm{v}^2(t)}{\mathrm{V}^2}$ is always +1.

But $\mathrm{A}^2\cos^2(\omega_o t + \theta) = \frac{\mathrm{A}^2}{2}[1 + \cos 2(\omega_o t + \theta)] = \frac{A^2}{2} + \frac{\mathrm{A}^2}{2}\cos 2(\omega_o t + \theta)$

Hence a band-pass filter is required to separate $\cos 2(\omega_o t + \theta)$.The frequency divider divides the frequency by 2, yielding $\cos(\omega_o t + \theta)$ as required.

At the synchronous demodulator, the signal of Eq (8.29) is multiplied by the locally recorded carrier, $\cos(\omega_o t + \theta)$. This product is

$$\cos(\omega_o t + \theta)V_{p(t)} = \frac{1}{2}\frac{v_{(t)}A}{V} + \frac{1}{2}\frac{v_{(t)}}{V}\cos 2(\omega_o t + \theta) \qquad 8.30$$

In which we are interested in $V_{(t)}$. If $V_{(t)}$ were band limited signal, then we might recover $V_{(t)}$ precisely through the use of a low pass filter. However, waveform is not band limited (at least in principle) because of the abrupt transitions in its waveform. Hence a low pass filter will introduce some distortion in v(t) and will also pass some of the sidebands of the double frequency carrier in Eq (8.30).

It may be mentioned here that we are not really interested in recovering $V_{(t)}$ but only in knowing whether $V_{(e)} = \pm V$ in each bit interval. If a bit interval extends over many cycles of the carrier $\cos\omega_o t$ then it will be easy to find a low-pass filter which will effect an adequate separation of the terms in Eq (8.30) to follow such determination.

ii. FSK: Two-ton frequency-shift keying is commonly used. In this system, two tones within the voice channel, which can be separated easily with band-pass filters, are assigned as carriers. One tone is transmitted for a bit, and the other for a 0 bit. Two band-pass filters, rectifiers and a differential amplifier demodulate the signal. Tele-type carrier systems need on the telephone system divide the 4 KHz voice band up into as many as 20 channels, with a channel separation of about 120 Hz for 60 r.p.m or as few as six 170 Hz channel for 100 r.p.m, eight level teletypes. Radio-telegraph systems quite often frequency modulate the RF carrier directly, transmitting the normal carrier frequency for the bit condition and transmitting a frequency about 70Hz lower for the 1bit. Alternatively, the two-tone FSK system can be either amplitude modulated or frequency modulated.

In frequency shift-keying, the binary signal is used to generate a waveform

$$v_f(t) = A\cos(\omega_c \pm \emptyset)t \qquad 8.31$$

In which + sign and – sign, depends on whether the bit is a 1 or a 0. Then the transmitted signal is of amplitude A and has an angular frequency $\omega_o + \emptyset$ or $\omega_o - \emptyset$ a constant angular deviation from the carrier frequency ω_c.

Such an FSK signal might be demodulated by applying the signal simultaneously to two sharply turned filters, one tuned to $\omega_c + \emptyset$ and the other tuned to $\omega_c - \emptyset$. We would then determine that 1 or a 0 had been transmitted in any bit interval depending on which filter yielding the larger output signal. However, a synchronous demodulator has the advantage that, it may be easily adopted to yield optimum performance in the presence of noise. As shown in Fig. 8.26, it requires two local carriers at angular frequencies $\omega_c + \emptyset$. If the received signal is $A\cos(\omega_c + \emptyset)t$, then the output of the difference amplifier will be

$$V_d = \frac{A}{2} - \frac{A}{2}[\cos 2\emptyset t + \cos 2\omega_c t - \cos 2(\omega_c + \emptyset)t] \qquad 8.32$$

If the receiver signal is $A\cos(\omega_c - \emptyset)t$ the output of the difference amplifier will be:

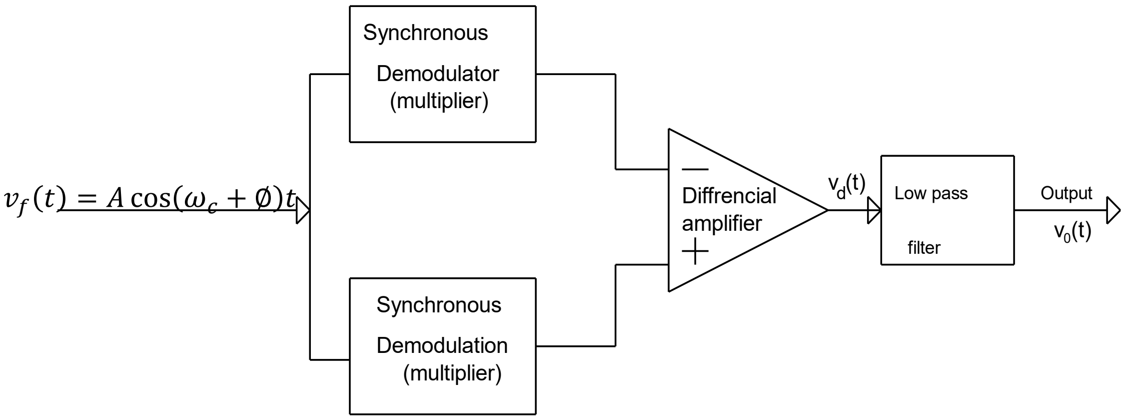


Figure 8.26 The FSK System

$$V_d = -\frac{A}{2} + \frac{A}{2}[\cos 2\emptyset t + \cos 2\omega_c t - \cos 2(\omega_c t - \emptyset)] \qquad 8.33$$

Similar to PSK system a low-pass filter is used to have adequate separation of D.C components in Eqs (8.32) and (8.33) to permit a determination of whether a 1 or a 0 has been transmitted. Ordinarily $\omega_c \gg \emptyset$; hence the lowest frequency, by far in the brackets in Eq (8.32) is the term $\cos 2\emptyset t$. To permit an easy separation of D.C components, it is therefore necessary that the bit interval extends over a time T such that, in that interval, $\cos 2\emptyset t$ include many cycles. Hence we need that

$$2\emptyset T \gg 2\pi$$

iii. DPSK: Phase keying receives a local oscillator at the receiver which is accurately synchronized in phase with the unmodulated transmitted carrier and in phase by π-radians can be difficult to achieve. Differential phase shift keying (normally abbreviated as DPSK) overcomes the difficulty by utilizing the phase shift between successive bits, hence the name differential PSK.

The phase of the first message bit has to be compared to a reference bit, which may be arbitrarily chosen as 0 or 1. Fig. 8.27 shows a 0 reference bit. If the compared bits do not differ, that is encoded as a 1 or zero phase shift or carrier. If they do differ, this is encoded as a 0 or π radians shift or carrier. The figure shows the sequence events for the message 01100. Comparing the first message bit with the reference, they are seen to be, so this is encoded as a 1, and transmitted as zero phase-shift. Note that the first transmitted bit is the reference bit, a zero in this code.

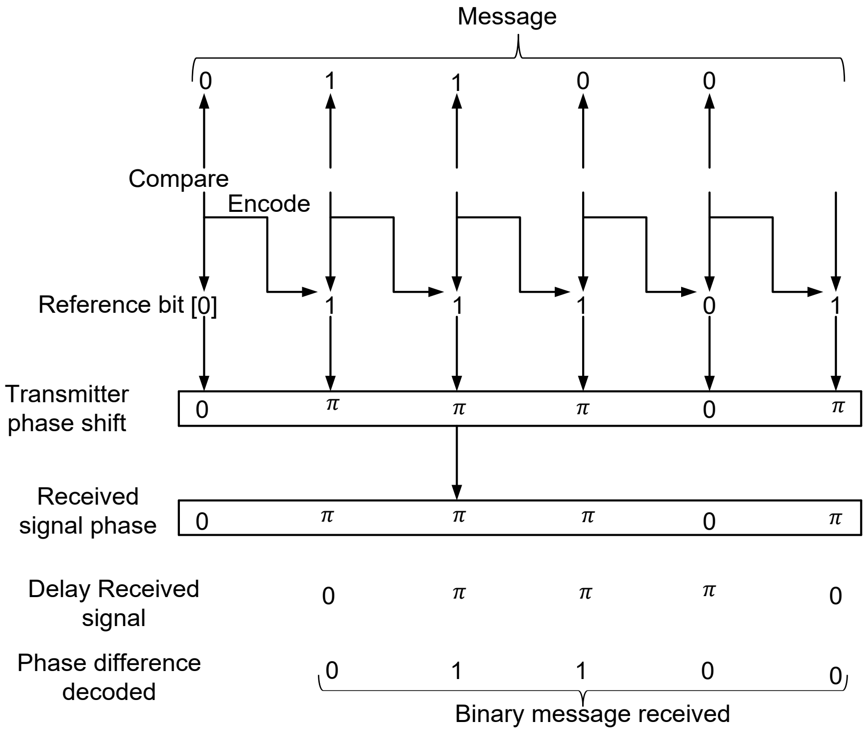


Figure 8.27 Encoding and Decoding a Message in DPSK

The second message bit is compressed with the second encoded bit both are seen to be 1, so this information is also encoded as a 1, and the carrier phase shift remains at zero. Continuing in this way, the third message is seen to be the same as the third encoded bit, so again this is encoded as a 1, or zero phase shift of carrier. The fourth message bit is compared with the fourth encoded bit. They are seen to be different, so this is encoded as a 0, and the carrier phase shift is π radians. The last message bit is compared with the fifth encoded bit, they are the same, so this is encoded a 1, or again, a carrier phase shift of zero. Fig. 8.28 shows how encoding might be achieved by digital hardware.

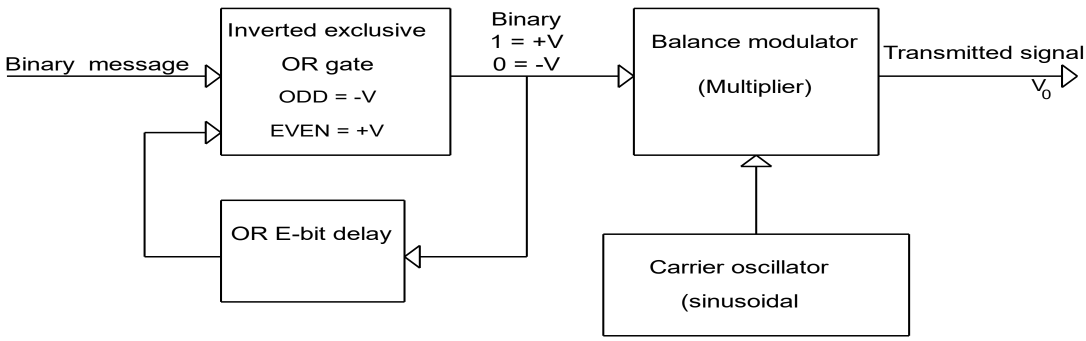


Figure 8.28 An Encoding Arrangement

At the receiver, the signal is multiplied by a one bit delayed version of itself. This is shown schematically in Fig. 8.28. If the delayed signal is in phase with the direct signal, the output following the low-pass filter will have a positive D.C component and will be decoded as a 1.

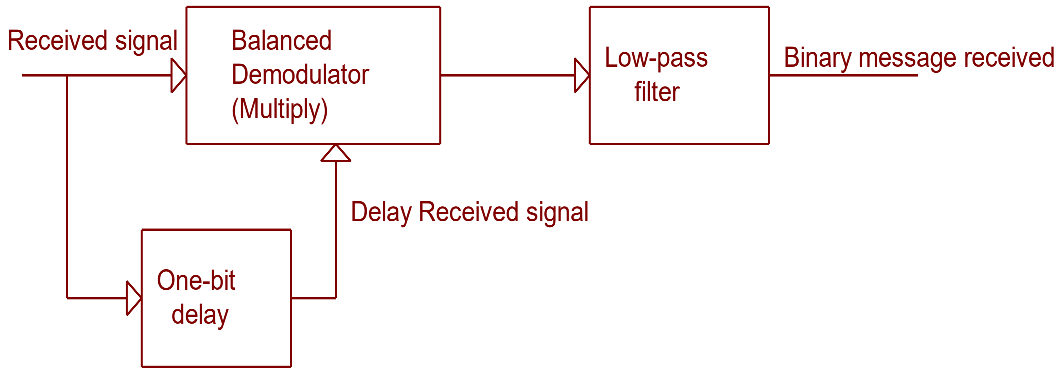


Figure 8.29 The DPSK Receiver

If the two signals differ in phase by π radians, the D.C output will be negative interpreted as a binary 0. The sequence of events at the receiver is also illustrated in Fig. 8.29.

8.18 The Complete PCM System

We have already discussed the PCM system in piece-meal, i.e. first sampling the quantizing leading to quantizing error or noise followed by its generation and reception and finally its bandwidth and SNR characteristics. We will now study the entire PCM system in compact form which will be just putting together what we have already discussed in the preceding articles.

i. **Encoder**: The block diagram of a PCM communication system is drawn in Fig. 8.30. In this analog signal $m_{(t)}$ is sampled, and these samples are quantized. The quantized samples are applied to encoder. The encoder responds to each such sample by the generation of a unique and identifiable binary pulse (or binary level) pattern. At the receiver, we must be able to identify the level from the pulse pattern. Thus it is clear that not only does the encoder numbered the level, it also assigns to its identification code.

As shown in Fig. 8.30, the combination of quantizer and encoder is called an analog-digital converter normally abbreviated as A/D converter. Thus A/D converter accepts an analog signal and replaces it with a succession of code symbols, each symbol consisting of a train of pulses in which each pulse may be interpreted as the representation of a digit in an arithmetic system. Therefore, the signal transmitted over the communication channel in PCM system is termed as digitally encoded signal.

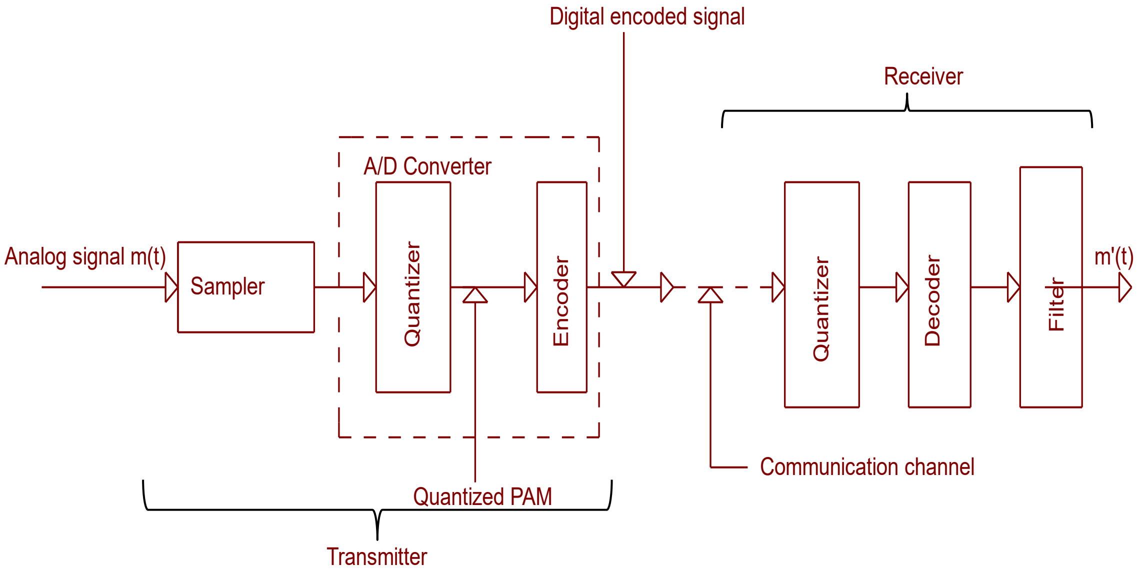


Figure 8.30 The PCM Communication System

ii. **Decoder:** The first step when the digitally encoded signal arrives at the receiver (or repeater) is to separate the signal from noise (which has been picked up during the transmission along the channel). As we know such a separation is possible because of the quantization of the signal, and hence process of re-quantization is done. Accordingly the first block in the receiver is quantizer. For each pulse interval this quantizer has to make the simple decision of whether a pulse has or has not been received or which of the two voltage levels has occurred.

The receiver quantizer then, in each pulse slot, makes a guess about whether a positive pulse or a negative pulse was received and transmits its decision in the form of a reconstituted or regenerated pulse train to the decoder. If repeater is to be used, the generated pulse train is simply raised in level and sent along next section of the transmission channel. The decoder does inverse operation of the encoder and is called digital-to-analog converter normally abbreviated as D/A converter.

The decoder output is the sequence of quantized multilevel sample pulses. The quantized PAM signal is thus reconstructed, which is filtered to reject any frequency components lying outside the base band. The final output signal $m^1_{(t)}$ is identical to input $m_{(t)}$ except for quantization noise and the occasional error in Yes/No decision making at the receiver. If companding is to be inclined, the compressor precedes the sampler and the expander follows the filter. Bit (pulse) synchronizing has also not been included in the PCM system shown in Fig. 8.30. The receiver must be given timing information identifying the beginning and end of a pulse time slot.

Furthermore, if a number of a band signals are being multiplexed, frame synchronization information must also be transmitted.

Example 8.3

The signal to quantizing noise ratio of a binary PCM system is required to be at least 1000. Determine the required minimum number of binary digits to represent the quantizing level. (Grad. IETE Nov.1981)

Solution: It is given that

$$\frac{S}{N_q} = 2Q^2 \geq 1000$$

$$Q^2 \geq 500$$

$$Q \geq 22.4 \geq 23$$

Since Q has to be integer.

This is the binary encoding, quantization level Q can be chosen to be 32 requiring n to be 5.

Example 8.4

A typical PCM system, sampling at 8000 samples per seconds; uses 6bits/word for transmission. Determine the Niquest bandwidth and obtain S/N_q (AIMES 1982)

Solution: It is given that

F_s = 8000 and n = 6

Therefore, bit rate = nf_s

= 6 x 8000 = 48 kbps

(Kb is an abbreviation of Kilobits)

Hence Niquest Bandwidth = $\frac{1}{2}(\text{bit rates})$

$$= \frac{1}{2} \times 48\ \text{kHz} = \mathbf{24\ kHz}$$

From Eq (8.28), S/N_q is given by

$$\frac{S}{N_q} = 2Q^2 = 2(2^n)^2, \text{Since } Q = 2^n$$

$$\frac{S}{N_q} = 2 \times 2^{2n}$$

But $\beta = \frac{nf_s}{2}$

Hence $\frac{S}{N_q} = 2 \times 2^{(\frac{4\beta}{fs})}$

$$= 2 \times 2^{(4 \times \frac{24}{8})}$$

$$= 2 \times 2^{12} = 2^{13} = \mathbf{8192\ or\ 39\ dB}$$

Example 8.5

The signal to quantization noise ratio is not to be less than 1000. The sampling ratio is 800 samples per seconds and the channel bandwidth is limited to 1.5 kHz. Give suitable encoding parameter.

$$\frac{S}{N_q} = 2Q^2 \geq 1000 \text{ and } f_s = 8000 \text{ samples/s}$$

$Q^2 \geq 500$ or $Q \geq 23$

Since Q has to be an integer

Also $\beta = \frac{1}{2} nf_s = \frac{800}{2}\, n \leq 15 \times 10^3$

Because channel bandwidth is limited to 15 sHz

Therefore, $n \geq \frac{15000}{4000} = 3.75$

Since n is to be an integer

$m^3 \geq 23$

or $m \geq 3$

Hence we can use tertiary or quaternary system. Although a ternary system will suffice, it is desirable to have quaternary system in practice.

Example 8.6

A sinusoidal system of 1volt amplitude and 800 Hz frequency is transmitted by delta modulation. Determine the minimum step-size so that the overload distortion is avoided, if the sampling rate is 40,000 sample/sec. (Roorkee University 1980-81)

Solution

Using notion A_m = 1volt, f_m = 800 and f_s = 40,000

Hence, minimum step size = $\frac{2\pi \times 800 \times 1}{40000} = 0.1256 \text{ V} = \frac{\mathbf{1}}{\mathbf{8}}\ \mathbf{v}$

This shows that the approximated signal will have about 2 V i.e. just 18 different levels.

8.19 Chapter Review Problems

8.1 Are the following systems digital: (a) Pulse Width Modulation (PWM), (b) Pulse Position Modulation (PPM)? Why?

8.2 Which of the following is susceptible to quantization noise:
(a) Amplitude Modulation (AM), (b) Frequency Modulation (FM), and (c) Pulse-Code
Modulation (PCM)? Which of the above is most noise resistance and why?

8.3 (a) State the theorem which provides the guideline to choose the sampling frequency for a band-limited signal. (b) What is aliasing? (c) What is the purpose of using a low pass filter prior to sampling an analog signal?

What is Chirp Signal?

Complex signal expressed in the form

$x(t)$= exp $(j0.5at^2)$, $a > 0$

Whose magnitude is constant and the frequency varies linearly in time

8.4 Which signal is called a box, boxcar or pulse signal?

8.5 What is the common name of the signal represented by delta function $\delta(t)$ or the Dirac distribution? If the input to a system is the delta function, what is the output called and by
what symbol is it denoted?

8.6 Assuming a safety margin of 10 per cent, determine the Nyquist rate for correctly sampling the electrocardiogram (ECG) signal whose maximum useful frequency content extends up to 100 Hz.

8.7 Define and explain the analog modulation system? What is pulse modulation? Explain its advantage over CW modulation. Discuss the applications of pulse modulation.Enumerate the type or pulse modulation. Describe PDM system in detail.State and explain sampling theorem in time domain. Prove that if a signal whose highest frequency is W Hz has been sampled at rate of 2 W samples per seconds, the sampled signal may be reconstructed by passing the impulse train through an ideal low pass filter whose cut off frequency is W Hz.

8.8 Discuss a typical communication system using pulse amplitude modulation with special reference to its bandwidth and signal two-noise ratio requirement.

8.9 Explain the similarity between PPM and PDM.
Discuss the SNR characteristic of PAM and their compare it with PDM.

8.10 Write short notes on:

i. Advantages of pulse modulation.
ii. $PPM\ and\ PDM$.
iii. PAM.
iv. SNR Characteristic PPM.
v. Bandwidth requirement of PPM.

8.11 Discuss clearly the principle of $PPM\ |AM$ transmission and reception.

8.12 Explain with diagrams, low PPM waveform. Show the spectrum of a PPM waveform. Estimate the channel bandwidth requirement.

8.13 Write short notes on the following:

i. Detection of PDM signals.
ii. Pulse time modulation comparison of pulse modulation systems.
iii. SNR improvements in PPM. $(AMIE\ \ W\ \ 1980)$
iv. Demodulation of PPM. $(Grad.ITETE\ Dec.1983)$
v. Slicer

8.14. Explain with diagrams, how PDM signals are generated and the modulated signal is recovered from a PDM wave. Show the spectrum waveform

8.15 Explain with suitable diagrams how a PPM signal can be converted to PPM signals. Draw the waveforms of the various stages. $(Banaras\ University$ 1982)

8.16 Explain with suitable circuit diagram the generation of PPM signal and explain how these signals are demodulated. $(Rookee\ University$, 1984.85)

8.17 Carefully plot the spectrum of flat-top sampled (PAM) which has a KHz sine modulation signal, a sampling frequency of 8 KHz and a pulse width of 31 25 μs, up to 6th harmonies of

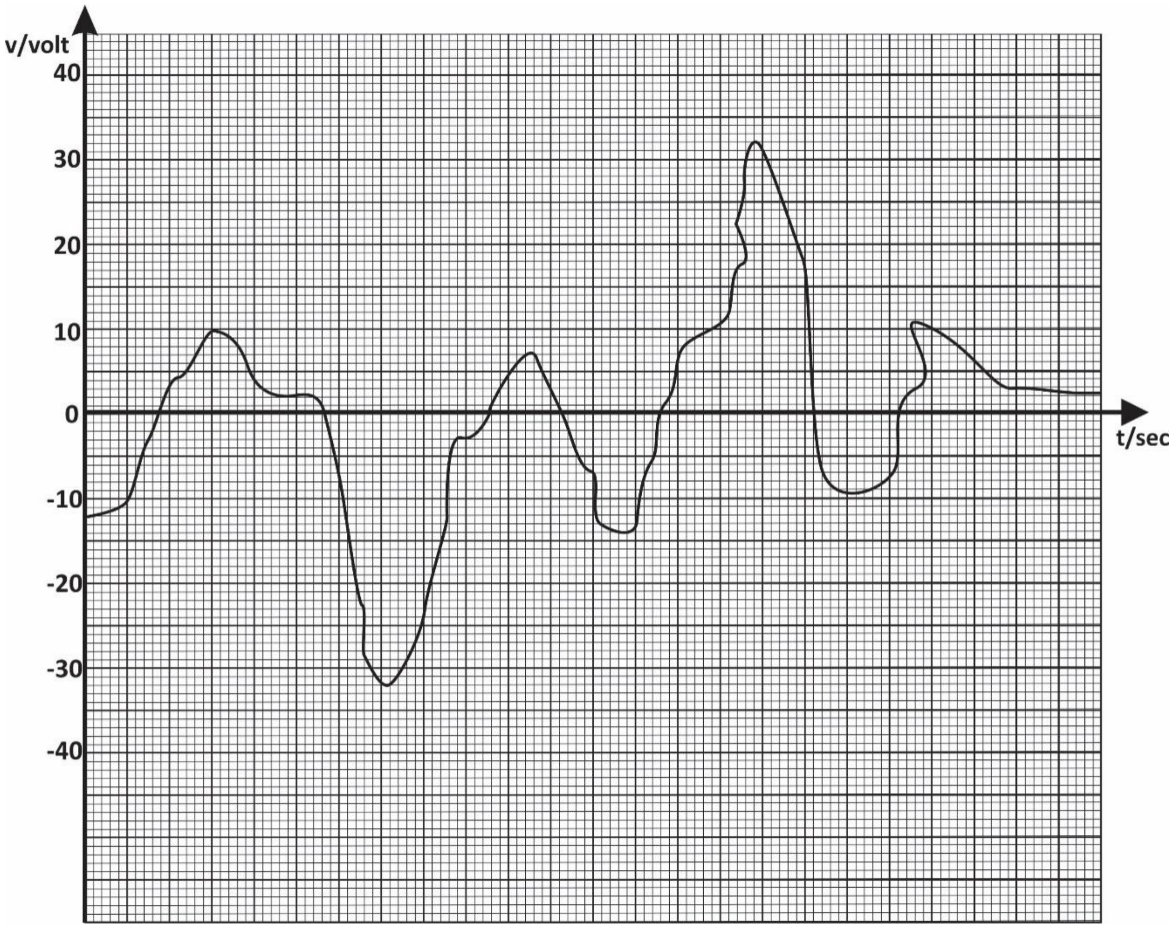


Fig 8Q1

8.18 (a) Explain briefly each of the following

(i) Pulse Modulation

(ii) Pulse Amplitude Modulation

(iii) Pulse Width Modulation

(b) Explain briefly

(i) Pulse Code Modulation

(ii) Quantization

(iii) Coding

(iv) Companding

(c) (i) State Nyquist Sampling Theorem

(ii) The maximum frequency of baseband signal is f_{max}. The baseband signal is pulse code modulated with n-bit per sample. What is the required minimum pulse code modulation bandwidth.

(iii) State the advantage and disadvantage of pulse code modulation.

(d) The signal-to-quantizing noise ratio of a binary PCM system is required to be at least 8000. Determine the required minimum number of binary digit to represent the quantizing level.

(e) Suppose the sinusoidal signal in fig(8Q1) is sampled at time t = 0, 5, 10, ………..60 s. The maximum amplitude of the signal is 32 V.

(i) Draw the PAM signal using single line

(ii) Draw the PCM signal. Using three bits per sample. Prepare a table to show the possible voltage range, quantization levels and corresponding PCM Codes. Determine the signal-to-quantizing noise ratio in dB.

8.19 (a) Briefly explain TDM. Why is it use in digital communication system? '***Note that sketches will be required for explanation'.***

(b) i. Explain with the aid of a diagrams, the concept of pre-emphases and de-emphases.

ii. Explain with the aid of diagrams and wave form, the action of a semi-conductor diode as a simple detector.

8.20 (a) (i) Draw the block diagrams of a Delta Modulation System showing the modulator and demodulator.

(ii) Explain briefly Delta Modulation.

(iii) With the aid of a diagram of waveform, explain "Start Up", "Hunting" and Slope overloads.

(iv) Compare Delta Modulation with Pulse Code Modulation.

(b) A sinusoidal signal of 4 Volt amplitude and 2000 Hz frequency is transmitted by delta modulation. Determine the minimum step size so that the overload distortion is avoided, if the sampling rate is 50,000 samples/second.

(c) (i) What is PSK?

(ii) Explain the PSK modulation process.

(iii) Draw the block diagram of a PSK system.

(University of Ibadan, TEL512-Communication system II 2010/2011 BSc degree Exam)

8.21 (a) Explain briefly

(i) Pulse Modulation

(ii) Pulse Amplitude Modulation and

(iii) Pulse Width Modulation

(b) Explain briefly

(i) Pulse Code Modulation.

(ii) Quantization

(iii) Coding

(iv) Companding

(c) (i) State Nyquist Sampling Theorem

(ii) The maximum frequency of baseband signal is f_{max}. The baseband signal is pulse code modulated with n-bit per sample. What is the required minimum pulse code modulation bandwidth.

(iii) State the advantage and disadvantage of pulse code modulation.

(d) The signal-to-quantizing noise ratio of a binary PCM system is required to be at least 4000.

