HSST-Physics

Electronics -Part 4

Modulation

Modulation is an important step of communication system. Modulation is defined as the process whereby some characteristics (amplitude, frequency and phase) of a high frequency signal wave (carrier wave) is varied in accordance with the instantaneous value of the voltage of low frequency signal wave (modulating wave). So the information which is a small frequency signal is added to a high frequency carrier signal by the process of modulation. Depending on which characteristic of the carrier is to be varied after the addition of information, we have different types of modulation schemes.

Two signals are involved in the modulation process. The baseband signal and the carrier signal. The baseband signal is the information signal or message which is to be transmitted to the receiver. The frequency of this signal is generally low. In the modulation process, this baseband signal is called the modulating signal. In radio broadcasting, radio program is the modulating signal and in TV, it is video signal. In telephone, our speech signal is the modulating signal.

The other signal involved in the modulation process is a high frequency sinusoidal wave. This signal is called the carrier signal or carrier. The frequency of the carrier signal is always much higher than that of the baseband signal. After modulation, the baseband signal of low frequency is transferred to the high frequency carrier, which carries the information in the form of some variations. After completing the modulation process, some characteristic of the carrier is changed such that the resultant variations carry the information

$e_c = Ec\ sin(\omega_c+\theta)$

the subscript 'c' indicates the carrier signal. The components of this equation are as follows:

- e_c : Instantaneous amplitude of the carrier
- E_e: Maximum amplitude of the carrier
- Angular frequency of the carrier, ω_c = 27πfc, where fc is the frequency of the carrier, also called the carrier frequency.
- θ : Initial phase of the carrier signal

Equation has three parameters namely, amplitudes (Ec), frequency (ω_c), and phase. In principle, these parameters have constant values for a particular sinusoidal wave. According to the definition of modulation, some characteristics of the carrier signal is varied in accordance with the modulating signal. After modulation any one of the three parameters of the carrier signal, namely, amplitude, frequency, or phase is varied keeping the remaining two constant. The baseband signal is then carried by these variations. The type of the modulation is decided by the parameter chosen to be varied.

If the amplitude of the carrier is chosen to be varied in accordance with the instantaneous amplitude of the baseband signal, keeping frequency and phase constant, the resulting modulation is called amplitude modulation (AM). Frequency modulation (FM) and phase modulation (PM) are also obtained in a similar way.

Low-frequency baseband signal is thus translated to a high frequency carrier such that the information is coded in the variations in one of