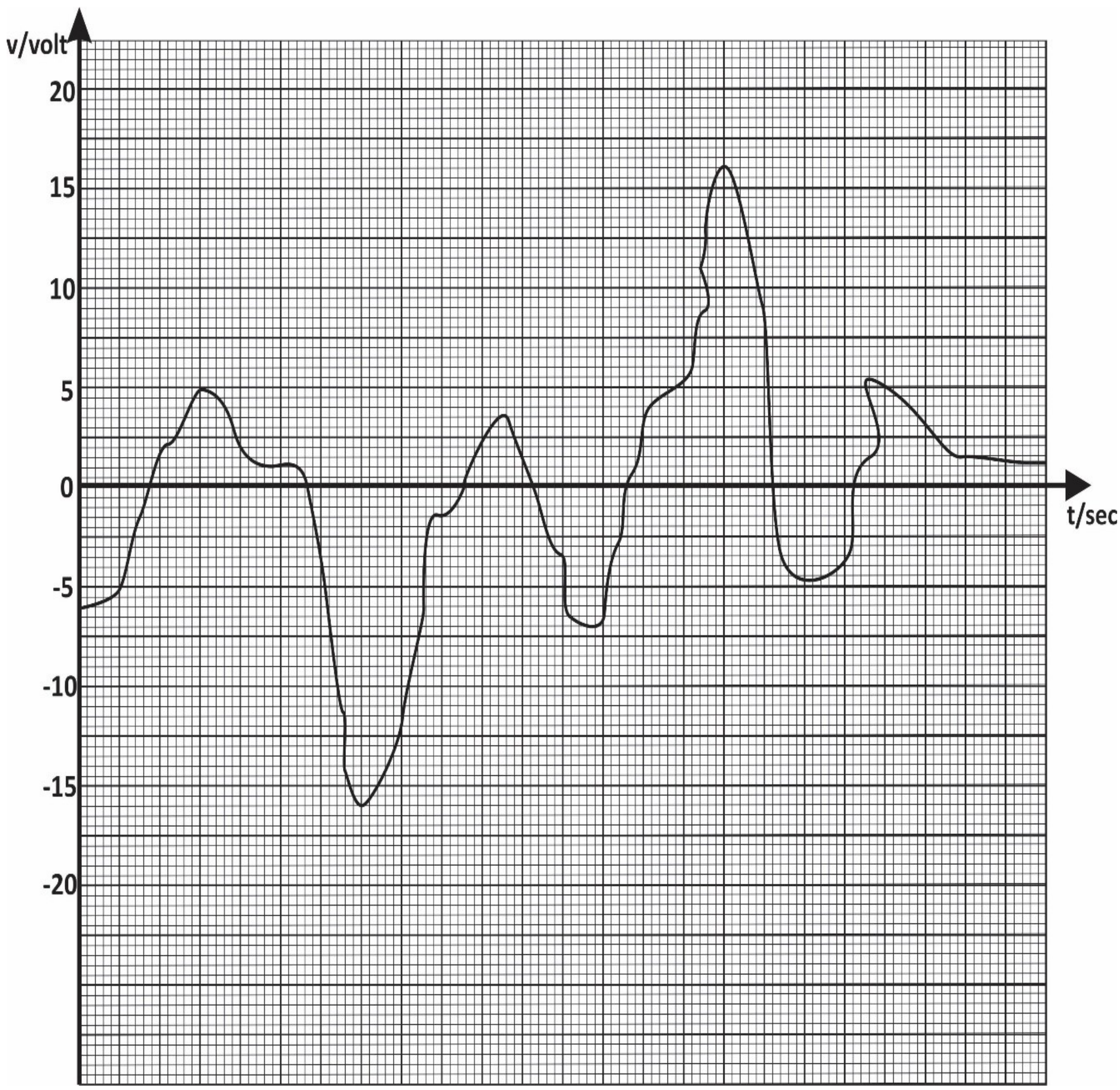


Fig 8Q2

Determine the required minimum number of binary digits to represent the quantizing level.

(e) Suppose the sinusoidal signal in Figure 8Q2 is sampled at time t = 0, 5, 10, ………..60 s. The maximum amplitude of the signal is 16 V.

(i) Draw the PAM signal.

(ii) Draw the PCM signal. Using three bits per sample. Draw a table to show the possible voltage ranges, quantization levels and corresponding PCM Codes.

(iii) Determine the signal-to-quantization noise ratio in dB.

(University of Ibadan, TEL512-Communication system II 2010/2011 BSc degree Exam)

8.22 (a) Explain briefly

(i) Pulse Modulation

(ii) Pulse Amplitude Modulation and

(iii) Pulse Width Modulation

(b) Explain briefly

(i) Pulse Code Modulation.

(ii) Quantization

(iii) Coding

(iv) Companding

(c) (i) State Nyquist Sampling Theorem

(ii) The maximum frequency of baseband signal is f_{max}. The baseband signal is pulse code modulated with n-bit per sample. What is the required minimum pulse code modulation bandwidth.

(iii) State the advantage and disadvantage of pulse code modulation.

(d) The signal-to-quantizing noise ratio of a binary PCM system is required to be at least 4000. Determine the required minimum number of binary digits to represent the quantizing level.

(e) Suppose the sinusoidal signal in Figure 8Q3 is sampled at time t = 0, 4, 8, ………..60 s. The maximum amplitude of the signal is 16 A.

(i) Draw the PAM signal.

(ii) Draw the PCM signal. Using three bits per sample. Draw a table to show the possible voltage ranges, quantization levels and corresponding PCM Codes.

(iii) Determine the signal-to-quantization noise ratio in dB.

(University of Ibadan, TEL512-Communication system II 2011/2012 BSc degree Exam)

8.23 (a) (i) Draw the block diagrams of a Delta Modulation System showing the modulator and demodulator.

(ii) Explain briefly Delta Modulation.

(iii) With the aid of a diagram of waveforms, explain ‟Start Up‟, ‟Hunting‟ and Slope overload‟.

(iv) Compare Delta Modulation with Pulse Code Modulation.

(b) A sinusoidal signal of 3 Volt amplitude and 1200 Hz frequency is transmitted by delta modulation. Determine the minimum step size so that the overload distortion is avoided, if the sampling rate is 40,000 samples/second.

(c) (i) What is PSK?

(ii) Explain the PSK modulation process.

(iii) Draw the block diagram of a PSK system.

(University of Ibadan, TEL512-Communication system II 2010/2011 BSc degree Exam)

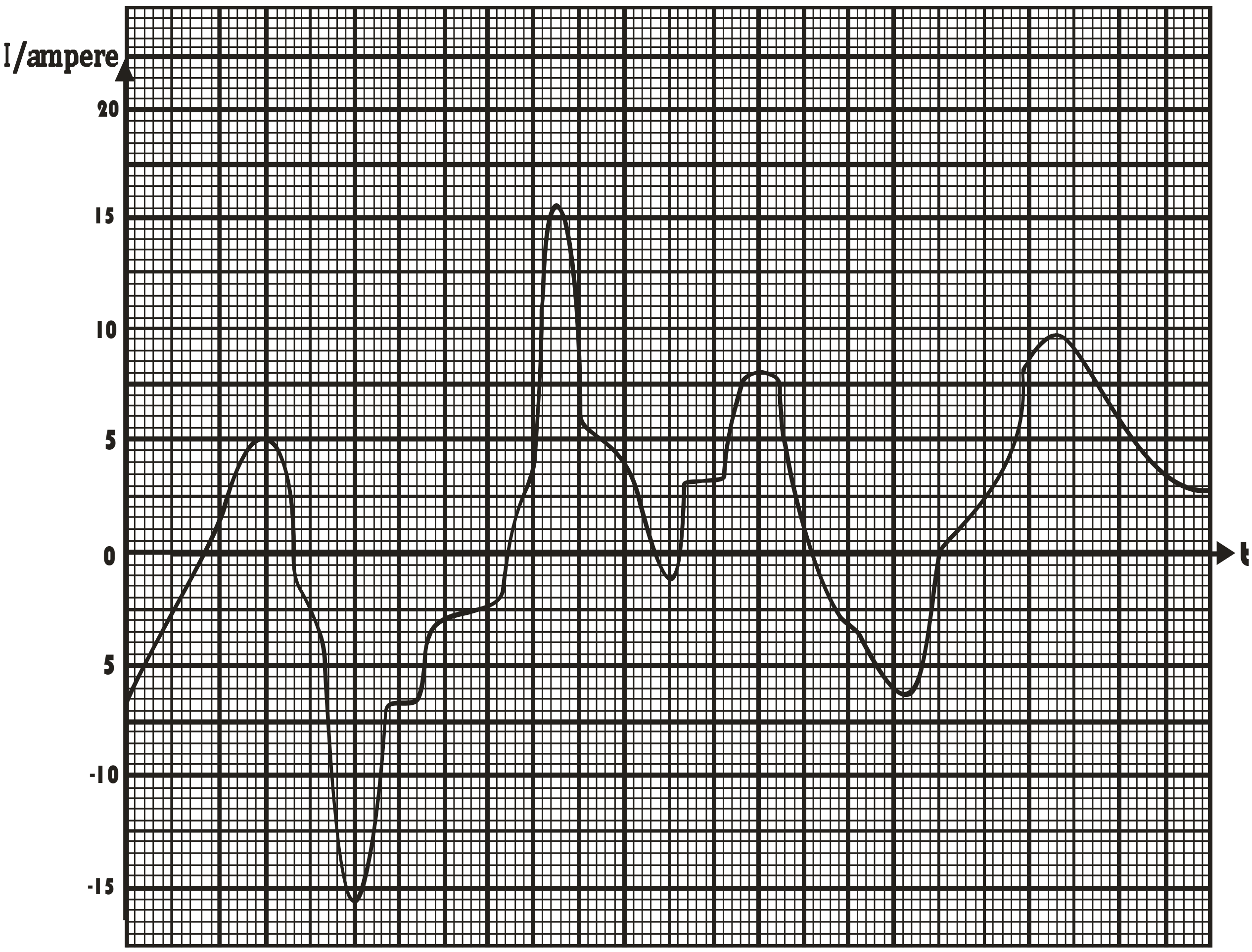


Fig 8Q3

8.24 (a) Explain briefly

(i) Pulse Modulation

(ii) Pulse Amplitude Modulation and

(iii) Pulse Width Modulation

(b) Explain briefly

(i) Pulse Code Modulation.

(ii) Quantization

(iii) Coding
(iv) Companding

(c) (i) State Nyquist Sampling Theorem

(ii) The maximum frequency of baseband signal is f_{max}. The baseband signal is pulse code modulated with n-bit per sample. What is the required minimum pulse code modulation bandwidth.

(iii) State the advantage and disadvantage of pulse code modulation.

(d) The signal-to-quantizing noise ratio of a binary PCM system is required to be at least 4000.

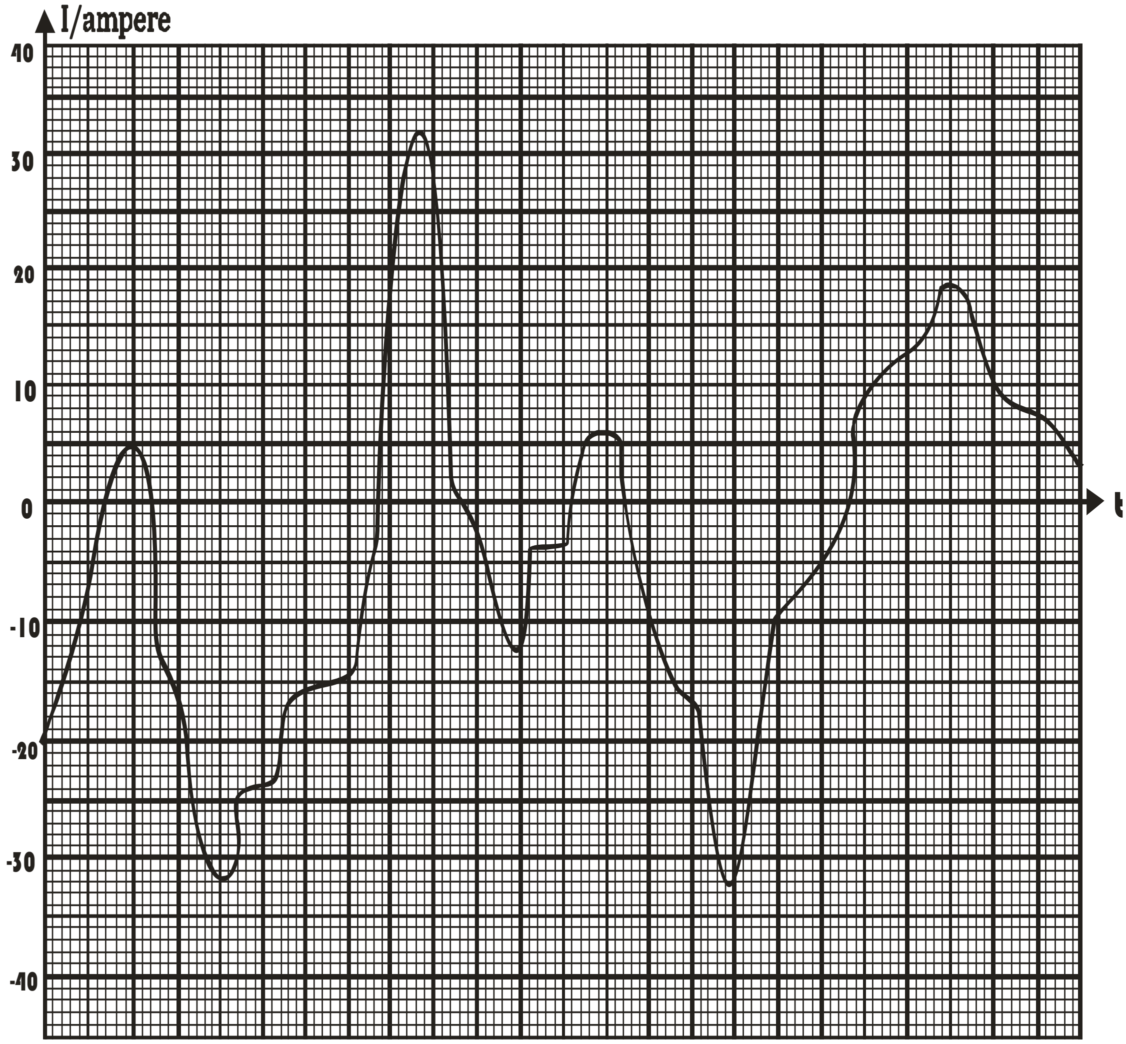


Fig 8Q4

Determine the required minimum number of binary digits to represent the quantizing level.

(e) Suppose the sinusoidal signal in Fig. 8Q4 is sampled at time t = 0, 5, 10, ………...60, s. The maximum amplitude of the signal is 32 A.

(i) Draw the PAM signal.

(ii) Draw the PCM signal. Using three bits per sample. Draw a table to show the possible voltage ranges, quantization levels and corresponding PCM Codes.

(iii) Determine the signal-to-quantization noise ratio in dB.

(Note: The signal varies from – 32 A through 0 to + 32 A)

(University of Ibadan, TEL512-Communication system II 2010/2011 BSc degree Exam)

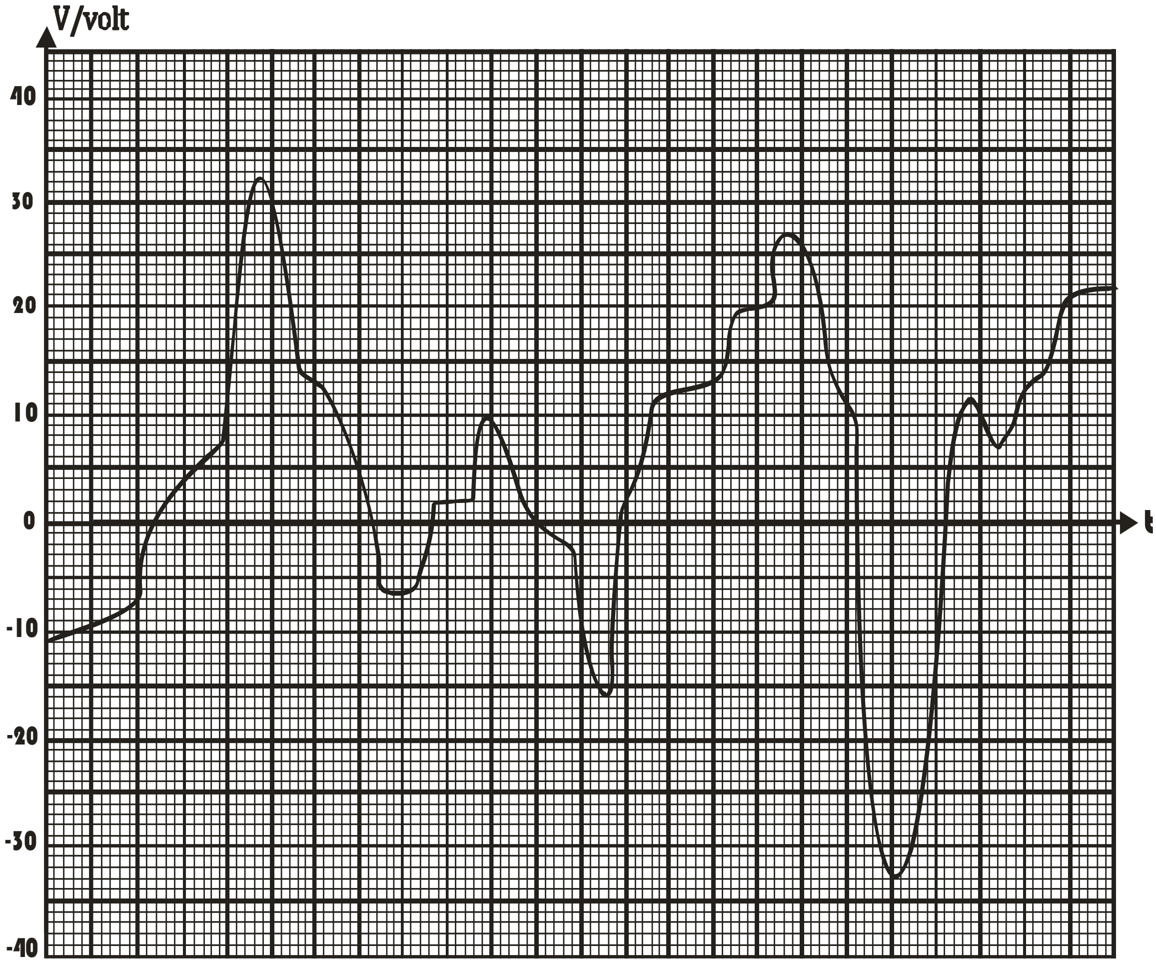


Fig 8Q5

8.25 (a) (i) Draw the block diagrams of a Delta Modulation System showing the modulator and demodulator.

(ii) Explain briefly Delta Modulation.

(iii) With the aid of a diagram of waveforms, explain "Start Up", "Hunting" and Slope overload".

(iv) Compare Delta Modulation with Pulse Code Modulation.

(b) A sinusoidal signal of 4 V amplitude and 2000 Hz frequency is transmitted by delta modulation. Determine the minimum step size so that the overload distortion is avoided, if the sampling rate is 50,000 samples/second.

(c) (i) What is PSK?

(ii) Explain the PSK modulation process.

(iii) Draw the block diagram of a PSK system.

(c) (i) State Nyquist Sampling Theorem

(ii) The maximum frequency of baseband signal is f_{max}. The baseband signal is pulse code modulated with n-bit per sample. What is the required minimum pulse code modulation bandwidth?

(iii) State the advantages and disadvantages of pulse code modulation.

(d) The signal-to-quantizing noise ratio of a binary PCM system is required to be at least 3000. Determine the required minimum number of binary digits to represent the quantizing level.

(e) Suppose the sinusoidal signal in Fig. 8Q5 is sampled at time t = 0, 5, 10, ………...60 s. The maximum amplitude of the signal is 32 V.

(i) Draw the PAM signal using single line

(ii) Draw the PCM signal. Using three bits per sample. Prepare a table to show the possible voltage range, quantization levels and corresponding PCM Codes. Determine the signal-to-quantizing noise ratio in dB. (Note: The signal varies from -32 V through 0V to +32 V)

(University of Ibadan, TEL512-Communication system II 2010/2011 BSc degree Exam)

8.26 (a) (i) Draw the block diagrams of a Delta Modulation System showing the modulator and demodulator.

(ii) Explain briefly Delta Modulation.

(iii) With the aid of a diagram of waveforms, explain "Start Up", "Hunting" and Slope overload".

(iv) Compare Delta Modulation with Pulse Code Modulation.

(b) A sinusoidal signal of 2 V amplitude and 1000 Hz frequency is transmitted by delta modulation. Determine the minimum step size so that the overload distortion is avoided, if the sampling rate is 50,000 samples/second.

(c) (i) What is PSK?

(ii) Explain the PSK modulation process.

(iii) Draw the block diagram of a PSK system.

(University of Ibadan, TEL512-Communication system II 2010/2011 BSc degree Exam)

CHAPTER 9

MULTIPLEXING

9.0 Introduction

Multiplexing is the process of simultaneously transmitting two or more individual signals over a single communications channel. Multiplexing has the effect of increasing the number of communication channels so that more information can be transmitted.

There are many instances in communications where it is necessary or desirable to transmit more than one voice or data signal. The application itself may require multiple signals and money can be saved by using a single communications channel to send multiple information signals. Telemetry and telephone applications are good examples. In satellite communications, multiplexing is essential to making the system practiced and for justifying the expense.

Telemetry is a good illustrative example. Telemetry is the process of measurement at a distance. Telemetry systems are used to monitor physical characteristic of some applications for the purpose of determining their status and operational conditions. This information may also be used as feedback in a closed-loop control system. Most spacecraft and many chemical plants, for example, use telemetry systems for monitoring their operations. Physical characteristics such as temperature, pressure, speed, light level, flow rate, and liquid level are monitored. Sensitive transducers convert these physical characteristics into electrical signals. These electrical signals are then processed in various ways and sent to a central monitoring location.

The most obvious way to send multiple signals from one place to another is to provide a single communications channels for each. For example, each signal could be sent over a single pair of wires. If long distances are involved, the signals will be degraded, and, therefore, special techniques must be used to prevent this. Using multiple wires is also an expensive process. When a very large number of signals must be monitored, many pairs of cables will be needed. This leads to extra cost and complexity. Ideally, it would be more economical if all the telemetry signals could somehow be combined and sent over a single cable.

In a spacecraft with multiple transducers, multiple transmitters would be required to send the signals back to earth. Again, this leads to incredible cost and complexity. In the case of a spacecraft, multiple transmitters would weigh much and consume an enormous amount of power, making them impractical. Again, the ideal situation would be to use a single transmitter and in some way combine all the various information signals and transmit them simultaneously over the single radio channel.

The telephone system is another example of the need for some means to increase the information-carrying capability of a single channel. There are hundreds of millions of telephones in this country, and each must be capable of being connected to any other telephone. This is what the telephone system is all about. If each telephone requires a two wire path, imagine the enormous number of wires required to make all the various interconnections. The problem is further compounded if you want to add in the ability to connect each telephone to any other telephone in the world. On such a large scale, cost is a major factor. Everything must be done to minimize the number of interconnecting wires. This leads to the use of some methods of combining multiple telephone conversations in such a way that they can be transmitted over a single pair of wires or a single radio communication channels.

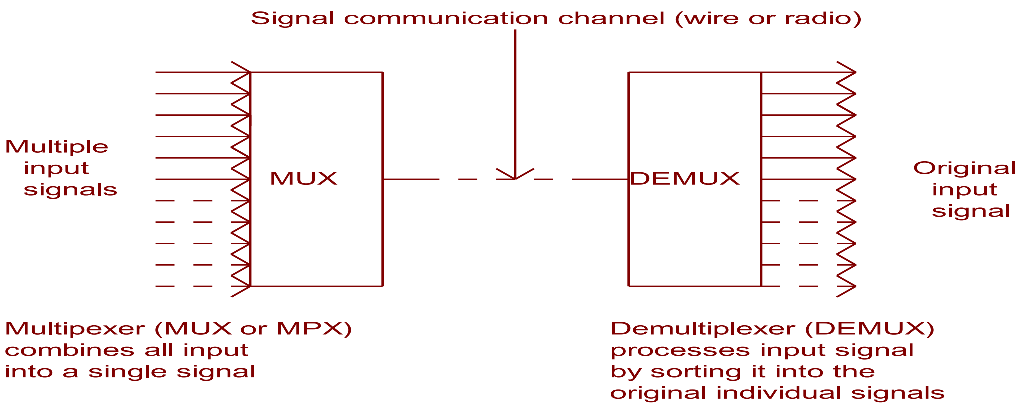


Figure 9.1 Concept of multiplexing

The concept of a simple multiplexer is illustrated in Fig. 9.1. Multiple input signals are combined by the multiplexer into a single composite signal that is transmitted over the communication medium. Alternatively, the multiplexed signal may modulate a carrier before transmission. At the other end of the communication link, a demultiplexer is used to sort out the signals into their original form.

There are two basic types of multiplexing; frequency division multiplexing (FDM) and time division multiplexing (TDM). Generally speaking, FDM systems are used to deal with analog information. Of course, TDM techniques are found in many analog applications as well because the process of analog-to-digital (A/D) and digital-to-analog (D/A) is so common. The primary difference between these techniques is that in FDM, individual signals to be transmitted are assigned a different frequency within a common bandwidth. In TDM, the multiple signals are transmitted in different time slots. In the following sections, we will discuss FDM and TDM in more detail.

9.1 Frequency Division Multiplexing

Frequency division multiplexing (FDM) is based on the idea that a number of signals can share the bandwidth of a common communications channel. The multiple signals to be transmitted over this channel are used to modulate a separate carrier. Each carrier is on a different frequency. The modulated carriers are then added together to form a single complex signal that is transmitted over the single channel.

9.1.1 FDM Concept

Fig. 9.2 shows a general block diagram of an FDM system. Each signal to be transmitted feeds a modulator circuit. The carrier for each modulation f_c is on a different frequency. The carrier frequencies are usually equally spaced from one another over a specific frequency range.

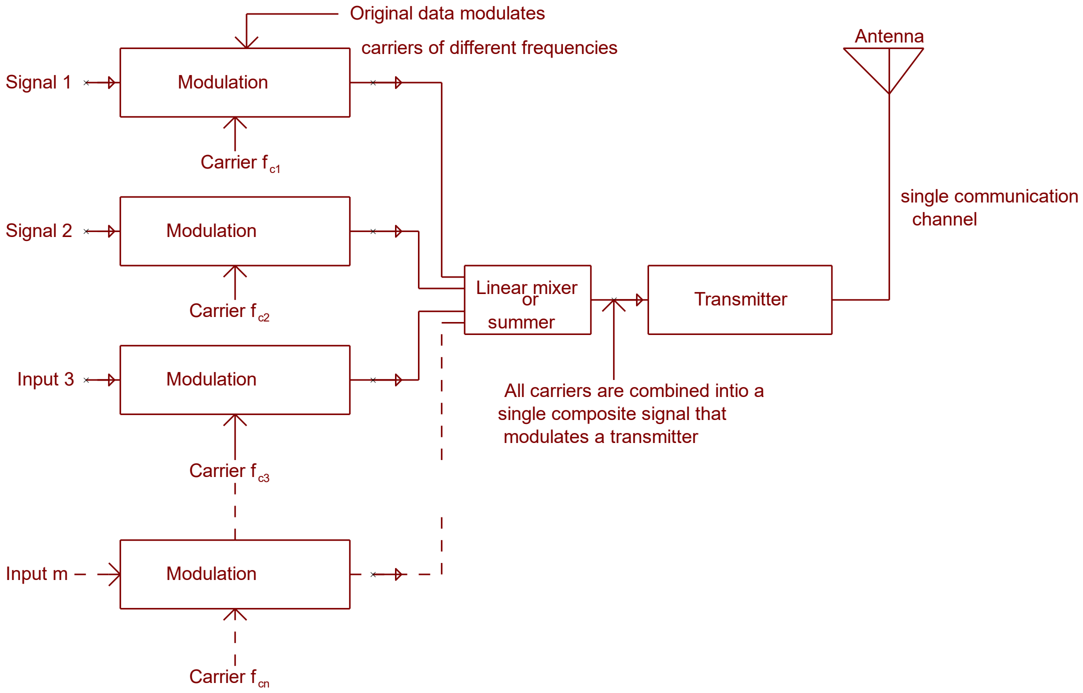


Figure 9.2 The Transmitting end of an FDM System

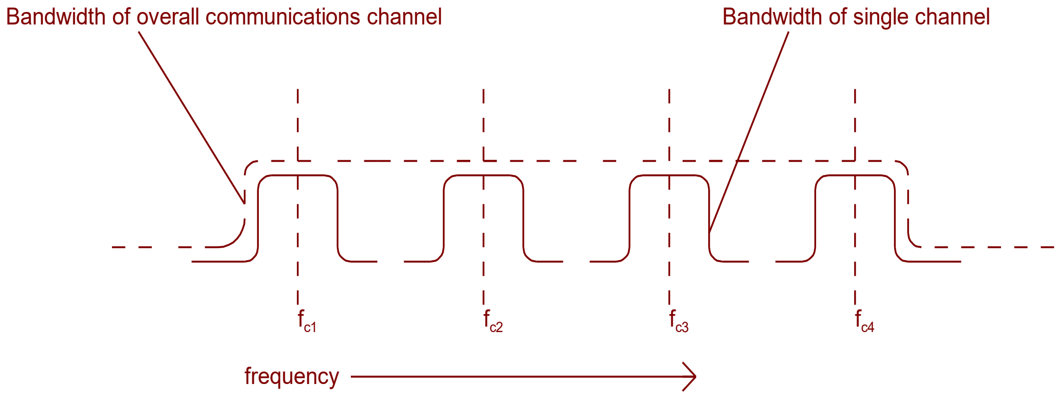


Figure 9.3 Spectrum of an FDM Signal

Each input signal is given a portion of the bandwidths. The result is illustrated in Fig. 9.3. As for the type of modulation, any of the standard kinds can be used including AM, SSB, FM or PM.

The modulator outputs containing the sideband information are added together in a linear mixer. In a linear mixer, modulation and the generation of sidebands do not take place. Instead, all the signals are simply added together algebraically. The resulting output signal is a composite of all carriers containing their modulation. This signal is then used to modulate a radio transmitter. Alternatively, the composite signal itself may be transmitted over the single communications channel. Another option is that the composite signal may become one input to another multiplexer system.

9.2 Demultiplexing FDM Signals

The receiving portion of the system is shown in Fig. 9.4. A receiver picks up the signal and demodulates it into the composite signal. This is sent to a group of bandpass filters (BPF), each centered on one of the carrier frequencies. Each filter passes only its channel and rejects all others. A channel demodulator then recovers each original input signal. To be specific about FDM systems, let’s consider three practical examples, telemetry, telephone, and FM stereo

9.2.1 FDM in Telemetry

Telemetry is one the most common uses for multiplexing techniques. In telemetry systems, many different physical characteristics are monitored by sensors. These generate electrical signals that change in some way to indicate the amplitude or measurement of the physical characteristics. An example of a sensor is a thermistor used to measure temperature.

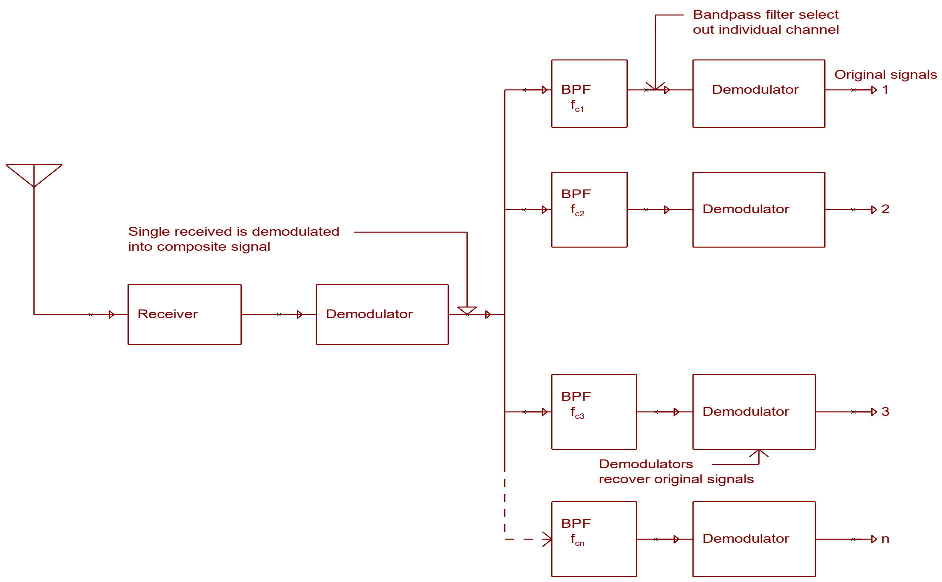


Figure 9.4 The Receiving end of an FDM System

A thermistor's resistance varies inversely with temperature: as the temperature increases the resistance decrease. The thermistor is usually connected to a dc voltage divider or bridge, and it is connected to a dc voltage source. The result is a dc output from this network whose voltage varies in accordance with the temperature. That varying dc level must then be transmitted to a remote receiver for measurement, readout, and recording. In this case, the thermistor signal becomes one channel of an FDM system.

Other sensors have different kinds of outputs. Many simply have varying dc outputs, and others may be ac in nature. Each of these signals is typically amplified, filtered, and otherwise conditioned before being use d to modulate a carrier. All the carriers are then added together to form a single multiplexed channel. In such systems, FM is normally used.

The conditioned transducer outputs are used to modulate a subcarrier. The varying direct or alternating current changes the frequency of an oscillator operating at the carrier frequency. Such a circuit is generally referred to as a voltage-controlled oscillator (VCO) or a subcarrier oscillator (SCO). The outputs of the SCOs are added together. A diagram of such a system is shown n Fig. 9.5.

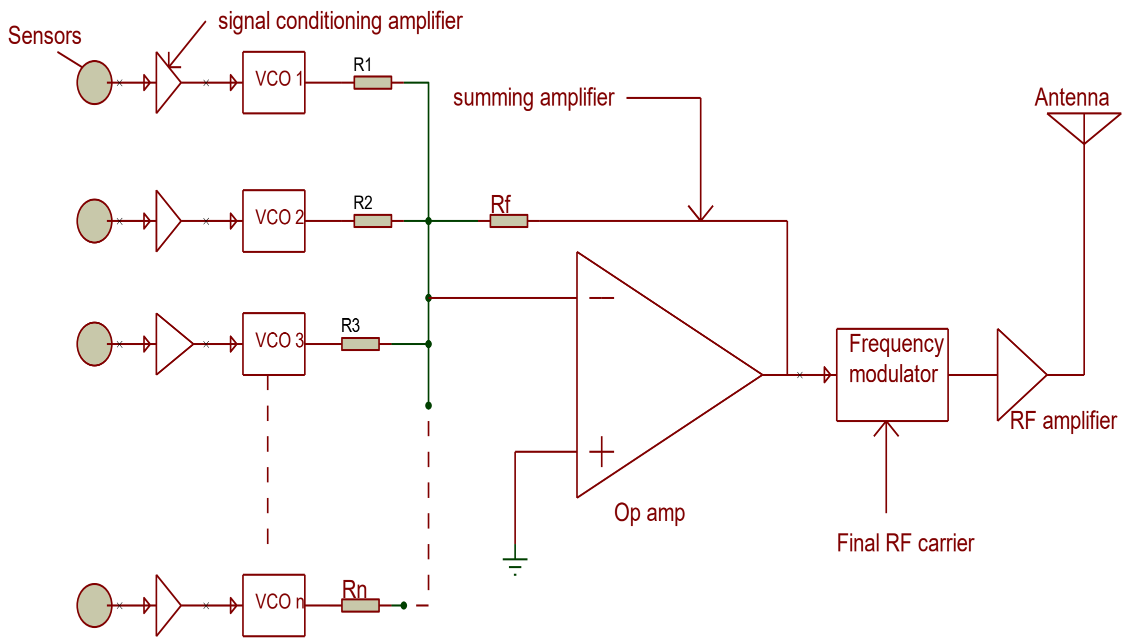


Figure 9.5 An FDM Telemetry Transmitting System

The output of the signal conditioning circuits is fed to the VCOs. Of course, to produce FDM, each VCO operates at a different frequency.

VCO Fig. 9.6(a) shows a block diagram of a typical VCO circuit. The VCOs are available as single IC chips. One version is called the 566. It consists of a dual-polarity current source that linearly charges and discharges an external capacitor C. The current value is set by an external resistor R_1. Together R_1 and C set the operating or center carrier frequency which can be any value up to about 1 MHz.

The current source may be varied by an external signal, either direct or alternating current. This is the modulating signal from a transducer or other source. The input signal varies the charging and current and, therefore, varies the carrier frequency, producing FM. The current source may be varied by an external signal, either direct or alternating current. This is the modulating signal from a transducer or other source. The input signal varies the charging and current and, therefore, varies the carrier frequency, producing FM.

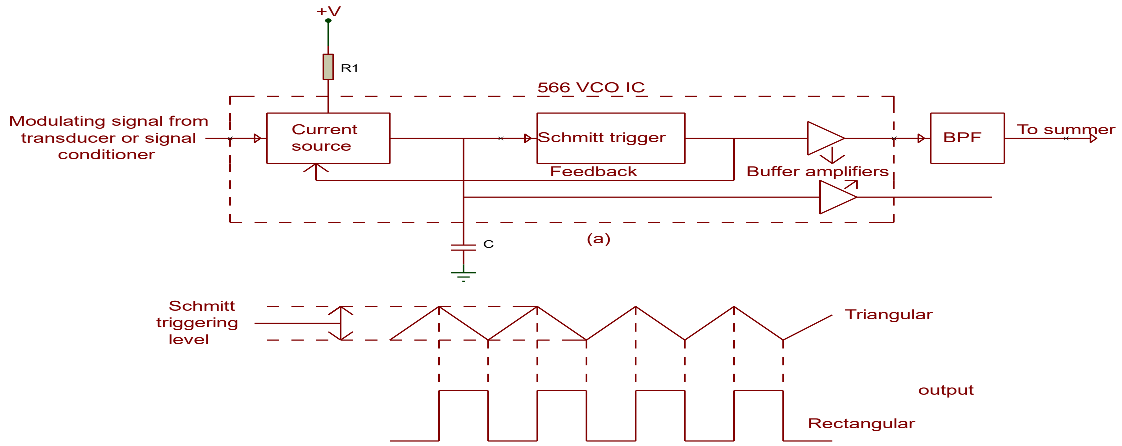


Figure 9.6 (a) Typical IC VCO Circuit AND (b) Waveform

The current source output is a linear triangular wave that is buffered by an amplifier for external use. This triangular waveform is also fed to an internal Schmitt trigger, which generates a rectangular pulse at the operating frequency. This is fed to a buffer amplifier for external use.

A Schmitt trigger output is also fed back to the current source, where it controls whether the capacitor is charged or discharged. For example, the VCO may begin by charging the capacitor. When the Schmitt trigger senses a specific level on the triangular wave, it switches the current source. Discharging then occurs. The waveform in Fig 9.6 (b) show this action. It is the feedback that creates a free-running, astable oscillator.

Most VCOs are astable multivibrators whose frequency is controlled by the input from the signal conditioning circuits. The frequency of the VCO changes linearly in proportion to the input voltage. Increasing the input voltage causes the VCO frequency to increase. The rectangular or triangular output of the VCO is usually filtered into a sine wave by a bandpass filter centered on the unmodulated VCO center frequency. This may be either a conventional LC filter or an active filter made with an op amp and RC input and feedback networks. The resulting sinusoidal output is applied to the linear mixer.

Linear Mixing: The linear mixing process in the FDM system can be accomplished with a simple resistor network as shown in Fig. 9.7. However, such networks greatly attenuate the signal. Typically, some voltage amplification is required in practical systems. A way to achieve the mixing and amplification at the same time is to use an op-amp summer like that shown in Fig. 9.5. Recall that the gain of the feedback resistor R_f to the input resistor value (R_1, R_2, etc.). The output is given by the expression

$$V_o = V_1 \frac{R_f}{R_1} + V_2 \frac{R_f}{R_2} + V_3 \frac{R_f}{R_3} + \cdots + V_n \frac{R_f}{R_n}$$

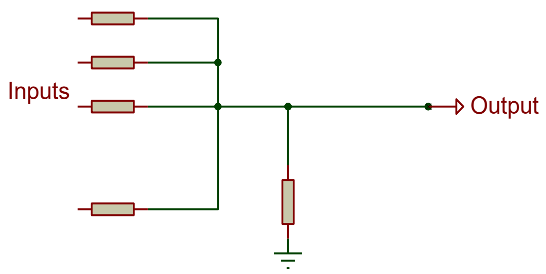


Figure 9.7 Resistive summing network

In most cases, the VCO FM output levels are the same, and, therefore, all input resistors on the summer amplifier are equal. However, if variations do exist, amplitude corrections can be made by making the summer input resistors adjustable. The output of the summer amplifier does invert the signal, but this has no effect upon the content.

Modulating Schemes: The composite output signal is then typically used to modulate a radio transmitter. Again, most telemetry systems use FM. A system that uses FM of the VCO subcarriers as well as FM of the final carrier is usually called an FM/FM system. However, keep in mind that other kinds of modulation schemes may be used.

Most FM/FM telemetry systems conform to standards established many years ago by an organization known as the Inter-Range Instrument Group (IRIG). These standards define specific channels as indicated in table 9.1. The center frequency of each channel is given along with the related channel number. A frequency deviation of ±7.5 percent is used on most of the channels, and this is increased to ±15 percent on the upper frequency channels. Also given in table 9.1 is the upper frequency range that the modulating signal can have on each channel. On channel 1, for example, with 400-Hz center frequency, the maximum signal frequency that can be used is 6 Hz. Most of the lowest frequency channels are used for direct current or very low frequency ac signals.

Table 9.1. The IRIG FM subcarrier brands and specifications.

Band Number	Center Frequency (Hz)	Lower Limit (Hz)	Upper limit (Hz)	Maximum Deviation (%)	Frequency Response (ops)
	400	370	430	7.5	6.0
	560	518	602	,,	8.4
	730	675	785	,,	11
	960	888	1,032	,,	14
	1,300	1,202	1,399	,,	20
	1,700	1,572	1,828	,,	25
	2,300	2,127	2,473	,,	35
	3,000	2,775	3,225	,,	45
	3,900	3,607	4,193	,,	59
	5,400	4,995	5,805	,,	81
	7,350	6,799	7,901	,,	110
	10,500	9,712	11,288	,,	160
	14,500	13,412	15,588	,,	220
	22,000	20,350	23,650	,,	330
	30,000	27,750	32,250	,,	450
	40,000	37,000	43.000	,,	600
	52,500	48,562	56,438	,,	790
	70,000	64,750	75,250	,,	1,050
	22,000	18,700	25,300	15	660
	30,000	25,500	34,500	,,	900
	40,000	34,000	46,000	,,	1,200
	52,500	44,625	60,375	,,	1,600
	70,000	59,500	80,500	,,	2,100

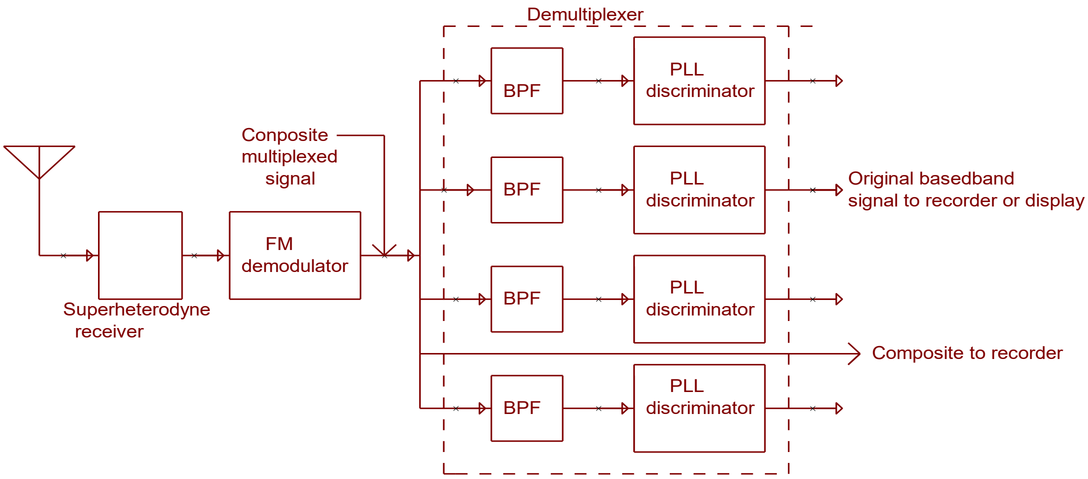


Figure 9.8 An FM/FM telemetry receiver

This set of standards is known as the proportional bandwidth FM/FM system. Since a fixed percentage of frequency deviation is specified, this means that the bandwidth is proportional to the carrier frequency. The higher the carrier frequency, the wider the bandwidth over which the modulating signal can occur.

Constant-bandwidth FM telemetry channels are also used. Carrier frequencies in the same approximate range as those given in table 9.1 are used. However, a fixed deviation of ±2 KHz is typically specified, thereby creating multiple channels with a 4 KHz bandwidth. These are spaced throughout the frequency spectrum with some guard space between channels to minimize interference.

The receiving end of a telemetry system appears as shown in Fig. 9.8. A standard super heterodyne receiver tuned to the RF carrier frequency is used to pick up the signal. An FM demodulator then reproduces the original composite multiplexed signal. This multiplexed signal is then fed to a demultiplexer that divides the signals and reproduces the original inputs. The outputs of the first frequency demodulator are fed simultaneously to multiple BPFs, each of which is tuned to the center frequency of one of the specified channels. Each filter passes only its subcarrier and related sidebands and rejects all the others. As you can see, the demultiplexing process is essentially that of using filters to sort the composite multiplex signal back into its original components. The output of each filter is the VCO frequency with its modulation.

These signals then, in turn, are applied to frequency demodulators. Also known as discriminators, these circuits take the FM signal and recreate the original dc or ac signal produced by the transducer. These original signals are then measured and otherwise interpreted to provide the desired information from the remote transmitting source. In most systems, the multiplexed signal is sent to a data recorder where it is stored for possible future use. The original telemetry output signals may be graphically displayed on a strip chart recorder or otherwise converted into usable outputs.

The demodulator circuit used in typical FM demultiplexers is either of the PLL or pulse averaging type. Of these two types, the PLL circuit is generally preferred because of its superior noise performance. However, it is typically more complex and expensive than the simpler pulse-averaging type. A PLL discriminator is also used to demodulate the receiver output.

9.3 FDM in Telephone Systems

Another example of a commonly used FDM system is the telephone system. For years telephone companies have been using FDM to send multiple telephone conversations over a minimum number of cables. The concepts are the same as those previously discussed. Here the original signal is voice in the 300-3000Hz range. The voice is used to modulate a subcarrier. Each subcarrier is on a different frequency. These subcarriers are then added together to form a single channel. This multiplexing process is repeated at several levels so that an enormous number of telephone conversations can be carried over a single communications channel, assuming its bandwidth is sufficient.

The frequency plan for a typical telephone multiplex system is shown in Fig. 9.9. The symbols are explained in Fig. 9.9. Here the voice signal amplitude modulates 1 of 12 channels in the 60-108 kHz range. The carrier frequencies begin at 60 kHz with a spacing of 4 kHz, slightly higher than the highest frequency used in a typical voice communication.

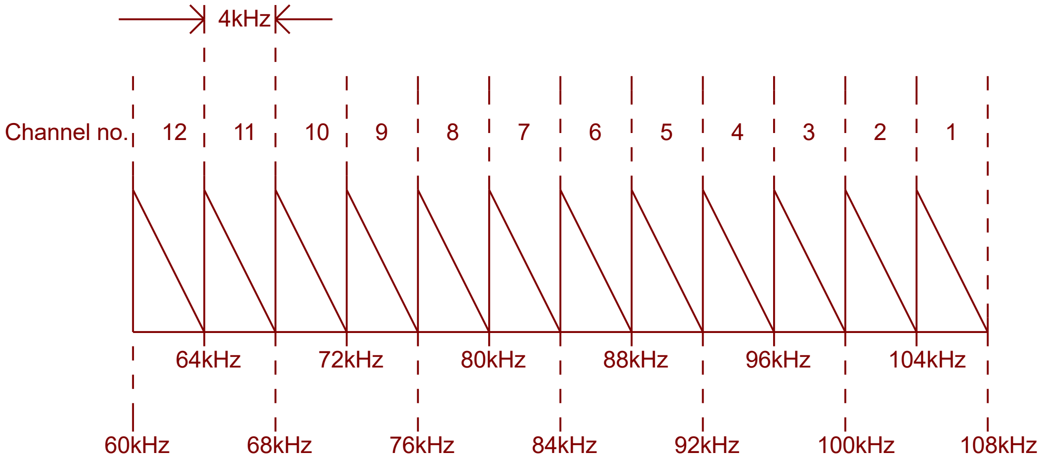


Figure 9.9 Basic Group Frequency Plan for FDM Telephone System

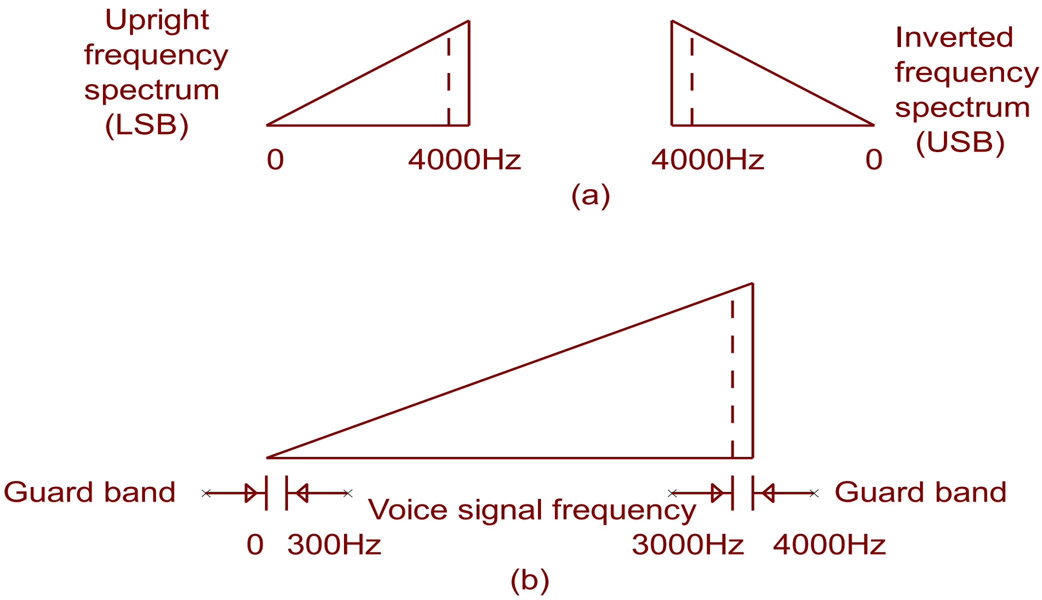


Figure 9.10 Meaning of symbols in Fig. 9.10 (a) Upper and Lower sidebands, and (b) sideband detail showing guarding bands

Single-sideband, suppressed-carrier modulation is used in telephone multiplex systems. The voice signal is applied to a balanced modulator along with a carrier. The output of the balanced modulator consists of the upper and lower sideband frequencies. The carrier is suppressed by the balanced modulator, and a highly selective filter is used to pass either the upper or lower sideband. The upper sidebands are selected here. The output of the filter is the sideband containing the original voice signal. All 12 SSB signals are then summed in a linear mixer to produce a single frequency multiplexed signal. This set of 12 modulated carriers is generally referred to as a basic group.

If more than 12 voice channels are needed, multiple basic groups are used. The outputs of these basic groups can then be further multiplexed onto higher-frequency subcarriers. In the telephone system, as many as five 12-channel basic groups can be combined. Carrier frequencies in the 360-552 kHz range are used. These carriers are spaced 48KHz apart. The frequency plan is shown in Fig. 9.12. The multiplexing process is similar, but here the output of each basic group modulates the higher-frequency carriers which are again summed to create an even more complex single-channel signal. Also SSB is used, and the lower sidebands are selected. Each of the 5 channels in this group carries 12 channels for a total of 60 voice signals. This composite signal is referred to as a super group.

This process may again be taken another step further. Up to 10 super groups can be used to modulate subcarriers in the 60-2540 kHz range. This allows a total of 5 (12) (10) = 600 voice channels to be carried. The output of these 10 multiplexers is referred to as a master group. The process can continue with six master groups being further combined into one jumbo group for a total of 3600 channels. These three jumbo groups can then be multiplexed again into one final output to achieve a total of 10,800 voice channels. Just keep in mind that as more and more levels of multiplexing are used, the bandwidth required carrying all the signals increases. A bandwidth of many megahertz is required to deal with the 10,800 channel composite signal mentioned above.

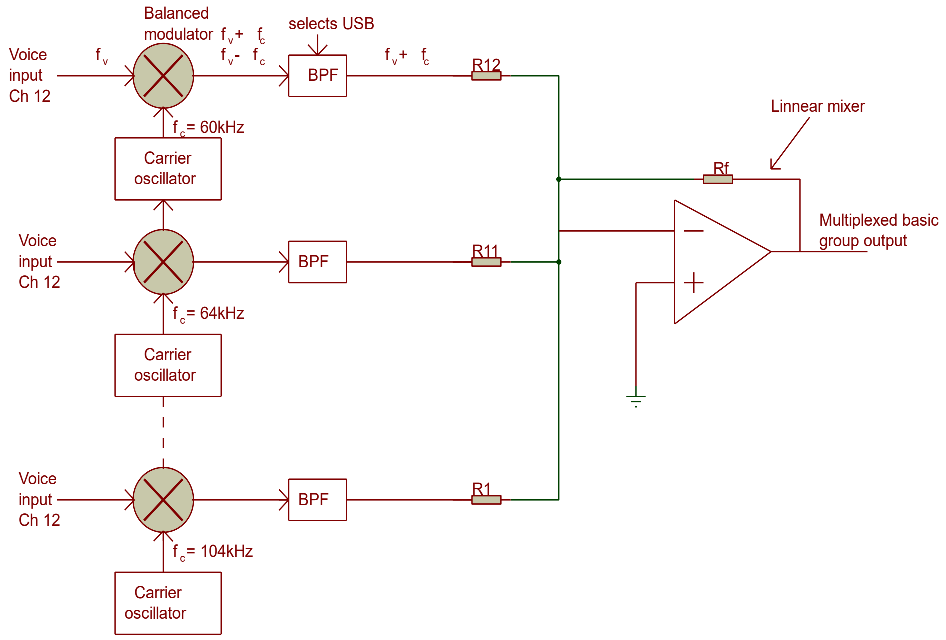


Figure 9.11 An FDM Telephone Multiplexer using SSB

The receiving end of the system is shown in Fig. 9.13. Bandpass filters select out the various channels, and balanced modulators are used to re-inject the carrier frequency and produce the original voice input. The telephone FDM system described here is no longer used in modern telephone systems. Instead, a digital multiplexing technique is used.

9.4 FDM in Stereo FM

Another well-known example of FDM is broadcast stereo FM. All FM broadcast stations use frequency multiplexing to transmit two channels of audio to the FM receiver in your car or to your Hi-Fi system at home. Let's take a look at how it works.

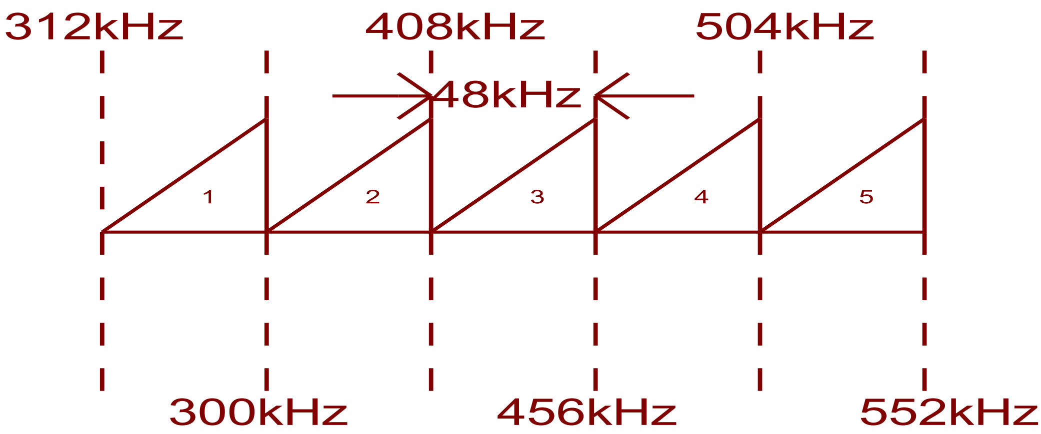


Figure 9.12 Five basic groups are used to form a super group

In stereo, two microphones are used to generate two separate audio signals. The two microphones pick up sound from a common source, such as a voice or band, but from different directions. The separation of the two microphones provides sufficient difference in the two audio signals to provide more realistic reproduction of the original sound. These two independent signals must somehow be transmitted by a single transmitter. This is done by frequency multiplexing techniques.

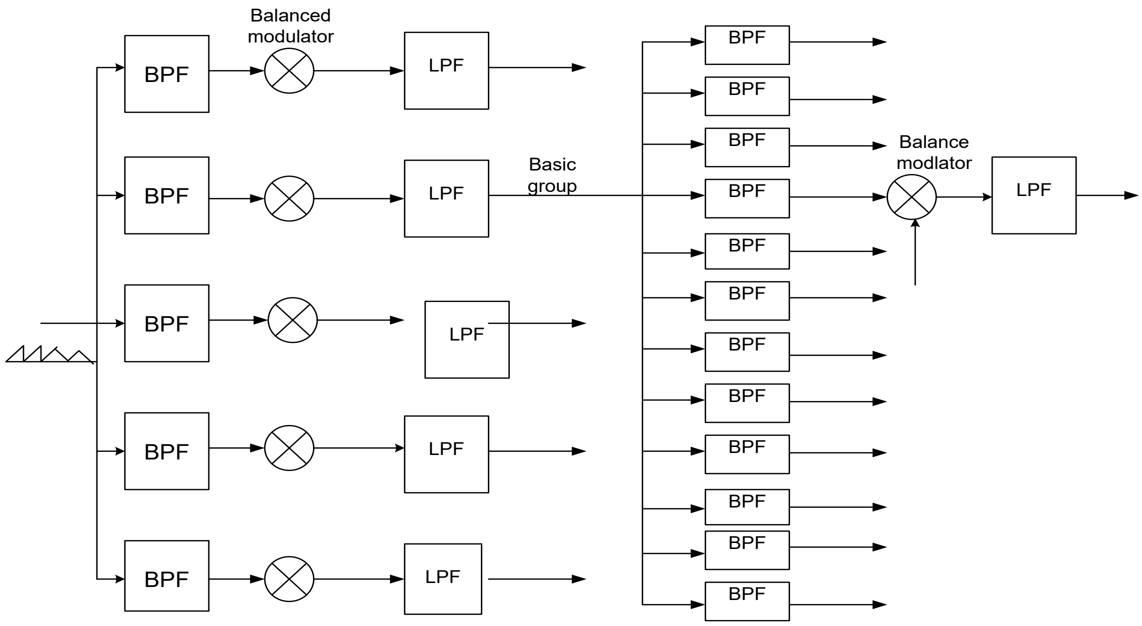


Figure 9.13 Demultiplexing the telephone signals

Fig. 9.14 is a general block diagram of a stereo FM multiplex modulator. The two audio signals generally called the left L and right R signals designate the positions of the microphones picking up the original sound. These two signals are fed to a circuit where they are combined to form sum L + R and difference L – R signals. The L + R signal is a linear algebraic combination of the left and right channels. The composite signal it produces is the same as if a single microphone were used to pick up the sound. It is the signal that a monaural receiver will hear. See Fig. 9.15.

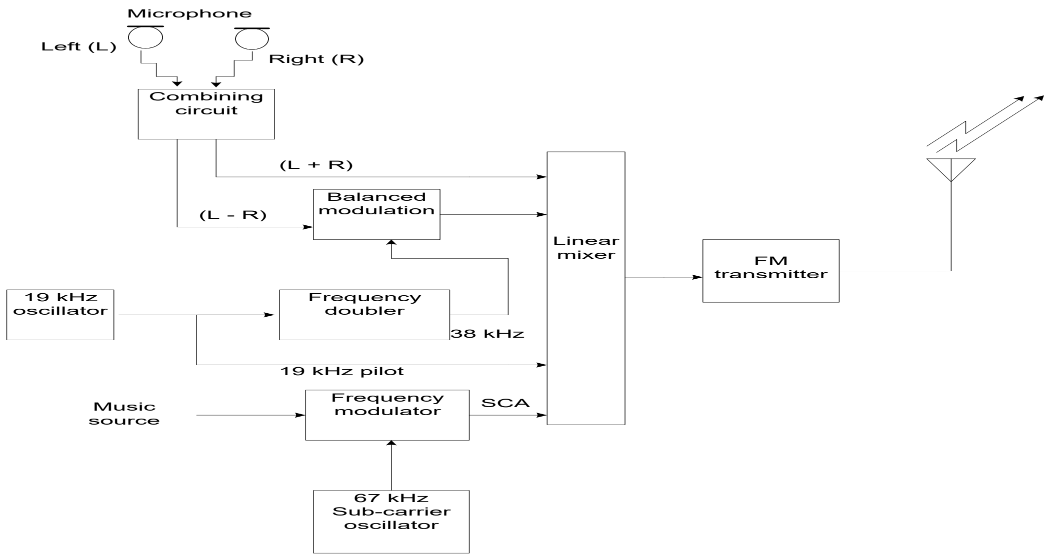


Figure 9.14 General block diagram of an FM stereo multiplex modulator, multiplexer and transmitter.

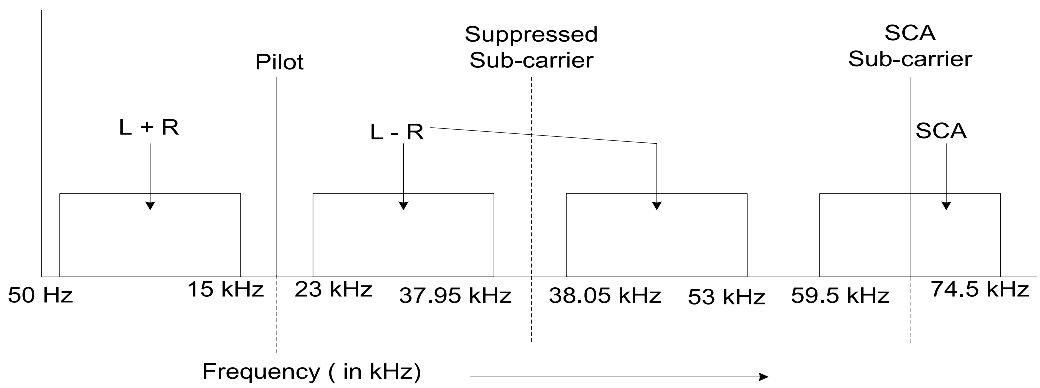


Figure 9.15 Spectrum of FM stereo multiplex broadcast signal. This signal frequency modulates the RF carrier.

The combining circuit inverts the right channel signal thereby subtracting it from the left channel signal to produce the L – R signal. These two signals, L + R and L – R, will be transmitted independently and recombined later in the receiver to produce the individual right and left hand channels.

The L – R signal is used to amplitude-modulate a 38 kHz carrier. This carrier is fed to a balanced modulator along with the L – R signal. The balanced modulator suppresses the carrier but generates upper and lower sidebands as shown in Fig. 9.15. Since the audio response of an FM signal is in the 0.05 to 15 kHz range, the sidebands are in the frequency range of 3 kHz ± 15 kHz or in the range of 23 to 53 kHz. This DSB signal will be transmitted along with the standard L + R audio signal. Also transmitted with the L + R and L – R signals is a 19 kHz pilot carrier. This is generated by an oscillator whose output will also modulate the main transmitter. Note that the 19 kHz oscillator drives a frequency doubler to generate the 38 kHz carrier for the balanced modulator.

Some FM stations also broadcast another signal referred to as the Subsidiary Communications Authorization (SCA) signal. This is a separate subcarrier of 67 kHz subcarrier with its music modulation will also modulate the FM transmitter. As in other FDM systems, all the subcarriers are added with a linear mixer to form a single signal. The spectrum of that composite signal is shown in Fig. 9.15. This signal is used to frequency-modulate the carrier of the broadcast transmitter. Again note that FDM simply provides a portion of the frequency spectrum for such independent signals to be transmitted. In this case there is sufficient spacing between adjacent FM stations so that the additional information can be accommodated. Some FM stations now transmit computer data over other subcarriers.

At the receiving end, the demodulation is accomplished with a circuit similar to that illustrated in Fig. 9.16. The FM super-heterodyne receiver picks up the signal, amplifier it, and translates it to an IF, usually 10.7 MHz. It is then demodulated. The output of the demodulator is the original multiplexed signal. The various additional circuits now sort out the various signals and reproduce them in their original form.

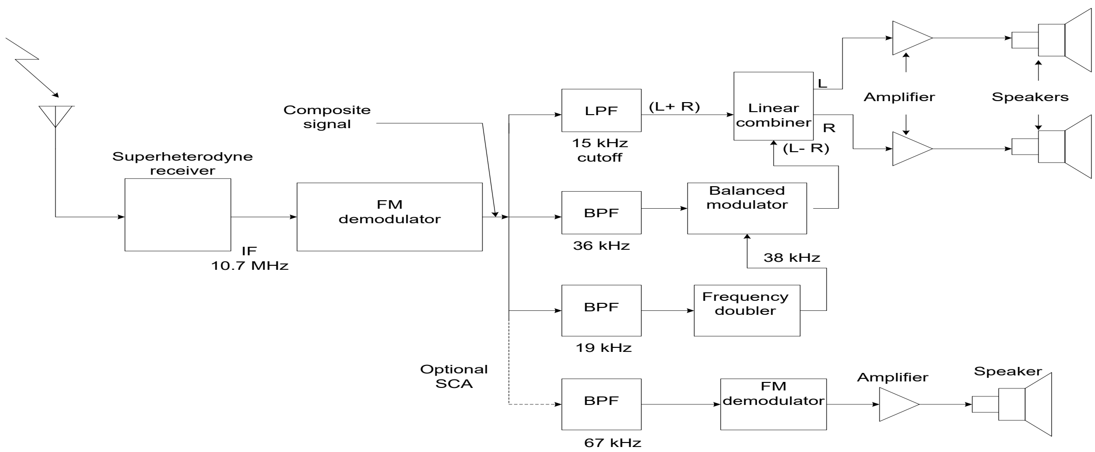


Figure 9.16 Demultiplexing and recovering the FM stereo and SCA signals

The original audio L + R signal is extracted by passing the multiplex signal through a low pass filter. Only the 50 to 15,000 Hz original audio is passed. This is signal is fully compatible with monaural FM receivers without stereo capability. In a stereo system, the L + R audio signal is fed to a linear matrix where it is mixed with the L – R signal to create the two separate L and R channels. The multiplexed signal is also applied to a bandpass filter that passes the 38 kHz suppressed subcarrier with its sidebands. This is the L – R signal that modulates the 38 kHz carrier. This signal is fed to a balanced modulator for demodulation.

The 19 kHz pilot carrier on the multiplexed signal is extracted by passing the multiplexed signal through a narrow bandpass filter. This 19 kHz subcarrier is then fed to an amplifier and frequency doubler circuit which produces a 38 kHz carrier signal. This is fed to the balanced modulator, of course, is the L – R audio signal. This is fed to the linear resistive matrix along with the L + R signal. If the SCA signal is used, a separate bandpass filter centered on the 67 kHz subcarrier would extract the signal and feed it to a frequency demodulator. The demodulator output would then be sent to an audio amplifier and speaker.

At this point the L + R and the L – R audio signals have been removed and fed to linear matrix. The matrix simply performs an algebraic operation on the two signals. The matrix both adds and subtracts these two signals. Adding the signals produces the left-hand channel.

$(L + R) + (L – R) = 2L$

Subtracting the two signals produces the right-hand channel.

$(L + R) – (L – R) = 2R$

The left and right hand audio signals are then sent to separate audio amplifiers and ultimately to the speakers. Similar multiplex systems are used in AM stereo radio and TV stereo.

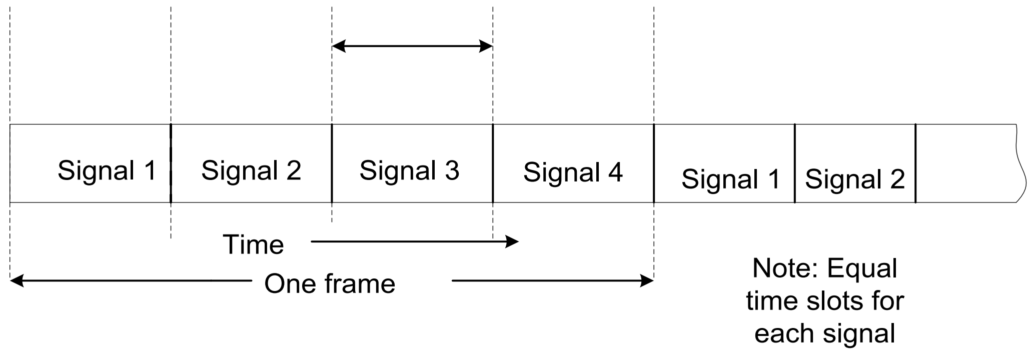


Figure 9.17 The basic TDM concept

9.5 Time Division Multiplexing

In FDM, multiple signals are transmitted over a single channel by sharing the channel band-width. This is done by allocating each signal a portion of the spectrum within that bandwidth. In TDM, each signal can occupy the entire bandwidth of the channel. However, each signal is transmitted for only a brief period of time. In other words, the multiple signals take turns transmitting over the single channel. This concept is illustrated graphically in Fig. 9.17. Here, four signals are transmitted over a single channel. Each signal is allowed to use the channel for a fixed period of time, one after another. Once all the signals have been transmitted, the cycle repeats again and again.

Time division multiplexing may be used with both digital and analog signals. To transmit multiple digital signals, the data to be transmitted is formatted into serial data words. For example, the data may consist of sequential bytes. One byte of data may be transmitted during the time interval assigned to a particular channel. For example, in Fig. 9.17, each time slot might contain 1 byte from each channel. One channel transmits 8 bits and then halts while the next channel transmits 8 bits. The third channel then transmits its data word and so on. One transmission of each channel completes one cycle of operation called a frame. The cycle repeats itself at a high rate of speed. In this way, the data bytes of the individual channels are simply interleaved. The resulting single channel signal is a digital bit stream that must somehow be deciphered and reassembled at the receiving end.

9.5.1 Pulse Amplitude Modulation

The transmission of digital data by TDM is straightforward in that digital data is incremental and can be broken up into words that can be easily assigned to different time slots. What is not obvious is how TDM can be used to transmit continuous analog signals. Yet, virtually any analog signals, be it voice, video, or telemetry measurements, can readily be transmitted by TDM techniques. This is accomplished by sampling the analog signal repeatedly at a high rate.

Sampling the process of "looking at" an analog signal for a brief instant of time, during this very short sampling interval, the amplitude of the analog signal is allowed to be passed or stored. By taking multiple samples of the analog signal at a periodic rate, most of the information contained in the analog signal will be passed. The resulting signal will be a series of samples or pulses that vary in amplitude according to the variation of the analog signal.

Modulation: Fig. 9.18 shows an analog signal. The resulting output is a series of pulses whose amplitudes are the same as those of the analog signal during the sample period. This process is known as pulse-amplitude modulation (PAM).

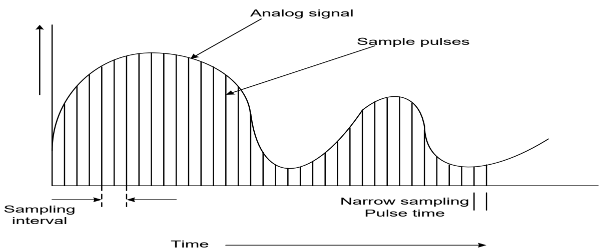


Figure 9.18 Sampling an analog signal to produce pulse amplitude modulation

The basic circuit for generating PAM is illustrated in Fig. 9.19. An astable clock oscillator drives a one-shot multivibrator which generates a narrow fixed width pulse. This pulse is applied to a gate circuit that is essentially a switch that will open and close in accordance with the one-shot signal. When the one shot is off, the gate is closed and the analog signal applied to it will not pass. When the clock triggers the one shot once per cycle, the gates opens for a short period of time allowing the analog signal to pass through. The gate circuit may be constructed with diodes or can be arrangement of bipolar or field-effect transistors.

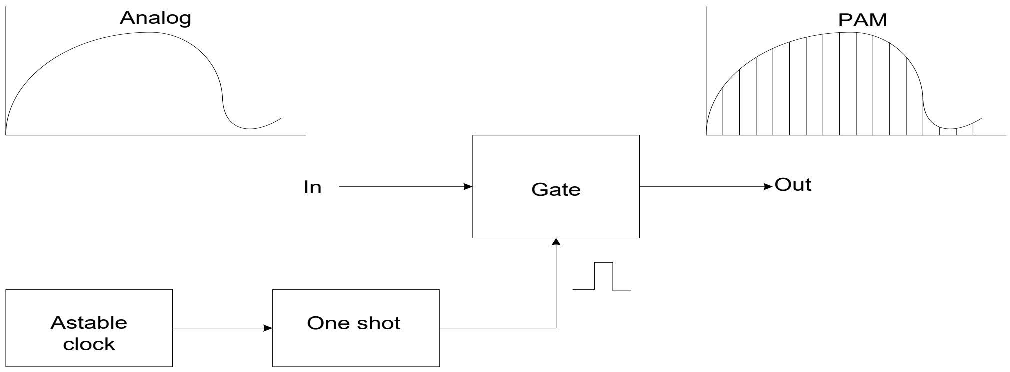


Figure 9.19 A Pulse Amplitude Modulator

DEMODULATION: To recover the original information, the transmitted pulses are simply passed through a low-pass filter. The upper cutoff frequency of the low-pass filter is selected to pass the highest-frequency components contained within the analog signal. All higher frequencies are eliminated. Since the pulses themselves represent a composite of many high frequency harmonics, these are effectively filtered out. The pulses, therefore, are smoothed into a continuous analog signal that is virtually identical in information content to the original transmitted signal. The process is similar to that used in a simple AM diode detector.

SAMPLING RATE: In order for the recovered signal to be an accurate representation of the original, the sampling rate must be high enough to ensure that rapid fluctuations are sampled a sufficient number of times. It has been determined that the sampling rate must be at least two times the highest frequency component of the original signal in order for the signal to be adequately represented. This relationship between the original analog signal and the sampling theorem, if the upper bandwidth value of the analog signal is known, the minimum sampling rate can be found by simply multiplying it by 2.

If the sampled signal is a simple sine wave, then the minimum sampling frequency can be twice the sine wave frequency. A 2 kHz sine wave would have to be sampled a minimum or 2 x 2 kHz, or 4 kHz. A more complex signal containing harmonics up to 650 kHz would have to be sampled at a 2 x 650 kHz = 1300 kHz, or 1.3 MHz rate or higher. The higher the sampling rate, the better the representation, most systems sample at a rate higher than the minimum 2 times to ensure good fidelity.

In telephone communications, the upper frequency value of the voice content is assumed to be 3 kHz. This dictates a 2 x 3 kHz or 6 kHz, sampling rate. In practice, the sampling rate for audio in telephone systems is 8kHz. The higher sampling rate provides more faithful reproduction of the audio signals. Although in many applications a sampling rate of twice the highest frequency content is satisfactory, usually the sampling rate is made much higher. The actual value depends upon the application, but typically the sampling rate is 4 to 5 times the highest frequency component in the analog signal. A sampling rate of 10 times the maximum analog bandwidth is ideal. This provides excellent representation of the signal.

9.6 Time Multiplexer

Now, by combining the concepts of TDM and PAM, you can see how multiple analog signals can be transmitted over a single channel. This is accomplished by a circuit called a multiplexer (usually abbreviated MUX or MPX). The multiplexer is simply a single-pole multiple positions mechanical or electronic switch that sequentially samples the multiple analog inputs at a high rate of speed. The simple rotary switch shown in Fig. 9.20 is an example.

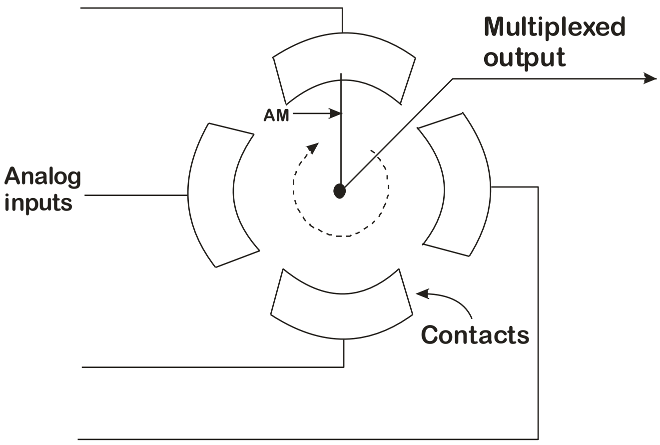


Figure 9.20 Simple Rotary Switch Multiplexer

The switch arm dwells momentarily on each contact allowing the input signal to be passed through to the output. It then switches quickly to the next channel and allows that channel to pass for a fixed duration. The remaining channels are sampled in the same way. After each signal has been sampled, the cycle repeats. The result is that four analog signals will be sampled, creating PAM signals that are interleaved with one another. Fig. 9.21 illustrates how four different analog signals are sampled by this technique. Be sure that you study the figure so that you can recognize each of the fours signals in the composite waveform.

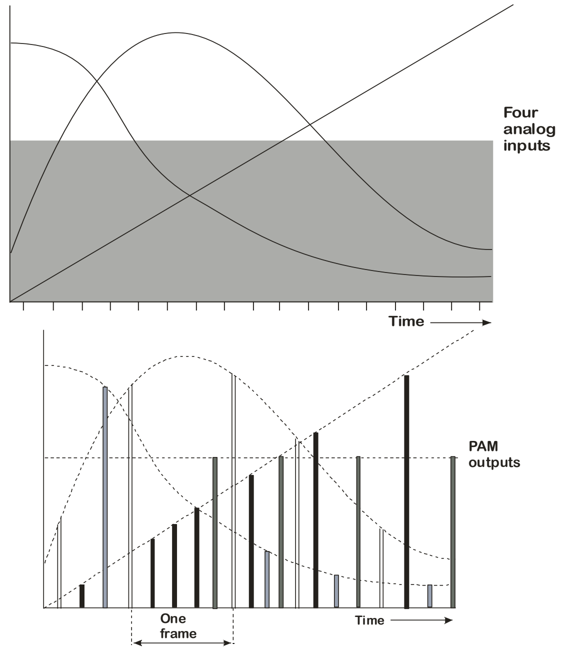


Figure 9.21 Four Channel PAM Time Division Multiplexer

Multiplexers used in early TDM/PAM systems used a form of rotary switch known as a commutator. Multiple switch segments were attached to the various incoming signals, and a high speed brush rotated by a dc motor rapidly sampled the signal as it passed over the contacts. Such commutators were used in early telemetry systems but have now been totally replaced by electronics circuits.

In practice, the duration of the sample pulses is shorter than the time which is allocated to each channel. For example, assume it takes the commutator or multiplexer switch 1ms to move from one contact to another. The contacts could be set up so that each sample to 1ms long. Typically, the duration of that sample is usually made about half that period, or in this example, 0.5 ms. One complete revolution for the commutator switch is referred to as a frame. In other words, during one frame, each input channel is sample one time. The number of contacts on the multiplexer switch or commutator sets the number of samples per frame. The number of frames completed in 1s is called the frame rate. If you multiply the numbers of samples per frame by the frame rate, you will get the commutation rate or multiple rates. This is the frequency of the pulses in the final multiplexed signal.

In our example in Fig. 9.21, the number of samples per frame is four. Assume that the frame rates are 100 frames per second. The period for one frame, therefore is 1/100 = 0.01 = 10 ms. During that 10 ms frame period, each of the four channels will be sampled once. Assuming equal sample durations, each channels would be allowed 10/4 = 2.5 ms.

As indicated earlier, the full 2.5 ms period would not be used. Instead, the sample duration during that interval might only be 1ms long. Since there are four samples taken per frame, the commutation rate would be 4(100) or 400 pulses per second. This would be the basic frequency of the composite signal to be transmitted over the communications channel. In practice TDM/PAM systems, electronic circuits are used instead of mechanical switches or commutators. The multiplexer itself is usually implemented with FETs, which are nearly ideal off/on switches that can turn off and on at very high speeds. A complete TDM/PAM circuit is illustrated in Fig. 9.22. Only four channels are used so as to simplify the discussion.

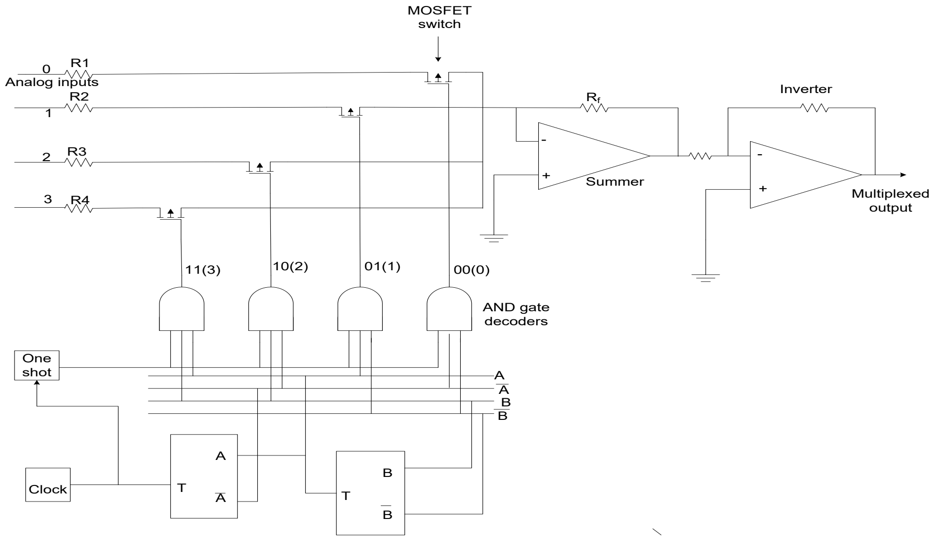


Figure 9.23 A PAM Multiplexer

The multiplexer is an op-amp summer circuit with FETs on each input resistor. When the FET is conducting it has a very low resistance and, therefore, acts as a closed switch. When the transistor is cut off, no current flows through it and, therefore, it acts an open switch. A digital pulse applied to the gate of the FET turns the transistor on. The absence of a pulse means that the transistor is cut off. The control pulses to the FET switches are such that only one FET is turned on at a time. These FETs are turned on in sequence by the digital circuit illustrated.

All the FET switches are connected in series with resistors R_1 to R_4 that in combinations with the feedback resistor R_f on the op-amp circuit determine the gain. For our discussion here, we will assume that the input and feedback resistors are all equal in value, meaning that the op-amp circuit has a gain of 1. Since this op-amp summing circuit inverts the polarity of the analog signals, it is followed by another op-amp inverter that gain inverts and restores the proper polarity. The digital control pulses are developed by the counter and decoder circuit shown in Fig. 9.23. Since there are four channels, four counter states are needed. Such a counter can be implemented with two flip-flops which can represent four discrete states. These are 00, 01, 10 and 11. These are the binary equivalents of the decimal numbers 0, 1, 2, and 3. We can therefore, label our four channels as channels 0, 1, 2, and 3.

A clock oscillator circuit triggers the two flip-flop counters. The clock and flip-flow waveform are illustrated in Fig. 9.23. The flip-flop outputs are applied to the decoder gates. These are AND gates that are connected to recognize the four binary combinations 00, 01, 10, and 11. The output of each decoder gate is applied to one of the multiplexer FET gates. The one-short multivibrator shown in Fig. 9.23 is used to trigger all the decoder AND gates at the clock frequency. This one-shot multivibrator produces an output pulse whose duration has been set to the desired sampling interval. Recall in our earlier discussion the sampling interval was 2.5 ms and the actual sample length was set at 1ms. Here the one shot would have a 1ms pulse duration.

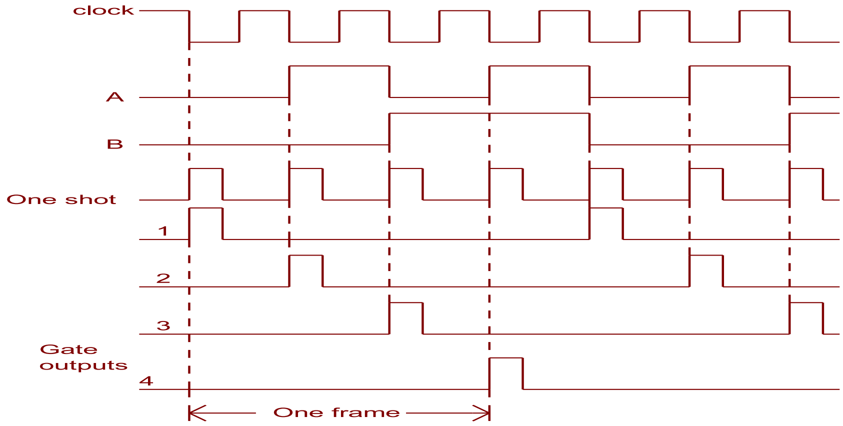


Figure 9.24 Waveforms for a PAM Multiplexer

Each time the clock pulse occurs, the one shot generates its pulse which is applied simultaneously to all four AND decoder gates. At any given time, only one of the gates is enabled. The output of the enabled gate will be a pulse whose duration is the same as that of the one shot. When the pulse occurs, it turns on the associated MOSFET and allows the analog signal to be sampled and passed through the op amps to the output. The output of the final op amp is the multiplexed PAM signal like that in Fig. 9.21.

9.7 Using PAM to Modulate a Carrier

The varying amplitude PAM signal is not transmitted as it is over the single channel. Instead, these varying amplitude pulses are used to modulate a carrier that is then transmitted over the communications medium. In most systems, the PAM signal is used to frequency-modulate either an RF carrier for radio transmission or a sub carrier which, in turn, modulates a final RF carrier. Two such arrangements are shown in Fig. 9.24. In the first arrangement, the PAM signals phase modulates a carrier. This system is, therefore, referred to as PAM/PM systems. In the second arrangements, the PAM signals phase modulates a subcarrier. These subcarriers are then linearly mixed and used to phase modulate the RF carrier, which is the final transmitted signal. This system is referred to as

PAM/PM/PM. The second system uses a combination of TDM and FDM schemes to create the final composite signal.

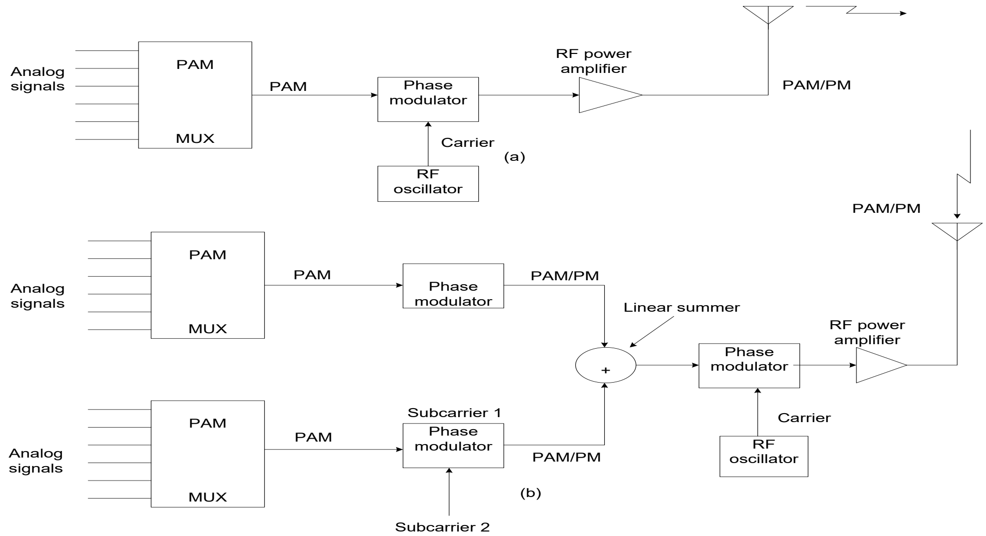


Figure 9.24 combining analog (frequency) and digital (PAM) multiplexing (a) PAM/PM, and (b) PAM/PM/PM

Once the composite signal is received, it must be demodulated and demultiplexed. In a PAM/PM/PM system, the signal is picked up by the receiver which ultimately sends the signal to a phase demodulator which recovers the original PAM data. In a PAM/PM/PM system, two levels of phase demodulation are required before the PAM signal is available. Once the composite PAM signal is obtained, it is applied to a demultiplexer (usually abbreviated DEMUX). The DEMUX is, of course, the reverse of a multiplexer. It has a single input and multiple outputs, one for each original input signal. Following with our four-channel example, a DEMUX for this system would have a single input and four outputs. Again, most DEMUXes use FETs driven by a pulse counter arrangement like that shown earlier in Fig. 9.22.

Demultiplexing: The main problem encountered in demultiplexing is synchronization. That is, in order for the PAM signal to be accurately demultiplexed into the original sampled signals, some method must be used to ensure that the clock frequency used on the DEMUX is identical to that used at the transmitting multiplexer. Further, even though the clock frequencies may be identical, the sequence of the DEMUX must be identical to that of the multiplexer so that when channel 1 is being sampled at the transmitter, channel 1 will be turned on in the receiver DEMUX at the same time. Such synchronization is usually carried out by a special synchronizing pulse included as a pan of each frame. Let's, take a look at some of the circuits used for clock frequency and frame synchronization.

Instead of using a free-running clock oscillator set to the identical frequency of the transmitter system clock, the clock for the DEMUX is derived from the received PAM signal itself. The circuits shown in Fig. 9.25 are typical of those used to generate the DEMUX clock pulses. They are called clockrecoverycircuits.In Fig. 9.25(a), the PAM signal is first applied to an amplifier-limiter circuit. This amplifies all the received pulses to a high level and then clips them off at a fixed level. The result is that the output of the limiter is a constant-amplitude rectangular wave whose output frequency is equal to the commutation rate. This is the frequency at which the PAM pulses occur. This, of course, is determined by the transmitting multiplexer clock.

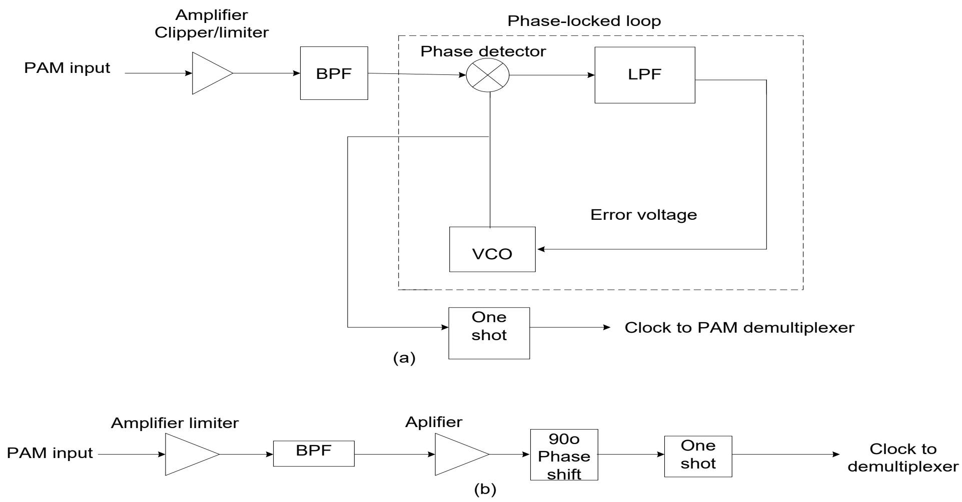


Figure 9.25 Two PAM Clock Recovery Circuits (a) Closed loop and (b) Open loop

The rectangular pulses at the output of the limiter are applied to a bandpass filter. This bandpass filter eliminates all the upper harmonics, creating a sine wave signal at the transmitting clock frequency. This signal is, applied to the phase detector circuit in a PLL along with the input from a VCO. The VCO is set to operate at the frequency of the PAM pulses However, the VCO frequency is controlled by a dc error voltage applied to its input. This input is derived from the- phase detector output which is filtered by low-pass filler into a dc voltage.

The phase detector compares the phase of the incoming PAM sine wave to the VCO sine wave. If a phase error exists, the phase detector will produce an output voltage that is translated into direct current to vary the VCO frequency. The system is stabilized or locked when the VCO output frequency is identical to that of the sine wave frequency derived from the PAM input. When the PLL is locked, the two sine waves are shifted in phase by 90°.If the PAM signals frequency changes for some reason, the phase detector picks up the variation and generates an error signal that is used to change the frequency of the VCO to match. Because of the closed-loop feature of the system, the VCO will automatically track frequency changes in the PAM signal. This means that the clock frequency used in the DEMUX will always perfectly match that of the original PAM signal regardless of any frequency changes that occur.The output signal of the VCO is applied to a one-shot pulse generator mat creates rectangular pulses at the proper frequency. These are used to step the counter in the DEMUX from which are derived the gating pulses for the FET 'DEMUX switches.

A simpler open-loop clock pulse circuit is shown in Fig. 9.25(b). Again, the PAM signal is applied to an amplifier-limiter and then a bandpass filter, just as it was in the previous circuit. The sine wave output of the bandpass filter is then amplified and applied to a phase shift circuit which produces 90° phase shift at the frequency of operation. This phase-shifted sine wave is then applied to a pulse generator which, in turn, creates the, clock pulses for the DEMUX. Although this circuit works satisfactorily, the phase-shift circuit is fixed to create a 90° shift at only one frequency, and, therefore, minor shifts in input frequency will produce clock pulses whose timing is not perfectly accurate. In most systems where frequency variations are not great, the circuit operates reliably.With clock pulses of the proper frequency, some means is now needed to synchronize the multiplexer channels. This is usually done with a special synchronizing pulse that is applied to one of the input channels at the transmitter.

In our example of a four-channel system, only three actual signals would be transmitted. The fourth channel would be used to transmit a special pulse whose characteristics would be unique in some way so that it could be easily recognized. The amplitude of the pulse may be higher than the highest-amplitude data pulse, or the width of the pulse may be wider than those pulses derived by sampling the input signals. Special circuits can men be used to detect the synchronizing(sync) pulse. Fig. 9.26 shows an example of a sync pulse that is higher in amplitude than the maximum pulse value of any data signal. The sync pulse is also the last to occur in the frame. At the receiver, a comparator circuit is used to detect the sync pulse. One input to the comparator is set to a dc reference voltage equal to slightly higher than the maximum amplitude possible for the data pulses.

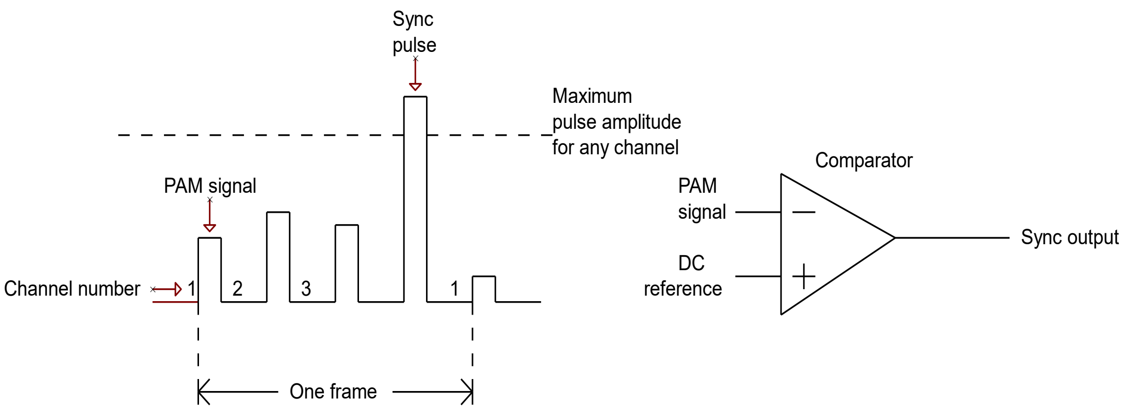


Figure 9.26 Frame Sync Pulse and Comparator Detector

When a pulse occurs that is greater than this amplitude, the comparator will generate an output pulse. The only time this occurs is when the sync pulse occurs. The output of the comparator will be a pulse. This pulse can then be used for synchronization purposes.Another method of providing sync is not to transmit a pulse during one channel interval. This leaves a blank space in each frame. This blank space can then be detected and used for synchronizing purposes. Such a circuit for doing this is illustrated in Fig. 9.27. Here the PAM multiplex signal is applied to an amplifier-limiter as it was in the clock circuits. The output is a series of pulses occurring at the pulse repetition rate of the PAM signal. However, since no pulse is transmitted during the sync interval, a blank space occurs. The resulting signal is inverted, so the blank space actually appears as a wider pulse, as shown in Fig. 9.27.

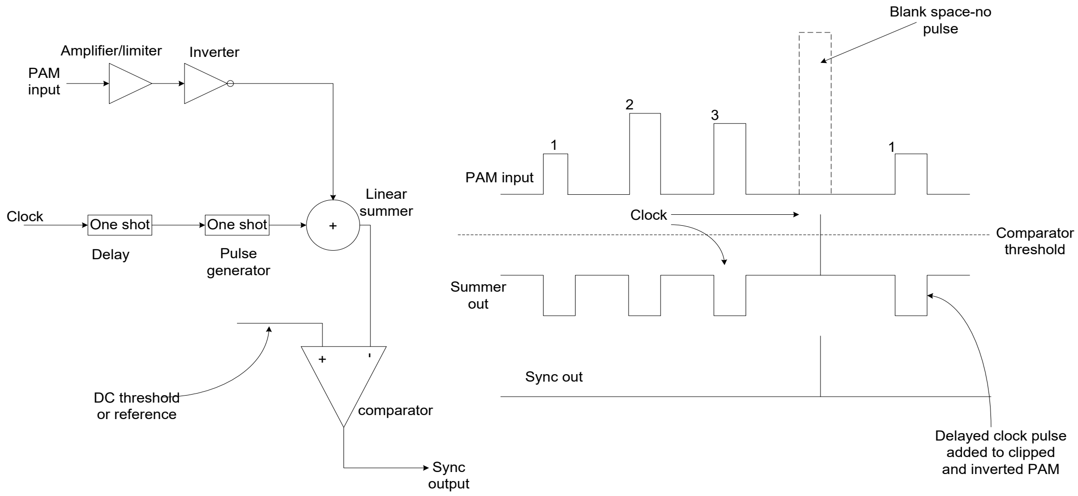


Figure 9.27 A PAM Sync Detector Circuit

This signal is then added to a series of narrow clock pulses. These pulses are delayed for one-half the clock period before they are linearly added to the clipped PAM signal. This delay is accomplished with a one-shot multivibrator whose pulse at the same frequency but at a predetermined width. These are then added in a linear resistive circuit to form the composite shown. The output of the linear mixer is applied to a comparator. The comparator threshold is set so that only the clock pulse added to the blank portion of the original signal is passed. Again, this synchronizing signal is used by the DEMUX.

As indicated, the sync pulse is usually the last one transmitted within a given frame. This sync pulse, when detected at the receiver, is used as a reset pulse for the counter.' in the DEMUX circuit. At the end of each frame, the counter is reset to zero, meaning that channel 0 is selected. Now when the next PAM pulse occurs, the DEMUX will be set': to the proper channel. Clock pulses then step the counter in the proper sequence for demultiplexing.Finally, at the output of the DEMUX, separate low piss filters are applied to each channel to recover the original analog signals. Fig. 9.28 shows the complete PAM DEMUX.

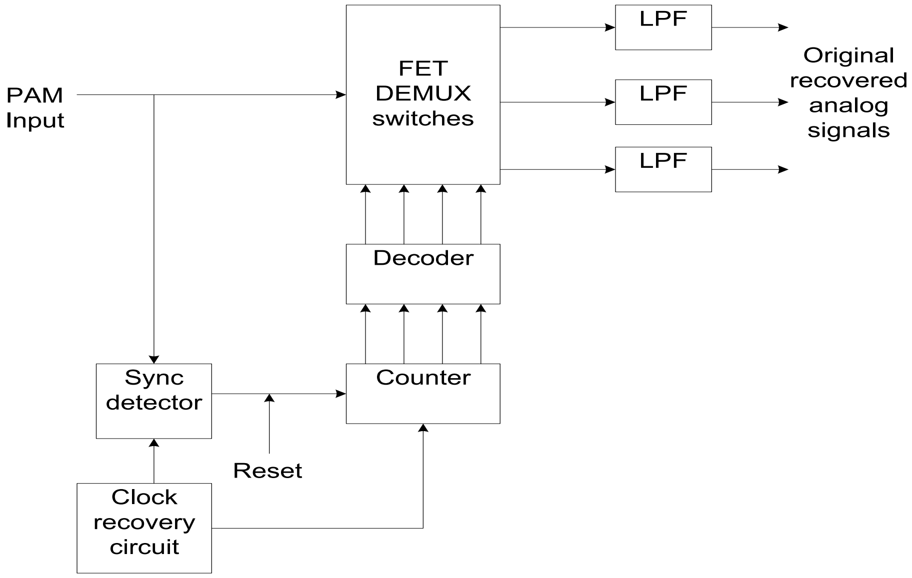


Figure 9.28 Complete PAM Multiplexer

Pulse-Code Modulation

The most popular form of pulse modulation used in TDM systems is pulse-code modulation (PCM). Pulse-code modulation is a form of digital modulation in which the code refers to a binary word that represents digital data. Multiple channels of serial digital data are transmitted with TDM by allowing each channel a time slot in which to transmit one binary word of data. The various channel data are interleaved and transmitted sequentially. Instead of transmitting a single pulse whose amplitude is the same as that of the analog signal being sampled, in PCM a binary number representing the amplitude of the analog waveform at the sampling point is transmitted. As this statement implies, analog signals may be transmitted by PCM. The analog signal is sampled as in PCM and is then converted into digital format by an analog-to-digital (ADC). The ADC converts the analog signal into a series of binary numbers where each number is proportional to the amplitude of the analog signal at the various sampling points. These binary words are converted from parallel to serial format and are then transmitted.

At the receiving end, the various channels are demultiplexed and the original sequential binary numbers are recovered. These are usually stored in a digital memory and then transferred to a digital-to-analog converter (DAC) which reconstructs the analog signal. Of course, the original data may be strictly digital in format, in which case no D/A conversion are required.

Pulse-code modulation systems allow the transmission of any form of digital data regardless of what it represents. Pulse-code modulation is used in telephone systems to transmit analog voice conversations, binary data for use in digital computers, and even video data. Most long distance space probes such as the Mariner and Voyager have on-board video cameras whose output signals are digitized and transmitted back to earth in binary format. Such PCM video systems make possible the transmission of pictures over incredible distances.

Multiplexing: Fig. 9.29 shows a general block diagram of the major components in a PCM system. We will assume that analog voice signals are the initial inputs. These are applied to ADCs as shown. The output of each ADC is an 8-bit parallel binary word. Since the digital data must be transmitted serially, the ADC output is fed to a shift register that produces a serial data output from the parallel input. The clock oscillator circuit driving the shift register operates at the desired frequency.

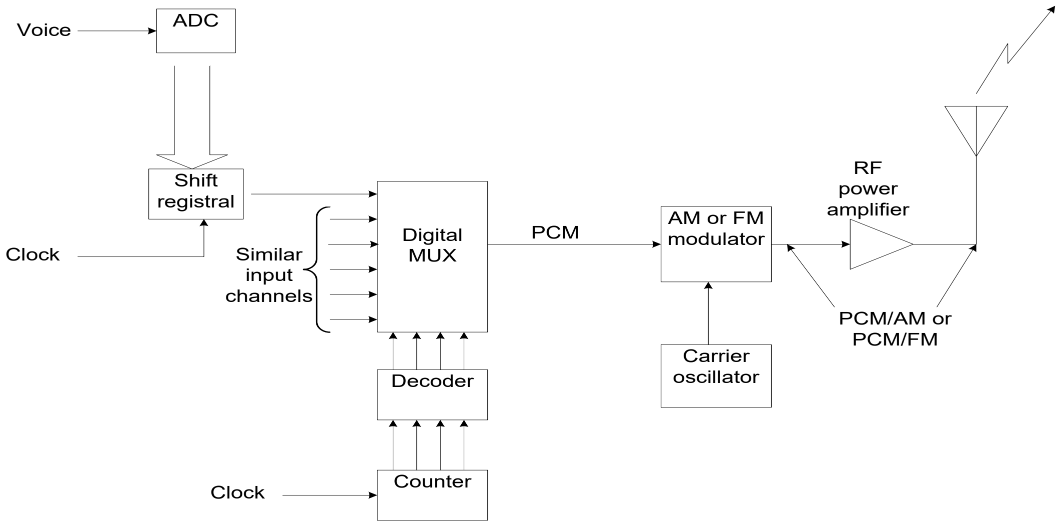


Figure 9.29 A PCM System

The multiplexing is done with a simple digital multiplexer. Since all the signals to be transmitted are binary in nature, a multiplexer constructed of standard AND or NAND gates can be used. A binary counter drives a decoder that selects the desired input channel.The multiplexer output is a serial data waveform of the interleaved binary words. This binary signal is used to modulate a carrier. Either AM or FM may be used in typical systems. This creates either PCM/AM or PCM/FM. The output of the modulator is then fed to a transmitter for radio communications or can otherwise be transmitted by wire or fiber optic cable. Additional levels of modulation may also be used.

Demultiplexing: At the receiving end of the communications link, the RF signal is picked up by a receiver and then demodulated. Refer to Fig. 9.30. The original serial PCM binary waveform is recovered. This is fed to a shaping circuit, such as a Schmitt trigger, to clean up and rejuvenate the binary pulses. The original signal is then demultiplexed. This is done with a digital DEMUX using AND or NAND gates. The binary counter and decoder driving the DEMUX are kept in step with the receiver through a combination of clock recovery and sync-pulse detector circuits similar to those used in PAM systems. The demultiplexed serial output signals are fed to a shift register for conversion to parallel data and are then sent to a DAC followed by a low-pass filter. The result is a very accurate reproduction of the original voice signal.

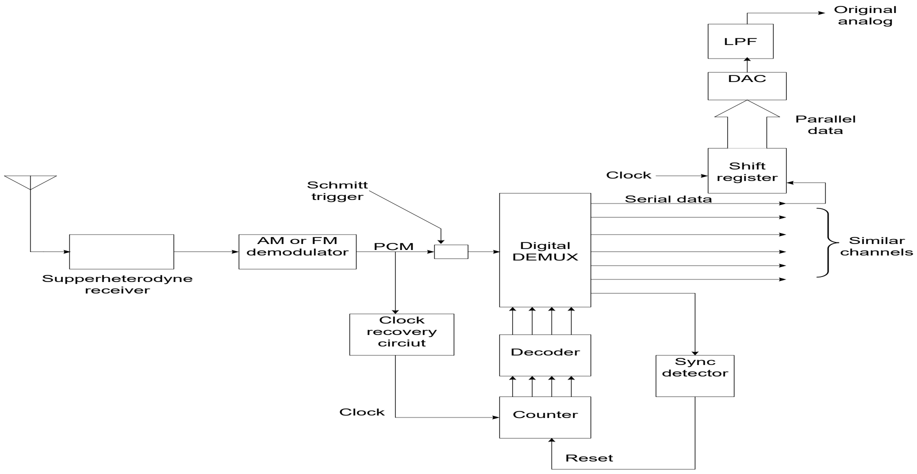


Figure 9.30 A PCM Receiver-Demultiplexer

9.8 Sample/Hold Circuit

Quantizingis the name given to the process of translating amplitude samples of an analog waveform into a binary code word. Quantizing is really the same thing as A/D conversion. You will also hear the term digitizing used to designate the same process. The first step in the quantizing process is virtually identical to that in PAM. The analog signal is sampled at periodic intervals. This is done us described previously; with a gate. Another method of sampling is to use a sample and hold (S/H) amplifier. An S/H amplifier, also called a track/store circuit, accepts the analog input signal and passes it through, unchanged, during its sampling mode. In the hold mode, the amplifier remembers a particular voltage level at the instant of sampling. The output of the S/H amplifier is a fixed dc level whose amplitude is the value at the sampling time.

Fig. 9.31 shows a simplified drawing of an S/H amplifier. A high-gain dc differential op-amp is the basic element. The amplifier is connected as a follower with 100 percent feedback. Any signal applied to the non-inverting(+) input will be passed through unaffected. The amplifier has unity gain and no inversion. A storage capacitor is connected across the very high input impedance of the amplifier. The input signal is applied to the storage capacitor and the amplifier input through a MOSFET gate. A depletion mode MOSFET is normally used.This MOSFET acts as an on/off switch, when the gate is at 0V the transistor tin us on, acting as a very low resistance and connecting the input signal to the amplifier. The charge on the capacitor follows the input signal. This is the sample or track mode for the amplifier. The output is simply equal to the input. When the gale voltage is made positive with a pulse, the MOSFET cuts off, in this mode, it acts as an open switch.

During the sample mode, the charge on the capacitor and the op-amp output simply follows the input signal. When the S/H control signal goes high, the transistor is rut off. The charge on the capacitor remains. The very high input impedance of the amplifier allows the capacitor to retain the charge for a relatively long period of lime. The output of the S/H amplifier then is the dc voltage value of the input signal at the instant the S/H control pulse switches from low (sample) to high (hold). It is this voltage that is applied to the ADC for conversion into a proportional binary number.The primary benefit of an S/H amplifier is that it stores the analog voltage during the sampling interval. In some high-frequency signals, the analog voltage may actually change during the sampling interval, either increasing or decreasing. This is an undesirable condition since it will confuse the A/D converter and introduce some error. The S/H amplifier, however, stores the voltage on the capacitor, which remains constant during the sampling interval and thus ensures more accurate quantizing.

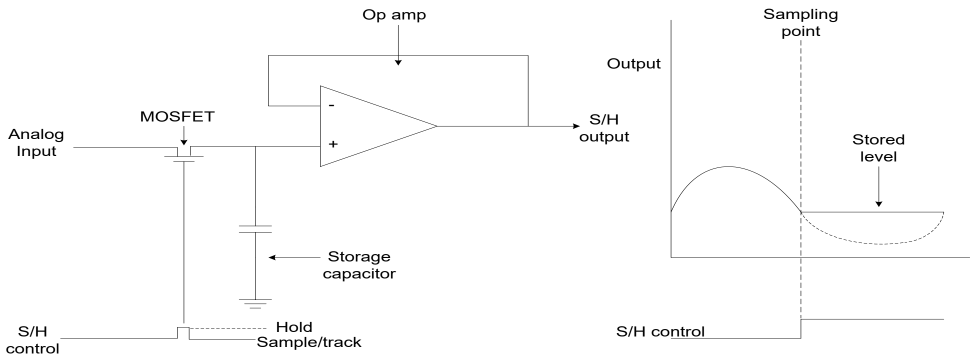


Figure 9.31 Am S/H Amplifier

9.8.1 A/D Conversion

In the quantizing process, we are effectively dividing the total analog signal amplitude range into a number of equal amplitude increments. Each one of these increments will be represented by a specific binary code. For example, assume that the total analog amplitude voltage range is 0 to 15 V. We could represent each voltage increment by a 4-bit binary number where 0V was represented by 0000 and 15 V was represented by 1111. During the sampling of the analog wave, the amplitudes of the samples can assume any one of a number of infinite values between 0 and 15 V. In the quantizing process, each of those values will be converted into an even or integer value. For example, one of the analog samples may be 9.2 V. This is closest to the integer 9 and, therefore, the 9.2 V value will be represented by the value 9 in binary form, or 1001. An analog value of 12.7 V might be represented as the integer 13, or 1101. As you can see, the quantizing process introduces some error. This is called quantization. The result is that the analog signal is somewhat distorted by the process. The quantized analog signal is only an approximation of the real thing.

Although 15 levels provide only crude quantization, improved representations of the analog signal can be used by providing more quantizing, increments. The greater the number of individual voltage increments or levels provided in the quantizer, the more closely the analog signal can be approximated. It has been determined that the range of voice amplitude levels in the telephone system is approximately 1000 to 1. In other words, the largest amplitude voice peak is approximately 1000 times the smallest voice signal. This voltage ratio of 1000:1 represents a 60 dB range. If a quantizer with 1000 increments were used, very high quality analog signal representation would be achieved. For example, an ADC with a 10-bit word can represent 1024 individual levels. A 10-bit ADC would provide excellent signal representation. If the maximum peak audio voltage were 1V, then the smallest voltage increment would be one-thousandth of this or 1 mV.

In practice, it has been found that it is not necessary to use this many quantizing levels for voice. In most practical PCM systems, a 7 or 8 bit ADC is used for quantizing. One popular format is to use an 8bit code where 7 bits represent 128 amplitude levels and the eighth bit designates polarity (0 = +, 1 = -). Overall, this provides 256 levels, one half positive, and the other half negative.

Companding

As indicated earlier, the analog voltage range of a typical voice signal is approximately 1000 to 1. It turns out, however, that lower-level signals predominate. Most of the conversations take place at a normal low level. Therefore, the upper end of the quantizing scale is not often used. It may be reached during momentary peaks of loud talking, shouting, or emotional outbursts, but for most general conversations, the lower level signals will be more typical. Since most of the signals are low level, the quantizing error will be larger. In other words, the smallest increment of quantization becomes a larger percentage of the lower-level signal. It is a smaller percentage of the peak amplitude value, of course, but that is irrelevant when the signals are much lower in amplitude. The increased quantizing error can produce garbled or distorted sound.

In addition to increased quantizing error, low level signals are also more susceptible to noise. Noise represents random spikes or voltage impulses added to the signal. The result is static that interferes with the low level signals and makes intelligibility difficult. The most common means of overcoming the problems of quantizing error and noise is to use a process of signal compression and expansion known as companding. At the transmitting end, the voice signal to be transmitted is compressed. That is, its dynamic range is decreased. The lower-level signals are emphasized, and the higher-level signals are de-emphasized. This compression can take place prior to quantizing. But in some systems, companding is accomplished digitally in the analog-to-digital convener (ADC) by having unequal quantizing steps, small ones at low levels and larger ones at higher levels.

At the receiving end, the recovered signal is fed to an expander circuit that does the opposite, de-emphasizing the lower-level signals and emphasizing the higher-level signals, thereby returning the transmitted signal to its original condition. Companding greatly improves the quality of the signal being transmitted. One type of compression circuit is a non-linear amplifier that amplifies lower-level signals more than it does upper-level signals. Figure 8.15 shows a graph illustrating the companding process.

The curve shows the relationship between the input and output of the compander. Note that at the lower input voltages, the gain of the amplifier is high and produces high output voltages. As the input voltage increases, the curve begins to flatten, producing proportionately lower gain. The nonlinear curve compresses the upper-level signals while bringing the lower-level signals up to higher amplitude. Such compression greatly reduces the dynamic range of the audio signal. Instead of an amplitude ratio of approximately 1000:1, compression reduces this to approximately 60:1. The actual degree of compression, of course, can be controlled by carefully designing the gain characteristics of the compression amplifier. Thus the 60 dB voice range is reduced to around a 36 dB range.

In addition to minimizing quantizing error and the effects of noise, compression also lowers the dynamic range so that fewer binary bits are required to digitize the audio signal. A 64:1 voltage ratio could be easily implemented with a 6-bil ADC. In practice, a 7-bit ADC is used. Fig. 9.32 shows a simplified diagram of a compression amplifier.

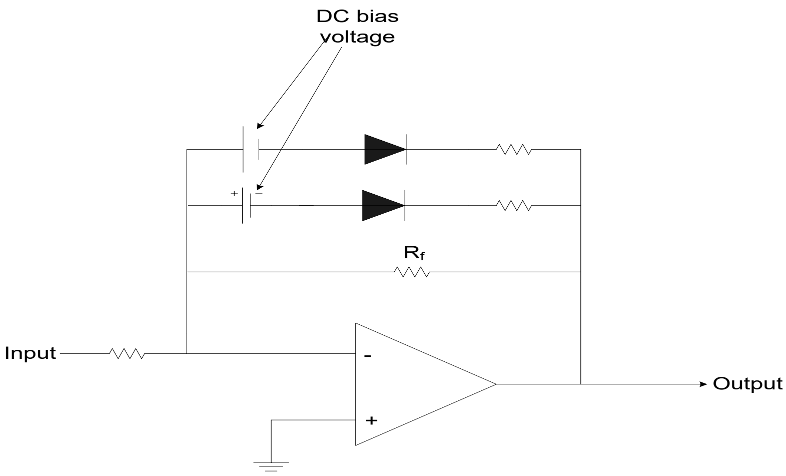


Figure 9.32 Compression Amplifier

Two biased diodes and their resistors are connected as the feedback elements in an op-amp. Although only two diodes are shown, typically multiple diodes each biased to a different voltage are used. At the lower-level signals, the amplifier output is basically linear. But as the input signals yet larger, the amplifier output will grow larger and at some point the diodes will begin to turn on. When the diodes turn on, they begin to shunt the feedback resistor R_F and thereby reduce the amplifier gain. Once the audio signal has been compressed, it is then applied to the ADC for quantizing.

At the receiving end, the digital signals are translated into analog signals. An analog signal is then passed through an expander amplifier that performs the opposite function of the compressor. A typical circuit is shown in Fig. 9.33. Again, biased diodes are used to shape the amplifier gain curve. Here the biased diodes are used on the input resistance to an op-amp.

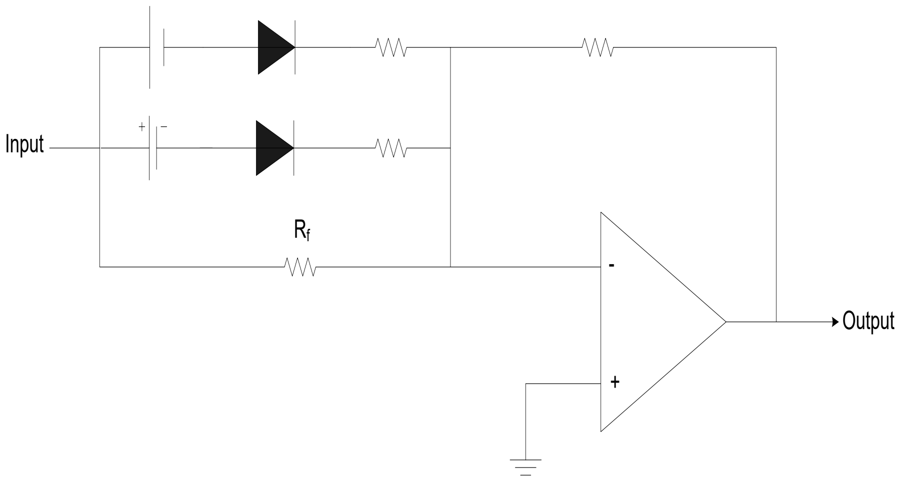


Figure 9.33 Expansion Amplifier

Fig. 8.15 shows the variation of the output voltage with respect to the input voltage during expansion. The lower-level signals are amplified less than the higher-level signals. If the compression and expansion curves are equal and opposite, the result will be a highly accurate reproduction of the original transmitted signal. In Fig. 9.33, as the input, increases, the output increases up to a certain level. Beyond that point, the diodes cut in, shunting the input resistor R_i thereby increasing the gain of the circuit in steps at the higher-input levels.

There are two basic types of companding used in telephone systems. One is called the μ-law compander and the other is called the A-law compander. The μ-law compander has a slightly different compression and expansion curve than the A-law compander. The μ-law compander is used in telephone systems in the United States and Japan, and the A-law compander is used in European telephone networks. The two are incompatible, but conversion circuits have been developed to make them compatible. These circuits convert μ-law to A-law and vice versa. According to international telecommunications regulations, those using μ-law are responsible for the conversions.

Although some companding is analog in nature as described, digital companding is far more widely used. The most common method is to use a nonlinear ADC. These converters provide a greater number of quantizing steps at the lower levels than they do at the higher levels, providing compression. On the receiving end, a matching nonlinear DAC is used to provide the opposite, compensation, expansion effect.

9.9 Chapter Review Problems

9.1 Briefly explain TDM. Why is it use in digital communication system? Sketches will be required for explanation'.

9.2 i Explain with the aid of a diagrams, the concept of pre-emphases and de-emphases.
ii. Explain with the aid of diagrams and wave form, the action of a semi-conductor diode as a simple detector.

9.3 (a) Explain briefly
(i) Pulse Modulation
(ii) Pulse Amplitude Modulation and
(iii) Pulse Width Modulation

(b) Explain briefly
(i) Pulse Code Modulation.
(ii) Quantization
(iii) Coding
(iv) Companding

9.4 Prior to sampling a ____________ is used to attenuate the high frequency components of the signal that lies outside the band of interest.

9.5 (a) Explain the Frequency Division Multiplexing with diagram. Give one example.
(b) Explain the Time Division Multiplexing with diagram. Give one example.
(c) Explain the Wavelenght Division Multiplexing with diagram. Give one example.

9.6 (a) Differentiate between Frequency Division Multiplexing and Time Division Multiplexing
(b) If a TDM Multiplexer was connecting four 64 kbps circuit to a remote site what bandwidth would be required for the remote link.
(c) Explain why digital system cannot use Frequency cannot use Frequency Division Multiplexing.

9.7 i. Explain with the aid of a diagrams, the concept of pre-emphases and de-emphases.
ii. Explain with the aid of diagrams and wave form, the action of a semi-conductor diode as a simple detector.

CHAPTER 10

DETECTORS

10.0 Demodulators

A *demodulator* is a circuit that accepts a modulated signal and recovers the original modulating information. Also know as a *detector,* a demodulator circuit is the key circuit in any radio receiver. In fact, the demodulator circuit may be used alone as the simplest form of radio receiver.

10.1 Diode Detector

The simplest and most widely used amplitude demodulator is the diode detector shown in Fig. 10.1. The AM signal is usually transformer-coupled as indicated. It is applied to a basic

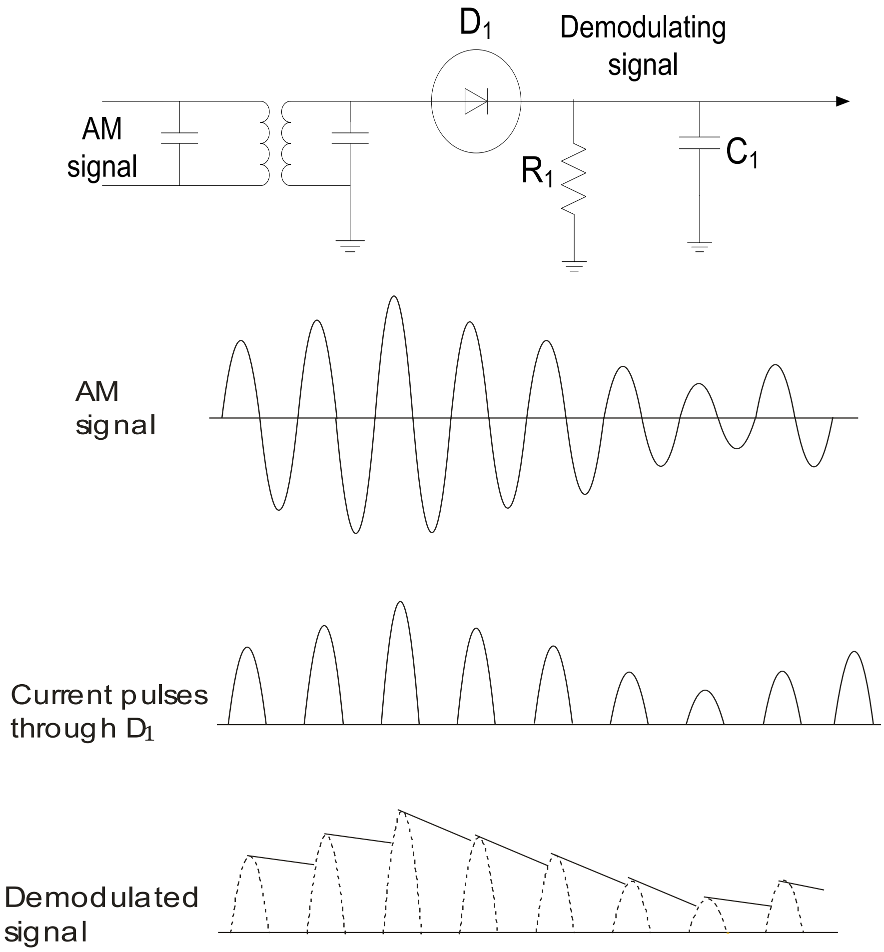


Figure 10.1 The diode detector Am Demodulator.

half-wave rectifier circuit consisting of and D_1 and R_1The diode conducts when the positive half, cycles of the AM signals occur. During the negative half cycles, the diode is reverse-biased and no current flows through it. As a result, the voltage across R_1 is a series of positive pulses whose amplitude varies with the modulating signal.
To recover the original modulating signal, a capacitor is connected across resistor R_1. It's value is critical to good performance, The value of this capacitor is carefully chosen so that it has a very low impedance at the carrier frequency. At the frequency of the modulating signal, it has a much higher impedance. The result is that the capacitor effectively shorts or filters out the carrier, thereby leaving the original modulating signal.
Another way to look at the operation of the diode detector is to assume that the capacitor charges quickly to the peak value of the pulses passed by the diode. When the pulse drops to zero, the capacitor retains the charge but discharges into resistor R_1. The time constant of C and R_1 is chosen to be long compared to the period of the carrier. As a result, the capacitor discharges only slightly during the time that the diode is not conducting. When the next pulse comes along, the capacitor again charges to its peak value. When the diode cuts off, the capacitor will again discharge a small amount into the resistor. The resulting waveform across the capacitor is a close approximation to the original modulating signal. Because of the capacitor
Although most technicians are employed in service jobs, some are employed to assist engineers in the development and testing of new products or equipment.
Charging and discharging, the recovered signal will have a small amount of ripple on it. This causes distortion of the demodulated signal. However, because the carrier frequency is usually many times higher than the modulating frequency, these ripple variations are barely notice able. In Fig. 10.1, the ripple is quite pronounced because the carrier frequency is low. The output of the detector is the original modulating signal. Because the diode detector recovers the envelope of the AM signal, which is the modulating signal, the circuit is sometimes, referred to as an *envelope detector.*

10.1.1 The Crystal Radio

The basic diode detector circuit is really a complete radio receiver in its own right. In fact, this circuit is the same as that used in the *crystal radio receivers* of the past. The crystal refers to the diode. In Fig. 10.2, the diode detector circuit is redrawn, showing an antenna connection and headphones. A long wire antenna picks up the radio signal, which is inductively coupled to the tuned circuit. The variable capacitor C_1 is used to select a station. The diode detector D_1 recovers the original modulating information which causes current flow in the headphones. The headphones serve as the load resistance, whereas capacitor C_l removes the carrier. The result is a simple radio receiver with very weak reception because no amplification is provided. Typically a germanium diode is used because its voltage threshold is lower than that of a silicon diode and permits reception of weaker

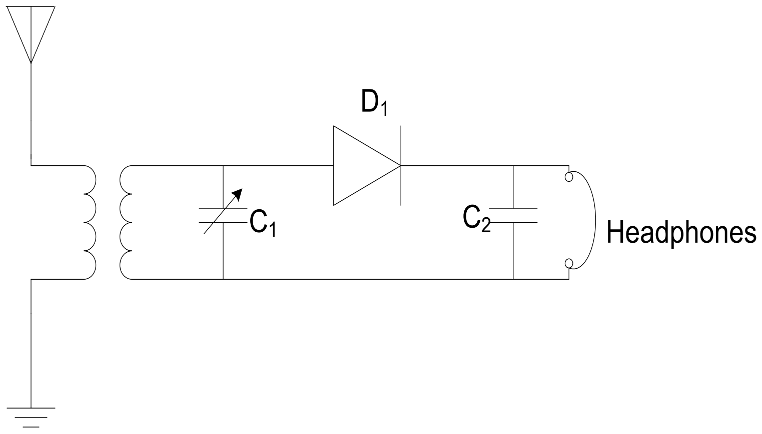


Figure 10.2 A crystal radio receiver.

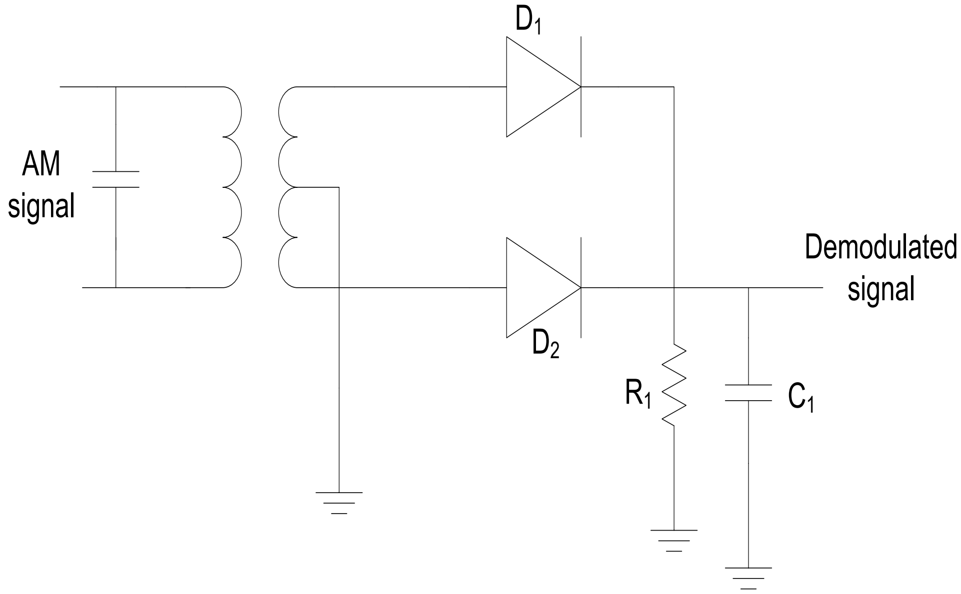


Figure 10.3 A full-wave diode detector for AM signals.

signals. Such a receiver can easily be built to receive standard AM broadcasting stations. The performance of the basic diode detector can he improved by using a full-wave rectifier circuit as shown in Fig. 10.3. Here, two diodes and n center-lapped secondary on the RF transformer are used to form a standard full-wave rectifier circuit. With this arrangement, diode D_l will conduct on the positive half cycle, while D_2 will conduct on the negative half cycle. This diode detector produces a higher average output voltage which is much easier to filter. The capacitor value necessary to remove the carrier can be half the size of the capacitor value used in a half-wave diode detector. The primary benefit of this circuit is that the higher modulating frequencies will not be distorted by ripple or attenuated as much as in the half-wave detector circuit

10.2 Frequency Demodulators

There are literally dozens of circuits used to demodulate or detect FM and PM signals. The well-known *Foster-Seeley discriminator* and the ratio detector were among the most widely used frequency demodulators at one time, but today these circuits have been replaced with more sophisticated 1C demodulators. Of course, they are still found in older equipment. The most widely used detectors today include the pulse-averaging discriminator, the quadrature detector, and the phase-locked loop. In this section, we will take a look at all these widely used demodulator circuits.

10.2.1 Foster-Seeley Discriminator

One of the host frequency demodulators is the Foster-Seeley discriminator shown in Fig. 10.4. The FM signal is applied lo the primary of the RF transformer T_1. The primary and secondary windings are resonated at the carrier frequency with C_1 and C_2. The parallel tuned circuit in the primary of T_1 is connected in the collector of a limiter amplifier Q_1that removes amplitude variations from the FM signal.

The signal across the primary of T_1 is also passed through capacitor C_3 and appears directly across an RFC. The voltage appearing across the RFC is exactly the same as that appearing across the primary winding simply because C_3 and C_5 are essentially short circuits at the carrier frequency. The voltage across this RFC is designated V_3.

The current flowing in the primary winding of T_1 induces a voltage in the secondary winding. Because the secondary winding is center-tapped, the voltage across the upper portion V_1 will be 180° out of phase with the voltage across the lower portion V_2. The voltage induced into the secondary winding is 90° out of phase with the voltage across the primary winding. When both the primary and the secondary windings of an air-core transformer are tuned resonant circuits, the phase relationship between the voltages across the primary and secondary will be 90°. This means that the voltages V_1 and V_2 will also be 90° out of phase with V_3, the voltage across the RFC. This phase relationship is shown in the vector diagrams in Fig. 10.4. In Fig. 10.4(a), the input is the unmodulated earlier frequency.

The remainder of the circuit consists of two diode detector circuits similar to those used for AM detection. The voltage $V_{1\text{-}3}$ applied to D_1, R_l. and C_4 is the sum of voltages $V_1$and V_3 The voltage $V_{2\text{-}3}$ applied to D_2. R_2, and C_s is the sum of voltages V_2 and V_3. Since voltages V_1 and V_2 are out of phase with voltage V_3, their respective sums $V_{1\text{-}3}$ and $V_{2\text{-}3}$, are vector sums, as illustrated in Figure10.4(a).

On one half cycle of the primary voltage, D_1 conducts and current flows through R_1 and

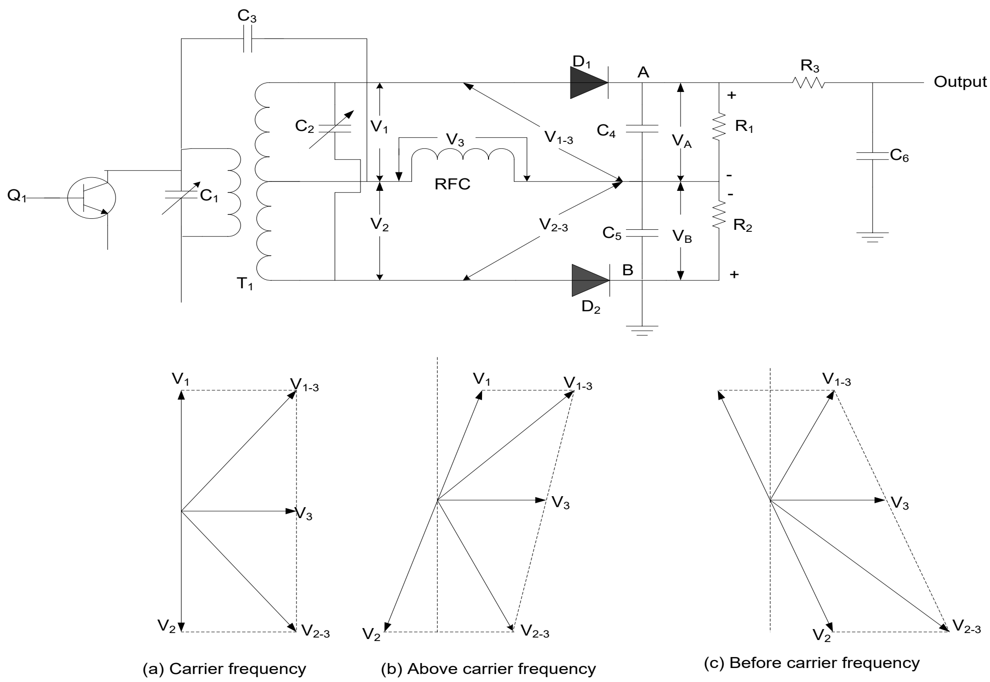


Figure 10.4 The Foster-Seeley Discriminator.

charges C_4. On the next half cycle, D_2 conducts and current flow through R_2 and charges C_5 the voltages across R_1 and R_2, designated V_A and V_B, are identical because $V_{1\text{-}3}$ and $V_{2\text{-}3}$ are the same. Since these two voltages are equal but of the opposite polarity, the voltage between point A and ground is zero. At the carrier center frequency with no modulation, the modulator output is therefore zero.

The secondary of T_1 and capacitor C_2 form a series resonant circuit. The reason for this is that the voltage induced into the secondary appears in series with that winding. At resonance, the inductive reactance of the secondary winding equals the capacitive reactance of C_2. At that time, the current flowing in the circuit is exactly in phase with the voltage induced into the secondary. The output is derived from a portion of the secondary winding which is an inductor. The voltage across the secondary winding, therefore, is 90° out of phase with the current in the circuit.

Should the input frequency change as would be the case with FM, the secondary winding will no longer be at resonance. For example, if the input frequency increases, the inductive reactance will be higher than the capacitive reactance, making this circuit inductive. This causes the phase relationship between V_1 and V_2 to change with respect to V_3 If the input is above the resonant frequency, then V_1 will lead V_3 by a phase angle less than 90°. Since V_1 and V_2 remain 180° out of phase, V_2 will then lag V_3 by an angle of more than 90°. This change in phase relationship is shown in chapter 6 and 7 respectively. When the new vector sums of V_1 and V_3 and $V_2$and V_3 are computed, it is found that the voltage $V_{1\text{-}3}$ applied to D_1 is greater than the voltage $V_{2\text{-}3}$ applied to D_2. Therefore, the voltage across R_1 will be greater than the voltage across R_2 and the net output voltage will be positive with respect to ground.

If the frequency deviation is lower than the center frequency, voltage V_1 will lead by an angle of more than 90° while voltage V_2 will lag by an angle less than 90°. The resulting vector additions of V_1 and V_3 and V_2 and V_3 are shown in Fig. 10.4*(c).* This lime, $V_{2\text{-}3}$ is greater than $V_{1\text{-}3}$. As a result, the voltage across R_2 will be greater than the voltage across R_1 and the net output voltage will be negative.

As you can see, as the frequency deviates above and below the center frequency, the output at point A increases or decreases

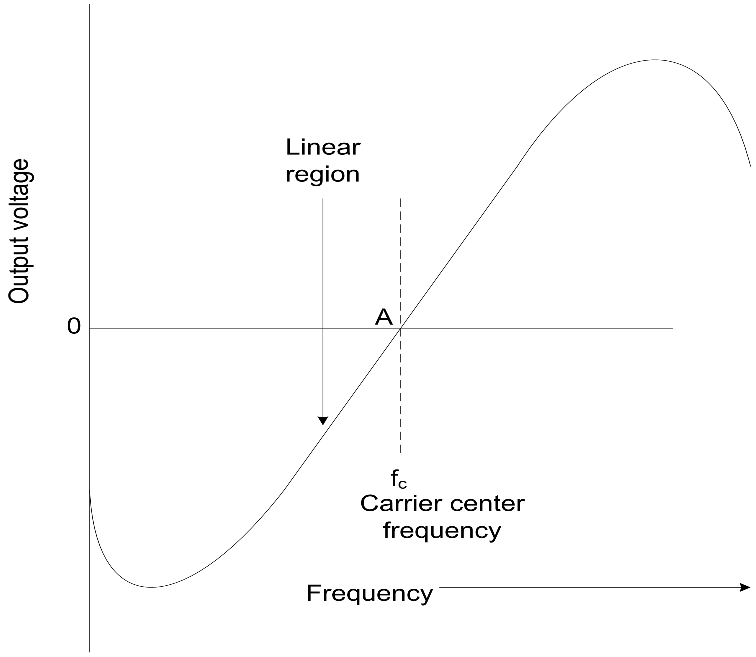


Figure 10.5 Output Voltage of The Discriminator.

therefore, the original modulating signal is recovered. This signal is passed through the de-emphasis network consisting of R_3 and C_6. The output is then applied to an amplifier or to other circuits as required. Fig. 10.5 shows the output voltage occurring across point A and ground with respect frequency deviation. It is the linear range in the center that produces the accurate reproduction of the original modulating signal.

The discriminator circuit is sensitive to input amplitude variations. Higher inputs will produce a greater amount of signal at the output, whereas lower inputs create less output. These output variations will be interpreted as frequency changes, and the output will be an incorrect reproduction of the modulating signal. For that reason, all amplitude variations must be removed from the FM signal before being applied to the discriminator. This is usually handled by limiter circuits that will be discussed later.

10.2.2 Ratio Detector

Another widely used demodulator is the *ratio* detector *detector.* It is similar in appearance to the discriminator but has several important differences as is shown in Fig. 10.6. The FM signal is applied to the RF transformer T_l with its center tapped secondary. The FM signal is also passed through capacitor C_3 and applied across the RFC as in the discriminator. The circuit uses two diodes, but note that the direction of D_2 is reversed from that in the discriminator. The voltages V_{1-3} and V_{2-3} applied to D_1 and D_2, respectively, are again a composite of V_1 and V_3 (V_{1-3}) and V_2 and V_3 ($V_{2/3}$) as before.

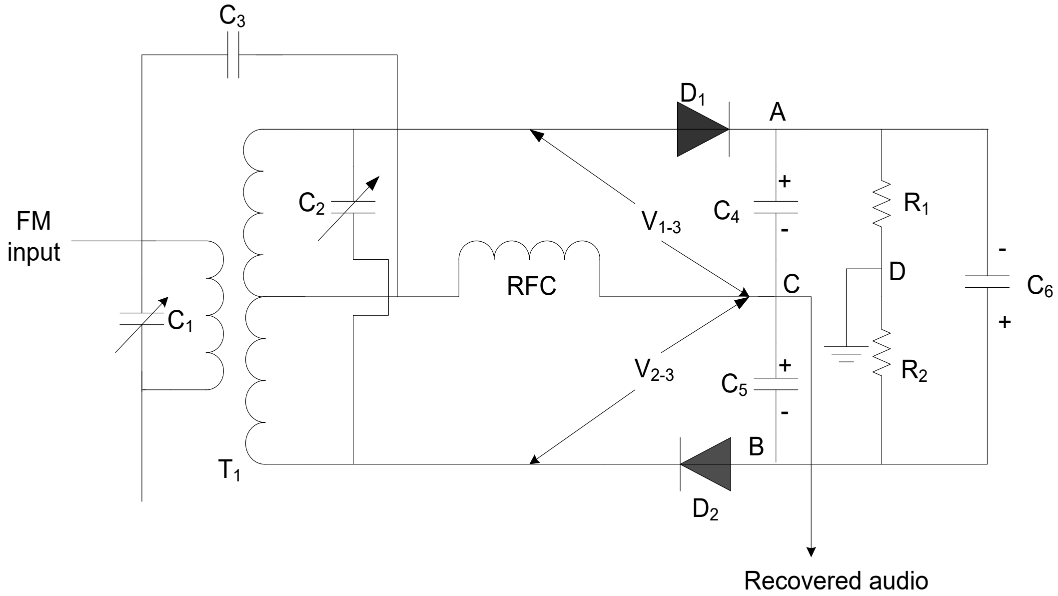


Figure 10.6 A Ratio Detector.

Another major difference in the radio detector is the use of a very large capacitor C_6 connected across the output. The load resistors R_1 and R_2 are equal in value, and their common connection is at ground. The output is taken from between points C and ground in the circuit. Capacitors C_4 and C_5 and resistors R_1 and R_2 form a bridge circuit. The voltage across capacitors C_4 and C_5 is the bridge input voltage, while the output is taken between points C and D.

With no modulation on the carrier, the voltage V_{1-3} applied to D_1 is the same us the voltage D_{2-3} applied to D_2. Therefore, capacitors C_4 and C_5 charge to the same voltage with the polarity shown. Since C_6 is connected across these two capacitors, it will charge to the sum of their voltages. Because C_6 is a very large capacitor, usually tantalum or electrolytic, it takes several cycles of the input signal for the capacitor to charge fully. However, once it charges, it will maintain a relatively constant voltage. Since R_1 and R_2 are equal, their voltage drops will be equal. Also, the voltage drops across C_4 and C_5 are equal. The bridge circuit, therefore, is balanced. Looking between points C and D, you will see 0 V because the potential is the same.

Assume that at the center earner frequency, the voltage drops across C_4 and C_5 are each *2* V. This means the charge on C_6 is 4 V. Then the voltages across R_1 and R_2 are each *2* V.

If the frequency increases, the phase relationship in the circuit will change as described previously fertile discriminator circuit [Fig. 10.4 (a)]. This will cause the voltage across C_4 to be greater than the voltage across C_5. Assume that the voltage across C_4 is 3 and the voltage across C_4 is 3 V. The voltages across R_1 and R_2 remain the same at 2 V each becausethe charge on C6 does not change. The bridge is now unbalanced, and an output voltage will appear between points C and D in the circuit. Using point B as a reference, the voltage at point C is 1 V positive, and the voltage across R_2 is 2 V positive. Therefore, the voltage difference at C is — I V.

If the frequency decreases, then the phase relationship will be such that the charge on C_5 will be greater than the charge on C_4. If the voltage across C_5 is +3V with respect to B and the voltage across R_2 remains 2 V, then point C is $+$ 1 V. The bridge is unbalanced, but in the opposite direction, and the output voltage is of the opposite polarity.

The primary advantage of the ratio detector over the discriminator is that it is essentially insensitive to noise and amplitude variations. The reason for this is the very large capacitor C_6 .Since it takes a long time for this capacitor to charge or discharge, short noise pulses or minor amplitude variations are totally smoothed out. However, the average dc voltage across C_6 is the same as the average signal amplitude. This voltage can, therefore, be used in automatic gain control applications.

As indicated earlier, the ratio detector and Foster-Seeley discriminator are no longer widely used because they are difficult to implement in integrated-circuit form. Besides, the quadrature demodulator and PLL offer far superior performance for comparable cost.

10.2.3 Pulse-Averaging Discriminator

A simplified block diagram of a *pulse-averaging discriminator* is shown in Fig. 10.7. The FM signal is applied to a zero-crossing detector or a clipper/limiter which generates a binary voltage

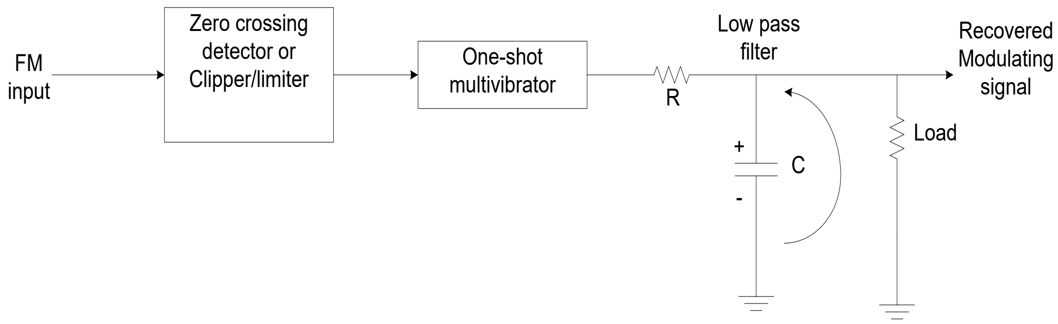


Figure 10.7 A Pulse-Averaging Discriminator for Frequency Demodulation.

level change each lime the FM signal varies from minus to plus, or plus to minus. The result is a rectangular wave containing all (he frequency variations of the original signal but without amplitude variations.

The FM square wave is then applied to a one-shot multivibrator. The one-shot or monostable multivibrator generates a fixed-amplitude, fixed-width dc pulse on the leading edge of each FM cycle. The duration of the one shot is set so that it is less than the period of the highest frequency expected during maximum deviation.

The one-shot output pulses are then fed to a simple *RC* low-pass filter that averages the dc-pulses to recover the original modulating signal.

The waveforms for the pulse-averaging discriminator are shown in Fig. 10.8. Note how at the low frequencies, the one-short pulses are widely spaced. At the higher frequencies, the one-shot pulses occur very close together. When these pulses are applied to the averaging filter, a dc output voltage is developed. The amplitude of this dc voltage is directly proportional to the frequency deviation.

When a one-shot pulse occurs, the capacitor in the filter charges to the amplitude of the

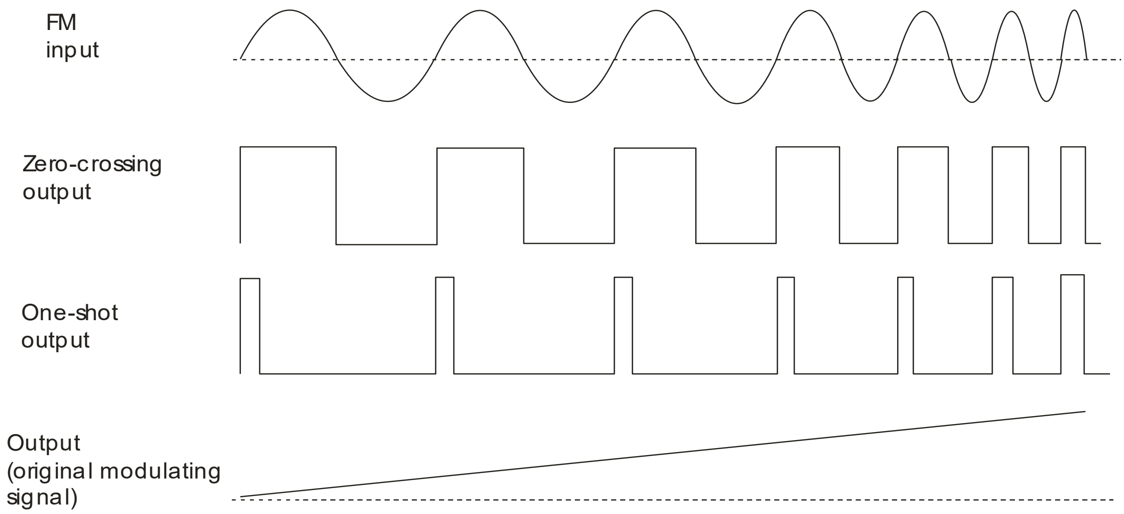


Figure 10.8 Waveforms for the Pulse-Averaging Discriminator.

pulse. When the pulse turns off, the capacitor discharges into the load. If the *RC* time constant is high, the charge on the capacitor will not decrease much. If the lime interval between pulses is long, however, the capacitor will lose some of its charge into the load, and so the average dc output will be low. When the pulses occur rapidly, the capacitor has little time to discharge between pulses, and therefore, the average voltage across it remains higher. As you can see, the filter output voltage varies in amplitude with the frequency deviation. Note in Fig. 10.8 how the voltage rises linearly as the frequency increases. The original modulating signal is developed across the filter output. The filler components are carefully selected to minimize the ripple caused by charging and discharging of the capacitor while at the same time providing the necessary high-frequency response for the original modulating signal.

The pulse-averaging discriminator is a very high quality frequency demodulator. Prior to the availability of low-cost ICs, its use was limited to expensive telemetry and industrial control applications. Today, the pulse-averaging discriminator is easily implemented with low-cost
ICs and has, therefore, found its way into many electronic products.

10.2.4 Quadrature Detector

Another popular frequency demodulator is the *quadrature detector. It* primary application is in TV audio demodulation. Although it is also used in some FM radio systems. Most 1C frequency demodulation are of the quadrature type. *The term quadrature* refers to a 90° phase shift between two signals. The quadrature detector uses a phase-shift circuit to produce a phase shift of 90° at the unmodulated carrier frequency. The most commonly used phase-shift arrangement is shown in Fig. 10.9. The FM signal is applied through a very small capacitor C1 to the parallel tuned circuit which is adjusted to resonance at the center carrier frequency. The tuned circuit appears as a high value of pure resistance at resonance. The small capacitor has a very high reactance compared the tuned circuit impedance. The output across the tuned circuit then al the earner frequency is very close to 90° leading the input. Now, when FM occurs, the carrier frequency will deviate above and below the resonant frequency of the tuned circuit. The result will be an increasing or decreasing amount of phase shift between the input and the output.

The two quadrature signals are then fed to a phase detector circuit. The phase detector is nothing more than a circuit whose output is a function of the amount of phase shift between two input signals. The most commonly used phase detector is a balanced modulator using differential amplifiers. The output of the phase detector is a series of pulses whose width varies with the amount of phase shift between the two signals. These signals are averaged in an *RC* low-pass filter to re-create the original modulating signal.

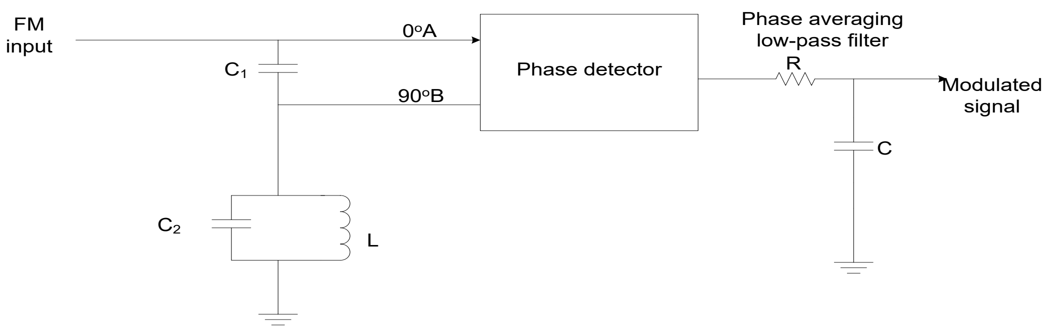


Figure 10.9 A quadrature FM detector

Normally the sinusoidal FM input signals to the phase detector arc at a very high level. Therefore, they will drive the differential amplifiers in the phase detector into cutoff and saturation. The differential transistors will act as switches, so the output will be as a series of pulses. No limiter is needed if the input signal is large enough. The output pulse duration is determined by the amount of phase shift. You might think of the phase detector as simply an AND gate whose output is on only when the two input pulses are on and is off if either one or both of the inputs is off.

Fig. 10.10 shows the typical waveforms involved in a quadrature detector. When there is no modulation, the two input signals are exactly 90° out of phase and, therefore, provide an output pulse width as indicated. When the frequency increases, the amount of phase shift decreases and. therefore, causes the output pulse width to be greater. The wider pulses averaged by the *RC* filter produce a higher average output voltage. This, of course, corresponds to the higher amplitude required to produce the higher carrier frequency.

When the FM signal frequency decreases, there is more phase shift and, as a result, the output pulses will be narrower. When averaged, the narrower pulses will produce a lower average output voltage, thus corresponding to the original lower modulating signal.

10.2.5 Differential Peal Detector

Another integrated-circuit FM demodulator is the *differential peak detector.* It is found in TV sets and other consumer electronic products. Usually the demodulator is only one of many other circuits on the chip. A general schematic diagram of the circuit is shown in Fig. 10.11(a). The circuit is an enhanced differential amplifier. Transistors Q_3 and Q_4 form the

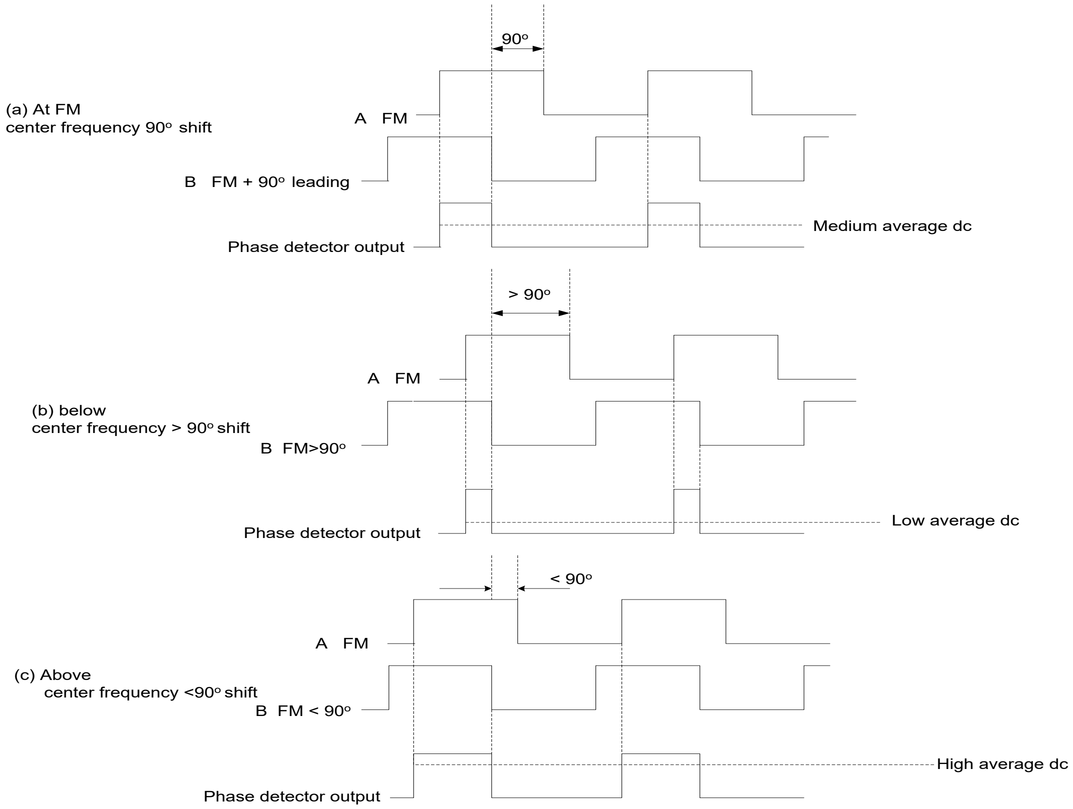


Figure 10.10 Waveforms in the Quadrature Detector.

differential amplifier, and the other transistors are emitter followers. Transistor Q_7 is a current source. Note that a tuned circuit is connected between the bases of Q_1 and Q_6. L_1, C_1 and C_2 are discrete components that are external to the chip. The FM input is applied to the base of Q_1. A typical input is a 4.5 MHz FM sound carrier in a TV set.

Q_1 and Q_6 are emitter followers that provide high input impedance and power amplification to drive two other emitter followers, Q_2 and Q_5. Q_2 and Q_5, along with on-chip capacitors C_a, and C_b, are peak detectors. C_a and C_b, charge and discharge as the voltages at the two inputs vary. For example, assume that the input is a carrier sine wave. When the input to Q_1 goes positive, the voltage at the emitter of Q_2 also goes positive, and C_a charges to the peak ac value that appears across the tuned circuit. When the peak of the sine wave is past, the voltage declines to zero and then reverses polarity for the negative half-cycle of the sine wave. As the voltage drops, capacitor Ca retains the positive peak value as a charge. The input impedance to Q_3 is relatively high and so does not significantly discharge Ca. Capacitor C_a simply acts as a temporary storage cell for the peak voltage that cycle.

Capacitor C_b, on Q_5 also forms a peak detector circuit. It too will charge up to the peak of the input from the opposite end of the tuned circuit. If the 4.5 MHz carrier is unmodulated, the two capacitors will charge up to the same value. These equal voltages appear at the inputs to the differential amplifier. Remember that the output of a differential amplifier is

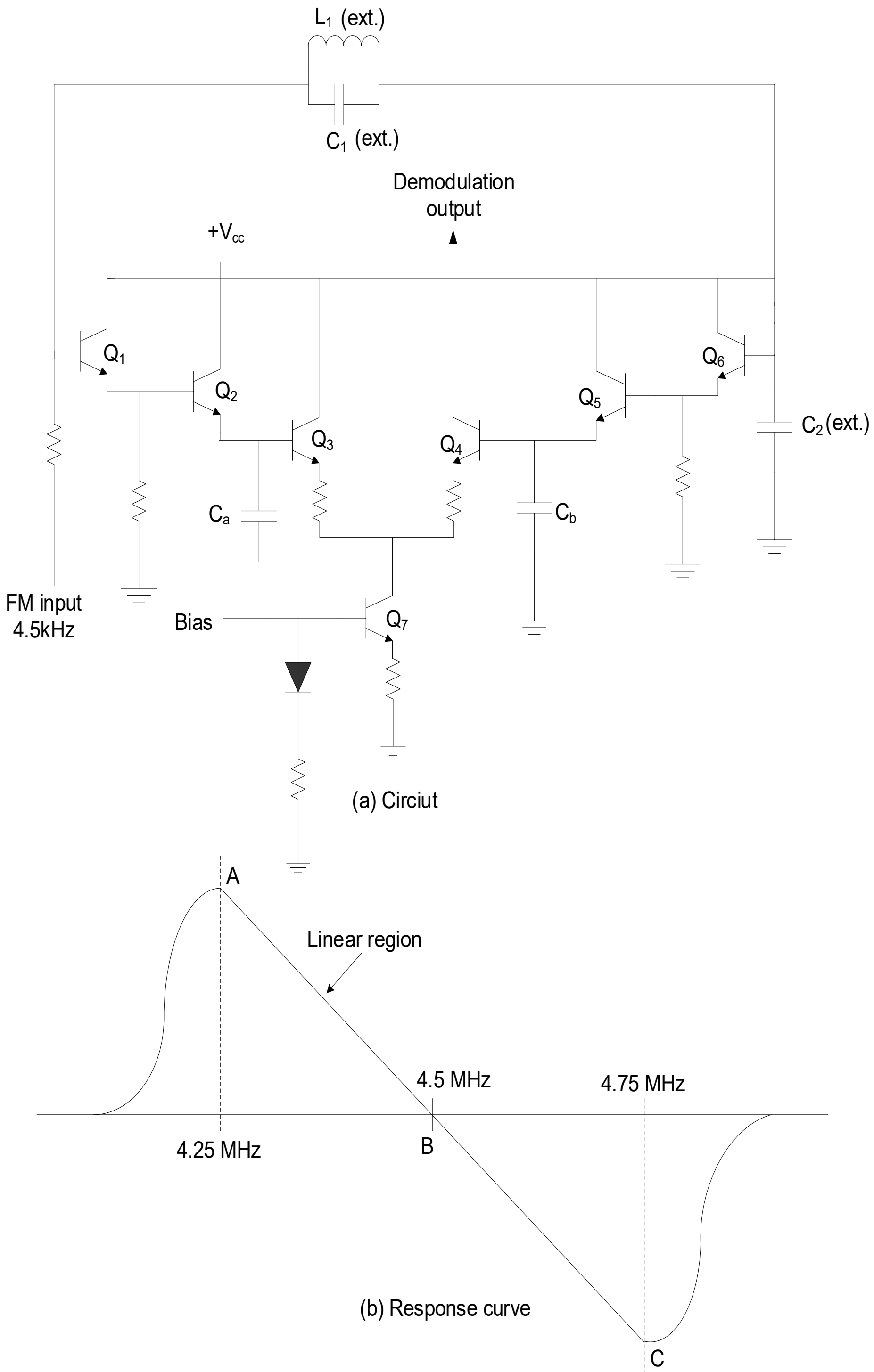


Figure 10.11 A Differential Peak Detector FM Demodulator.

proportional to the difference of the two inputs. If the inputs are equal, the difference is zero, and so is the output. The output current from Q_4 when this condition exists is used as the zero reference. A lower or higher current will represent a varying frequency.

The external tuned circuit is designed so hat L_1 forms a series-resonant circuit with C_2 it a frequency lower than the center frequency, usually 4.5 MHz. A typical series-resonant frequency is 4.25 MHz. L_1 forms a parallel resonant circuit, with C_1 at a frequency above the center frequency. This frequency is usually 4.75 MHz.

To see how the circuit works, assume that an input is gradually increasing in frequency from below 4.25 MHz to above 4.75 MHz. At very low deviation frequencies, the effect on the tuned circuit is minimal and the inputs to the differential amplifiers are about equal; therefore the output is near zero. Now, as the input frequency increases, at some point it reaches
the series-resonant point of L_1 and C_2. When this occurs, the total impedance of the series circuit is resistive and very low. This effectively shorts the input of Q_1 to ground, making it near zero. At the same time, the voltage across C_2 is at its peak value because of resonance, so the input to Q_6 is maximum. The result is a very high output current from Q_4. This condition is illustrated at point A in Fig. 10.11*(b).*

As the input frequency continues to rise, at some point the center frequency of 4.5 MHz is reached, at which time the output will be zero as described earlier. This is shown at point B in Fig. 10.11(b). The frequency continues to increase. At some point, L_1 will be parallel resonant with C_1. This produces a very high impedance at the base of Q_1, so the voltage is maximum at this time. The lower reactance of C_1 at this higher frequency makes the input to Q_6, very low. As a result, the difference between the two inputs is very high and the output current at Q_4 becomes very high in the opposite (negative) direction. This is point C in Fig. 10.11(b). As you can see, the variation in output current from point A to point C is very linear as the frequency increases. This change in output amplitude with a change in frequency results in excellent demodulation of FM signals.

10.2.6 Phase-Locked Loop Demodulator

The best frequency demodulator is the phase-locked loop (PLL). A *PLL* is a frequency- or phase-sensitive feedback control circuit. It is used not only in frequency demodulation but also in frequency synthesizers, as you will see in a later chapter, and in various filtering and signal detection applications.

All PLLs have the three basic elements illustrated in Fig. 10.12. A phase detector or mixer is used to compare the input or reference signal to the output of a VCO. The VCO frequency is varied by the dc output voltage from a low-pass filter. It is the output of the phase

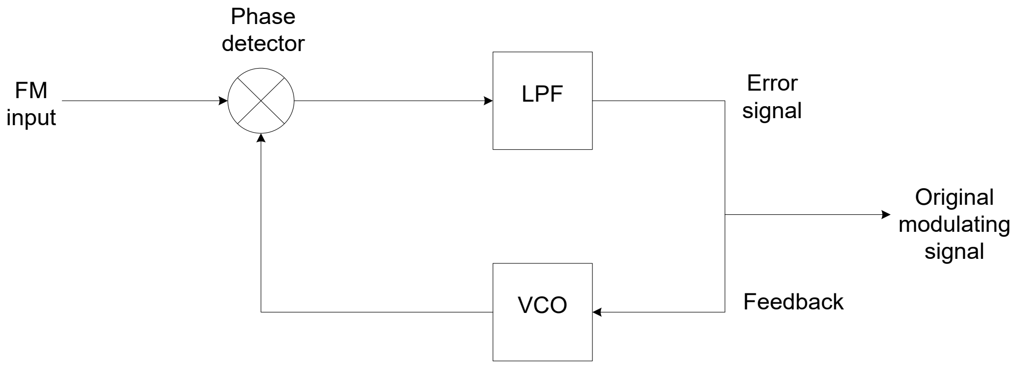


Figure 10.12 The Elements Of A Phase-Locked Loop (PLL).

detector that the low-pass filter uses to produce the dc control voltage.

The primary job of the phase detector is to compare the two input signals and generate an output signal that when filtered will control the VCO. If there is a phase or frequency difference between the input and VCO signals, the phase detector output will vary in proportion. The filtered output will adjust the VCO frequency in an attempt to correct for the original frequency or phase change. This dc control voltage is called the *error signal* and is also the feedback in this circuit.

To examine the operation of the PLL, assume initially that no input signal is applied. At this time, the phase detector and low-pass filter outputs are zero. The VCO then operates at what is called its *free-running frequency.* This is the normal operating frequency of the VCO as determined by its internal frequency-determining components.

Now assume that an input signal close to the frequency of the VCO is applied. The phase detector will compare the VCO free-running frequency to the input frequency and produce nit output voltage proportional to the frequency difference. The resulting dc voltage is applied to the VCO. This dc voltage is such that it forces the VCO frequency to move in a direction that reduces the dc error voltage. The error voltage forces the VCO frequency to change in the direction that reduces the amount of phase or frequency difference between the VCO and the input. At some point, the error voltage will eventually cause the VCO frequency to equal the input frequency. When this happens, the PLL is said to be in a *locked condition.* Although the input and VCO frequencies are equal, there will be a phase difference between them that produces the dc output voltage which causes the VCO to produce the frequency that will keep the circuit locked.

If the input frequency changes, the phase detector and low-pass filter will produce a new value of dc control voltage that will force the VCO output frequency to change until it is equal to the new input frequency. Because of this action, the PLL is said to track the input signal. Any input frequency variation will be matched by a VCO frequency change so that the circuit remains locked.

The VCO in a PLL, therefore, is capable of tracking the input frequency over a wide range. The range of frequencies over which the PLL will track the input signal and remain locked is known as the locked range. The lock range is usually a band of frequencies above and below the free-running frequency is out of the lock range, the PLL will not lock. When this occurs, the VCO output frequency jumps to its free-running frequency.

If an input frequency within the lock range applied to the PLL, the circuit will immediately adjust itself and remain in a locked con-lion. The phase detector will determine the phase difference between the VCO free-running and input frequencies and generate the error signal that will force the VCO to equal the input frequency. Once the input signal is captured, the PLL remains locked and will track any changes in the input signal as long as it remains within the lock range.

The range of frequencies over which the PLL will capture an input signal is known as the *capture range.* It is much narrower than the lock range, but it is generally centered around the free-running frequency of the VCO as is lock range. See Figure 10.13.

The characteristic that causes the PLL to cap-s signals within a certain frequency range also causes it to act like a bandpass filter. Phase-locked loops are often used in signal conditioning applications where it is desirable to pass signals only a certain range and reject signals outside of range. The PLL is highly effective in eliminating the noise and interference on a signal.

Since the PLL responds to input frequency variations, naturally you can understand why it

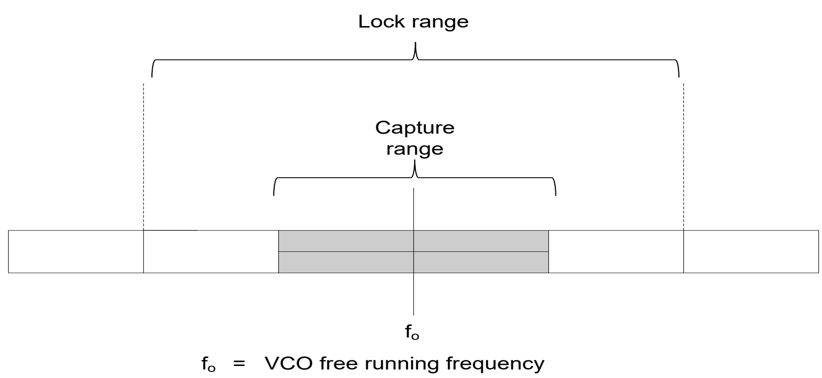


Figure 10.13 Capture and Lock Ranges of A PLL.

is useful in FM applications. If an FM signal is applied to the input, the VCO will truck it. The VCO is, in effect, a frequency modulator that produces exactly the same FM signal as the input In order for this to happen, however, the VCO input must be identical to the original modulating signal.

The VCO output follows the FM input signal because the error voltage produced by the phase detector and low-pass filter forces the VCO to track it. For that reason, the VCO output must be identical to the input signal if the PLL is to remain locked, and the error signal must be identical to the original modulating signal of the FM input. The low-pass filter cutoff frequency is designed such that it is capable of passing the original modulating signal.

The PLL is the best of all frequency demodulators in use because of the characteristics described earlier. Its ability to provide frequency selectivity and filtering give it a signal-to-noise ratio superior to any other type of FM detector. The linearity of the VCO ensures a highly accurate reproduction of the original modulating signal. Although PLLs are complex, they are easy to apply because they are readily available in low-cost 1C form.

10.3 Square Law Detector

A DSBC double side band plus carrier, signal may be express as:

$$f(t) = A(1 + mg(t))\ cos\,\omega_c t \qquad 10.1$$

Where g (t) = modulating signal

$m = modulation\ index\ (0 < m < 1)$

$\omega_c = 2\pi f_c$

$f_c = carrier\ frequency$

$A = amphtude\ of\ unmodulated\ carrier$

Let $g(\mathrm{t}) = \cos\omega_\mathrm{m}\mathrm{t}$

$\omega_m = 2\pi f_m$

f_m = frequency of the modulating signal

Then; $f(t) = \mathrm{A}(1 + \mathrm{m}\cos\omega_\mathrm{m}\mathrm{t})\cos\omega_\mathrm{c}\mathrm{t}$ 10.2

Which may be written as:

$$f(t) = A\,cos\,\omega_c t + A\,mcos\,\omega_m t\,cos\,\omega_c t$$

$$f(t) = A\,cos\,\omega_c t + A\,m\frac{1}{2}[cos(\omega_c + \omega_m)t + cos(\omega_c - \omega_m)t]$$

$$f(t) = A\,cos\,\omega_c t + A\,m\tfrac{1}{2}cos(\omega_c + \omega_m)t + Am\tfrac{1}{2}cos(\omega_c - \omega_m)t \qquad 10.3$$

10.3.1 Square Law Detection Of DSBC Signal

The output of a square – law detector is of the form $p = q^2$ 10.4

Where q → input signal

Semiconductor (solid state) divides have this type of response for small inputs. Substituting Eq (10.3) in Eq (10.4) yields.

$$f(t)^2 = [A\cos\omega_c t + \tfrac{Am}{2}\cos(\omega_m + \omega_C)t + \tfrac{Am}{2}\cos(\omega_m - \omega_c)t]^2 \qquad 10.5$$

$$f(t)^2 \left[A\cos\omega_c t + \tfrac{Am}{2}\cos(\omega_m + \omega_c)t + \tfrac{Am}{2}\cos(\omega_m - \omega_c)t\right] \times \left[A\cos\omega_c t + \tfrac{Am}{2}\cos(\omega_m + \omega_c)t + \tfrac{Am}{2}\cos(\omega_m - \omega_c)t\right] \qquad 10.6$$

$f(t)^2 \mathrm{A}^2 \cos^2 \omega_c t + \frac{A^2 m}{2} \cos \omega_c t \cos(\omega_m + \omega_c)t + \frac{A^2 m}{2} \cos \omega_c t \cos(\omega_m - \omega_c)t + \frac{A^2 m}{2} \cos(\omega_m + \omega_c)t \cos \omega_c t + \frac{A^2 m^2}{4} \cos^2(\omega_m + \omega_c)t + \frac{A^2 m^2}{4} \cos(\omega_m + \omega_c)t \cos(\omega_m - \omega_c)t + \frac{A^2 m^2}{4} \cos^2(\omega_m + \omega_c)t$

But from $\quad cos^2\theta = \frac{1}{2}\cos(2\theta + 1)$ 10.7

Substituting 10.7 into 10.8 yields 10.9

$f(t)^2 = \frac{A^2}{2}[\cos 2\omega_c t + 1] + \frac{A^2 m}{2} \cos \omega_c t \cos(\omega_m t + \omega_c t) + \frac{A^2 m}{2} \cos \omega_c t \cos(\omega_m t - \omega_c t) + \frac{A^2 m}{2} \cos(\omega_m t + \omega_c t) \cos \omega_c t + \frac{A^2 m^2}{8} [\cos 2(\omega_m t + \omega_c t) + 1] + \frac{A^2 m^2}{4} \cos(\omega_m t + \omega_c t) \cos(\omega_m t - \omega_c t) + \frac{A^2 m^2}{8} [\cos 2(\omega_m t - \omega_c t) + 1]$ 10.8

But $\cos A \cos B = \frac{1}{2}\{\cos(A + B) + \cos(A - B)\}$ 10.9a

$f(t)^2 = \frac{A^2}{2}[\cos 2\omega_c t + 1] + \frac{A^2 m}{4} \cos(\omega_m t + 2\omega_c t) + \frac{A^2 m}{4} \cos \omega_m t + \frac{A^2 m}{4} \cos 2\omega_C t + \frac{A^2 m}{4} \cos(\omega_m t + 2\omega_c t) + \frac{A^2 m}{4} \cos \omega_m t + \frac{A^2 m^2}{8} [\cos 2(\omega_m t + \omega_c t) + 1] + \frac{A^2 m^2}{8} \cos(2\omega_m t) + \frac{A^2 m^2}{8} \cos 2\omega_c t + \frac{A^2 m^2}{8} [\cos 2\,(\omega_m t + \omega_c t) + 1]$ 10.9b

The output of the square-law detector contains AF & RF components. After filtering out the latter we are left with: Eq (10.10)

$f(t)^2 = \frac{A^2 m}{4} \cos \omega_m t + \frac{A^2 m}{4} \cos \omega_m t + \frac{A^2 m}{4} \cos \omega_m t + \frac{A^2 m}{4} \cos \omega_m t + \frac{A^2 m^2}{4} \cos 2\omega_m t + \frac{A^2 m^2}{4} \cos 2\omega_m t$ 10.9c

$$f(t)^2 = \frac{4}{4}(A^2 m \cos \omega_m t) + \frac{2\,A^2 m^2}{8} \cos 2\omega_m t \qquad 10.9d$$

$$D(t) = mA^2 \cos \omega_m t + \frac{m^2}{4} A^2 \cos 2\omega_m t \qquad 10.10$$

The first term in Eq.10.10 resembles the modulation. The second term constitutes a distortion component.

Circuit operation: A solid-state diode rectifiers and RF signal and produces a DC voltage that is integrated by a capacitor. The diode not an ideal rectifier. The forward drop at high current is $0.3 - 0.7$ volts, depending on the semiconductor material see Fig. 10.14

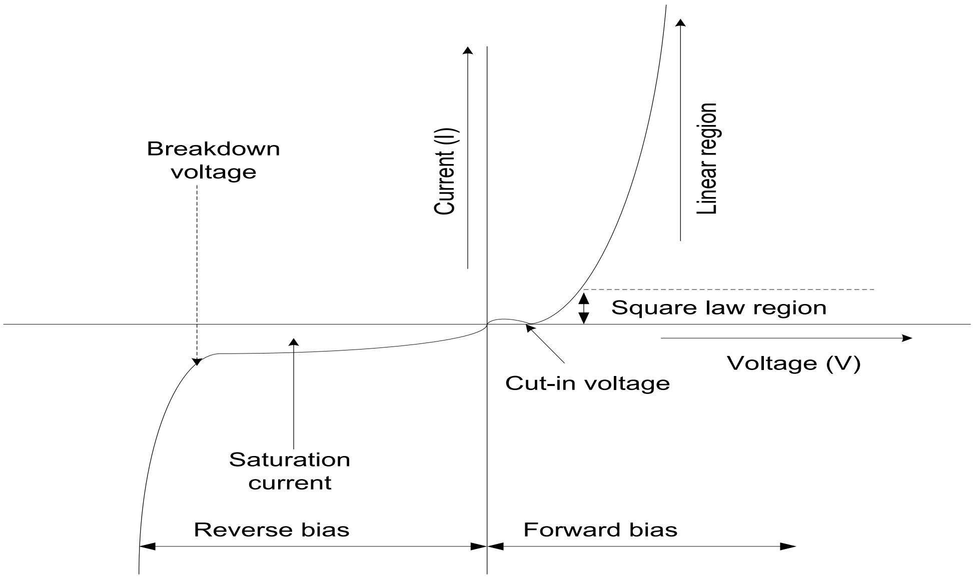


Figure 10.14 Diode Electrical Characteristics.

The rectification efficiency varies widely over the range of input voltage. At very low RF voltage, below the knee of the diode characteristics, the diode exhibits a square law response, where $V_0 = nP_i$ 10.11
In the square law region, the output voltage V_0 is proportional to the square of the input voltage V_i, thus V_o is proportional to the input power.

$$V_0 = nV_i^{\,2} = nP_i \qquad 10.11a$$

Pi is directly proportional to V_0, where n is the constant of proportionality.

At higher RF voltage, above knee, the output voltage is approximately the peak of the voltage, proportional to the square of the power. Consider the following circuit and the typical diode curve in the forward bias region. This looks like a half-wave rectifier circuit shown in Fig. 10.15

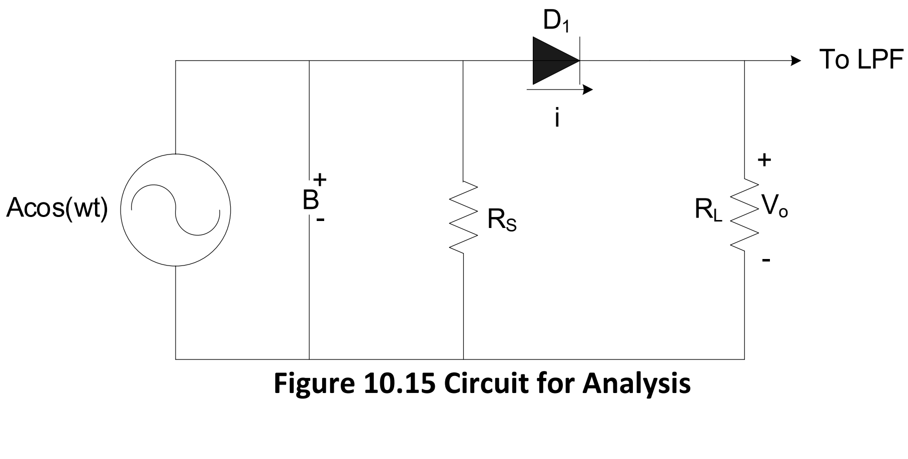


Figure 10.15 Circuit for Analysis

The following assumption should be considered

a. The diode has a forward bias applied so it is operating somewhere in the 'knee' of the curve. In this case, say about 500 nA. This operating point is still very non-linear.
b. $A\cos\omega t << V_{diode}$ (under these conditions, the signal is not rectifier)
c. The V-I curve of the series combination of D_1 and R_2 can be represented by a power series of the form i.e. $= ae + be^2 + ce^3 + \cdots \ldots\ldots\ldots\ldots$ 10.12

Since $v = A\cos\omega t$ 10.13

i_d may be represented by the polynomial as in Eq. 10.14

$i_d = aA\,cos(\omega t) + bA^2cos^2(\omega t) + cA^3cos^3(\omega t) + \cdots \ldots\ldots\ldots\ldots$ 10.14

Substituting in the familiar trigonometric identities we have;

$cos^2(\omega t) = ½[1 + cos\,(2\omega t)]\; and\; cos^3(\omega t) = ¼[\,cos(3\omega t) + 3\,cos(\omega t)]$ 10.15

Substituting Eq (10.14) into Eq (10.15), this gives us the Eq (10.16)

$$i_d = aA\,cos(\omega t) + \frac{bA^2}{2}[1 + cos\,(2\omega t)] + \frac{cA^3}{4}[cos\,(3\omega t) + 3\cos(\omega t) + \cdots \ldots\ldots..$$

$$= \frac{b}{2}A^2 + \left[aA + \frac{3}{4}cA^3\right]cos(\omega t) + \frac{b}{2}A^2cos(2\omega t) + \frac{C}{4}A^3cos(3\omega t)t + \cdots \ldots\ldots.. \quad 10.16$$

Which is of the form:

$\gamma_1A^2 + \gamma_2cos(\omega t) + \gamma_3\,cos(3\omega t) + \gamma_4\,cos(3\omega t) + \cdots \ldots\ldots....$ 10.17

where $\gamma_n$are constants. Since the voltage across $R_2(V_0)$ is what is of interest, and

$V_0 = i_dR_l$ 10.18

Substitute Eq (10.17) to Eq (10.18) we have Eq (10.19)

$V_0 = R_l\gamma_1A^2 + R_l\gamma_2cos(\omega t) + R_l\gamma_3\,cos(3\omega t) + R_l\gamma_4\,cos(3\omega t) + \cdots \ldots\ldots....$ 10.19

This represents a DC term all the harmonies of cos (ωt). By passing V_o through a low pass filter we get the DC output voltage.

$Vo = R_l\gamma_1A^2$

Remembering that A is the amplitude of RF signal, and that

$P = \frac{V^2}{R}$ 10.20

$P = \alpha V_0$ 10.21

Where α is a constant in other word, V_o is directly proportional to the power dissipated by R_s. The square land detector differs from linear diode detector because the applied input carrier voltage is of small magnitude and hence is restricted to the exclusively nonlinear portion of the dynamic characteristics, where as in linear diode detector, a large amplitude modulated carrier voltage is applied to the diode and most of the operation takes place over the linear region of the characteristic.

10.4 Product (Synchronous) Detection

If we multiply equation Eq (10.16) by the term $s(t) = B\cos w_c t \cos\omega_c$t, we obtain an expression for the output of a product or synchronous detector, whose schematic representation is given in Fig. 10.16

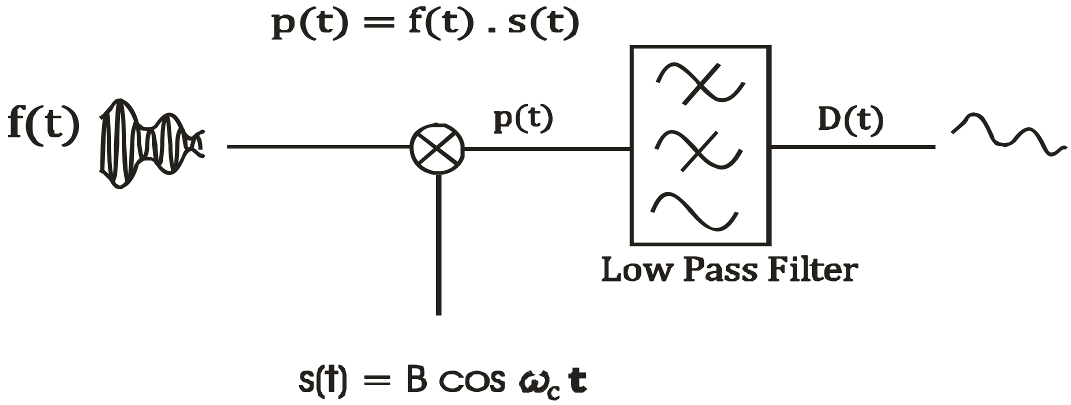


Figure 10.16 Product (Synchronous) Detector

The output of this type of detector is given by $P(t) = f(t). s(t)$, 10.22
that is

$$f(t). s(t) = ABcos^2(\omega_c t) + \frac{mAB}{2}\cos\omega_c t\cos(\omega_c + \omega)t + \frac{mAB}{2}\cos\omega_c t\cos(\omega_c - \omega_m)t \quad 10.23$$

Which may be written as:

$$f(t). s(t) = AB[1 + cos2(\omega_c t)] + \frac{mAB}{4}[\cos 2\omega_c + \omega_m)t + \cos\omega_m t] + \frac{mAB}{4}[\cos 2\omega_c - \omega_m)t + \cos\omega_m t] \quad 10.24$$

After filtering out RF components in Eq (10.24) we obtain Eq (10.25)

$$D(t) = \frac{mAB}{2}\cos\omega t \quad 10.25$$

and the distortion component of frequency $2\omega_m$ does not exist

10.5 Single Side Band (SSB) Detection

A SSB signal may be expressed by

$f(t) = A\, cos\, (\omega_c + \omega_m)t$ or $f(t) = A \cos(\omega_c - \omega_m)t$ 10.26

The plus sign is used for an upper side band (USB) signal and the minus sign for a lower side band (LSB) signal. Product and mixer-type demodulators are used for SSB detection. Consider a product detector with inputs.

$$f(t) = A\, cos(\omega_c + \omega_m)t \; and \; s(t) = B\, cos\omega_c t \qquad 10.27$$

The detector performs the mathematics function of Eq. 10.38

$P(t) = f(t).s(t) = AB\, cos\omega_c t.cos\,(\omega_c + \omega_m)t$ 10.28

This is identical to, Eq. 10.29

$P(t) = \frac{AB}{2}[cos(2\omega_c + \omega_m)t + cos\omega_m t]$ 10.29

If we filter out RF components in equ (10.29), we get the desired modulation signal given as $\frac{AB}{2}\cos\omega_m t$

10.6 Detection of DSBC Signals Using a Regenerative Receiver in the Oscillating Mode

Usually DSBC Signals are detected in regenerative receivers adjusting the regenerative gain slightly below the oscillation point, the nonlinearities of the active device being responsible for the demodulation process, usually of a square-law type. However, in the oscillating mode, DSBC Signal detection is also possible. When turned to the same frequency, a gently carrier. Both signals will be preset across the tank circuit and hence will be mixed by the active. If 'A' is the amplitude of unmodulated carrier and B is that of the oscillation across the turned circuit (having the same frequency as the carrier), then using eq (10.2) we get.

$$f(t) = A(1 + m\cos\omega_m t)\cos\omega_c t + B\cos\omega_c t \qquad 10.20$$

For the signals across the tank circuit. Eq (10.23) may be written as;

$f(t) = (A + B)\cos\omega_c t + \frac{mA}{2}\cos(2\omega_c + \omega)t + \frac{mA}{2}\cos(\omega_c + \omega_m)t$ 10.31

After square-law detection and filtering Eq (10.31), we are left with the recovered modulation term (ω_m) and unwanted distortion term; ($2\omega_m$)

$$D(t) = mA(A+B)\cos\omega_m t + \frac{m^2A^2}{4}\cos 2\omega_m t \qquad 10.32$$

Usually, harmonic distortion levels are tolerable.

10.7 Demodulation of SSB Signal in a Regenerative Receiver

Detection of SSB Signals requires the receiver to be working in the oscillating mode. Strong oscillations are needed so the receiver does not look to the incoming signal. Otherwise, the recovered audio will have the known 'quack-quack' type of sound. Alternatively, input signal can be alternated by the operator. Detection is again of the square law type, after a mixing process carrier out by the nonlinearities of the active device.

If $A\cos(\omega_c + \omega_m)t$ is the SSB signal and $B\cos(\omega_c + \Delta\omega)t$ the receivers' oscillation, both across the tank circuit, the mixers output will be:

$$P(t) = [A\cos(\omega_c + \omega_m)\,t + B\cos(\omega_c + \Delta\omega)t]^2 \qquad 10.33$$

Yielding eq (10.34) after filtering out RF frequency component.

$$D(t) = AB\cos(\omega_c + \Delta\omega)t \qquad 10.34$$

Carefully turning of the receiver will be necessary so that $\Delta\omega \to 0$. Also, the oscillation frequency should be very stable.

10.8 Demodulation of DSBSC Modulated Wave by Coherent Detection

The message signal $m(t)$ can be uniquely recovered from a DSBSC wave $s(t)$ by first multiplying $s(t)$ with a locally generated sinusoidal wave and then low pass filtering the product as shown.

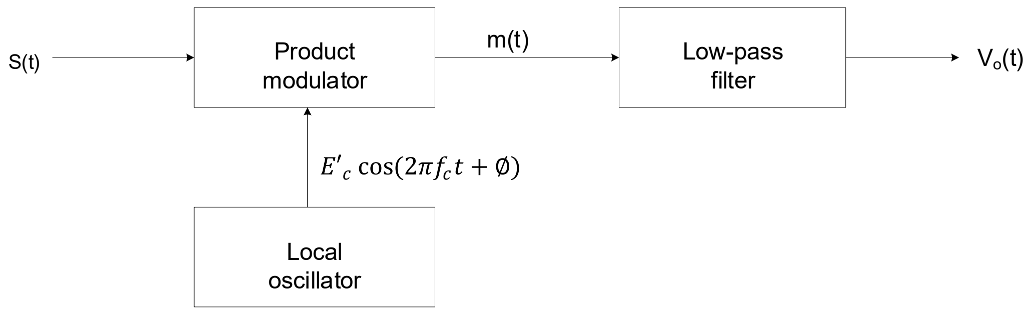


Figure 10.17 Coherent Detector

It is assumed that the local oscillator signal is exactly coherent or synchronized, in both frequency and phase, with then carrier wave $c(t)$ used in the product modulator to generate $s(t)$. This method of demodulation is known as coherent detection or synchronous detection.

The DSBSC wave $s(t)$ is applied to a product modulator in which it is multiplied with the locally generated carrier cos ($2\pi f_c t$) .

We assume that this locally generated carrier is exactly coherent or synchronized in both frequency and phase with the original carrier wave c(t) used to generate the DSB-SC wave.

This method of detection is therefore called as coherent detection or synchronous detection .

The output of the product modulator is applied to the low pass filter (LPF) which eliminates all the unwanted frequency components and produces the message signal

10.8.1 Analysis of Coherent Detection

Let $E'_c cos\ (2\pi f_c t + \emptyset)$ be the local oscillator signal, and $s(t)\ =\ E_c cos\ (2\pi f_c t) m(t)$be the DSBSC wave. Then the product modulator output v(t) is given by

$v(t)\ =\ E_c E'_c\ cos\ (2\pi f_c t)\ cos\ (2\pi f_c t + \emptyset)\ m(t)$ 10.35

$v(t)\ =\ \frac{E_c E'_c}{4} cos\ (4\pi f_c t + \emptyset)\ m(t)\ +\ \frac{E_c E'_c}{2} cos\ (\emptyset) m(t)$ 10.36

Let the output of the local oscillator be given by : $E_c cos\ (2\pi f_c t + \emptyset)$

Thus its amplitude is E'_c, frequency is f_c and the phase difference is arbitrary equal to Ø .

This phase difference has been measured with respect to the original carrier c(t) at the DSBSC generator .

PROOF:

If the output of the product modulator is given by:

$$m(t) = s(t).c'(t) \quad 10.37$$

$$s(t) = DSBSC\ input = x(t).E_c \cos(2\pi f_c t)$$

And $$c'(t) = Local\ carrier = \cos(2\pi f_c t + \emptyset)$$

Or $$m(t) = x(t).E_c \cos(2\pi f_c t) \cos(2\pi f_c t + \emptyset)$$

Therefore, $$m(t) = x(t).E_c \cos(2\pi f_c t + \emptyset) \cos(2\pi f_c t)$$

But from Eq.10.9a

Hence,

$$\cos(2\pi f_c t + \emptyset)\cos(2\pi f_c t) = \frac{1}{2}[\cos(2\pi f_c t + \emptyset + 2\pi f_c t) + \cos\emptyset]$$
$$= \frac{1}{2}[\cos(4\pi f_c t + \emptyset) + \cos\emptyset]$$

Thus,

$$m(t) = \frac{1}{2}x(t)E_c[\cos(4\pi f_c t + \emptyset) + \cos\emptyset]$$

Therefore,

$$m(t) = \frac{1}{2}x(t)E_c\cos\emptyset x(t) + \frac{1}{2}x(t).E_c\cos(4\pi f_c t + \emptyset) \qquad 10.38$$

The first term of equation 10.38 is scaled version of message signal $x(t)$ while the second term is known as an unwanted signal.

The Eq. 10.38 shows that the output of product modulator i.e., $m(t)$ consists of two terms. The first one represents the message signal x(t) with an amplitude of (1/2) $E_c\cos\emptyset$. Hence, this is the wanted term. The second term is an unwanted one .

Signal $m(t)$ is then passed through a low pass filter, which allows only the first term to pass through and will reject the second term .

Therefore, the filter output is given by :

$$v_0(t) = \frac{1}{2}x(t).E_c\cos\emptyset x(t)$$

Thus, output voltage of the coherent demodulator is proportional to the message signal x(t) if the phase error cos ∅ is constant .

The first term in the above Eq. 10.36 represents a DSBSC modulated signal with a carrier frequency $2f_c$ and the second term represents the scaled version of message signal. Assuming that the message signal is band limited to the internal $-\omega < f < \omega$, the spectrum of $v(t)$ is plotted as shown below.

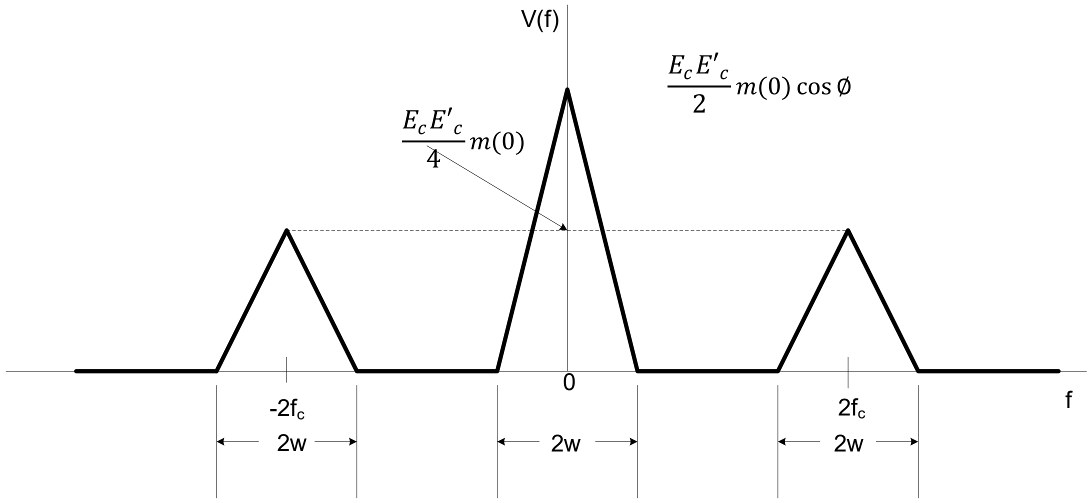


Figure 10.8: Spectrum of Output of the product Modulator

From the spectrum, it is clear that the unwanted component (first term in the expression) can be the low-pass filter, provided that the cut-off frequency of the filter is greater than ω but less than $2f_c$- ω. The filter output is given by

$$v_o(t) = \frac{E_c E'_c}{2} cos\ (\emptyset) m(t) \qquad 10.39$$

The demodulated signal $v_o(t)$ is therefore proportional to m(t) when the phase error f is constant.

10.8.2 Effect of Phase Error on the Demodulated Output

Let us consider the expression for the output of coherent detector is given by :

$$v_0(t) = \frac{1}{2} x(t). E_c \cos\emptyset x(t)$$

In this expression, Ø represents the phase error and the amplitude of the demodulated output is maximum and equal to (1/2) E_c when Ø = 0^o and the amplitude is zero when Ø = 90^o. This effect is called as the **quadrature null effect** of the coherent detector. Here, quadrature term represents the phase difference of 90^o. In other words, the phase error attenuates the demodulator output . In practice, the phase error varies randomly with time due to the random variations taking place in the communication channel . Hence, cosØ will vary randomly and the detector output also will vary in a random manner. This is undesirable . hence, circuitry must be provided in the detector to keep the locally generated carrier c'(t) in perfect synchronism, in both frequency and phase, with the original carrier c(t) .

10.8.3 Merit of coherent Detection Optical Transmission

A "coherent" optical transmission system is characterized by its capability to do "**coherent detection**," which means that an optical receiver can track the phase of an optical transmitter (and hence "phase coherence") so as to extract any phase and frequency information carried by a transmitted signal.

Coherent detection is well known in wireless communication systems. In those wireless systems, a radio frequency (RF) local oscillator (LO) is tuned to "heterodyne" ("heterodyne" is a signal processing technique which combine a high-frequency signal at f_1 with another at f_2 to produce a lower frequency signal at ($f_1 - f_2$), with a received signal through an RF mixer, as shown in Fig.10.9(a), so that both the amplitude and phase information contained in an RF carrier can be recovered in the following digital signal processor (DSP).

For an optical coherent system, a narrow-line width tun-able laser, serving as an LO, tunes its frequency to "intradyne" with a received signal frequency through an optical coherent mixer, as shown in Fig.10.9(b), and thereby recovers both the amplitude and phase information contained in a particular optical carrier. Here, "intradyne" means that the frequency difference between an LO and a received optical carrier is small and within the bandwidth of the receiver, but does not have to be zero. This implies that the frequency and phase of an LO do not have to be actively controlled to an extreme accuracy, therefore avoiding the use of a complicated optical phase locked loop.

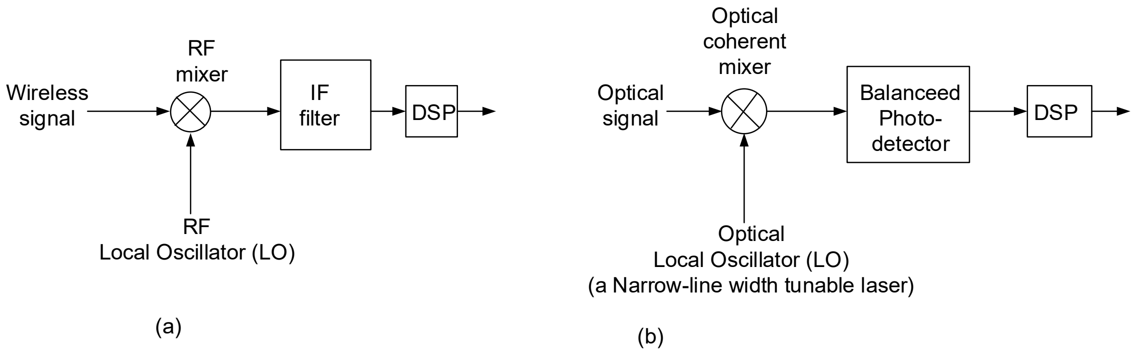


Figure 10.9 A simple illustration of how coherent detection works for (a) wireless systems, and (b) optical coherent systems

In contrast to coherent detection is direct detection, typically used by 10 Gb/s or lower-speed systems. In a direct detection receiver, its photo-detector only responds to changes in the receiving signal optical power, and cannot extract any phase or frequency information from the optical carrier.

Coherent detection therefore offers several key advantages compared to direct detection:

1) Greatly improved receiver sensitivity.
2) Can extract amplitude, frequency, and phase information from an optical carrier, and consequently can achieve much higher capacity in the same bandwidth.

3) Its DSP can compensate very large chromatic and polarization mode dispersion due to optical fibers, and eliminate the need for optical dispersion compensators and the associated optical amplifiers. This saves not only significant capex, but also simplifies optical network design tremendously, because the complicated dispersion map associated with 10 G and 40 G direct detection systems is no longer needed.
4) When using balanced detectors with a high common mode noise rejection ratio (CMRR), not only signal-to-noise ratio (SNR) can be improved further, but also agile wavelength selection can be achieved by LO tuning without the use of an optical filter or demultiplexer. This feature enables the next-generation colorless and directionless (CD), or even colorless, directionless, and contentionless (CDC), reconfigurable optical add-drop multiplexers (ROADM).

It should also be noted that laser phase noise of an LO is an important impairment in coherent systems as it impacts the "phase coherence". The "linewidth" parameter of a laser diode is directly related to its phase noise.

Although its fundamental concept is derived from wireless communications systems, coherent detection has started another paradigm shift in optical fiber communications. Its impact has become as huge as what commercial laser diodes and erbium-doped fiber amplifiers have brought to the industry. Consequently, long-haul optical networks have already become coherent-centric today, while metro optical networks are bound to become coherent-centric in the next few years.

10.9 Chapter Review Problems

10.1 Draw a diagram for the demodulation of single-sideband (SSB) amplitude modulated signals where the carrier is suppressed. Indicate the bandwidth of the band-pass filter.

10.2 Explain the demodulation of FM signal using Slope detector method.

10.3 Explain the demodulation of FM signal using Ratio detector method.

10.4 Explain the terms "synchronous detection", "envelope detection", "coherent detection", and "non-coherent detection"

10.5 Sketch the non-coherent receiver for the detection of binary FSK signals. Discuss the need for an envelope detector following the matched filter in the presence of phase error.

10.6 Explain Binary PSK and Binary FSK modulation and demodulation with the help of block diagrams.

10.7 (a) Name and explain briefly types of AM detectors.
(b) Differentiate between diode and square law detectors.
(c) Deduce the output of the square law detector from first principle.

10.8 (a) Explain briefly the Detection of DSBC signals using a regenerative receiver in the oscillating.

(b) Explain briefly the Demodulation of SSB signals in a regenerative receiver and derive the output waveform.

10.9 (a) What are the drawbacks of the slope detector and derive the output wave form from first principle

(b) What are the drawbacks of the balanced slope detector?

10.10 Describe briefly the following terms:

(a) Foster-Seeley Discriminator

(b) Ratio detector

10.11 With the use of a diagram explain the basic operation of the ration detector.

10.12 Explain the term "Quadrature Detector" using a simple diagram.

10.13 What is the phase lock loop (PLL)? Describe the configuration of the PLL FM detector.

10.14 Explain the demodulation of FM signal using the PLL.

10.15 Explain with the aid of block diagram Demodulation Of DSBSC Modulated Wave By Coherent Detection. Deduce its output waveform.

Name: ……………………………………………………………………………..
Reg. Number: …………………………………………………………….
Date: ……………………………..

Suppose the sinusoidal signal in **Fig. A** is sampled at time t = 0, 5, 10, ………..40 s. The maximum amplitude of the signal is 32 A.

(i) Draw the PAM signal using single line

(ii) Draw the PCM signal. Using three bits per sample. Prepare a table to show the possible voltage range, quantization levels and corresponding PCM Codes. Determine the signal-to-quantizing noise ratio in dB.

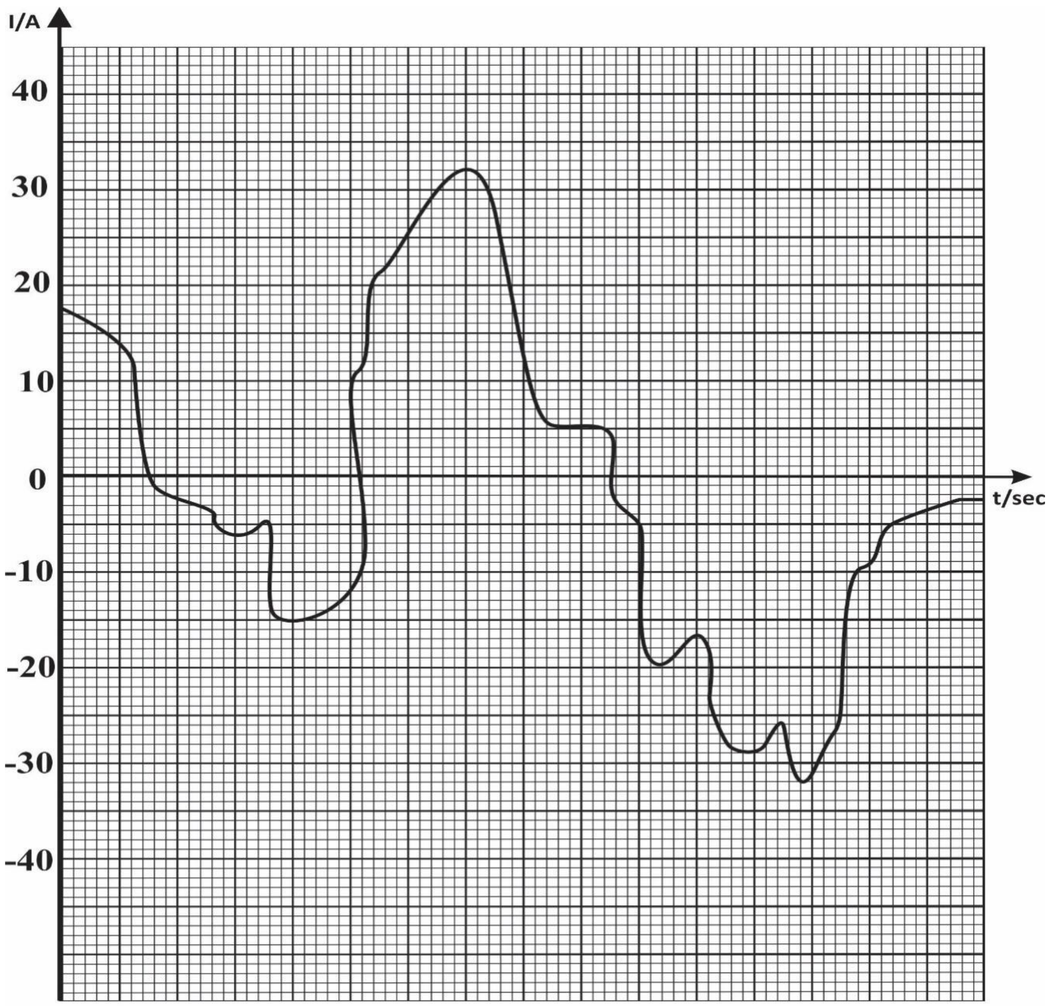


Figure A

BIBLIOGRAPHY

1. Charles K. Alexander and Mathew N.O Sadiku, (2002) Fundamentals of Electric Circuits (2^{nd} edition) McGraw-Hill, New York.
2. Frenzel Jr, Louis E (2002). Communication Electronics (3^{rd} edition). Columbus, McGraw-Hill Publishing Company, USA.
3. George Kennedy, et al. (1992), Electronic Communication Systems (5^{th} edition). Columbus, McGraw-Hill Kogaku Sha.

INDEX

E

Q

R

S

T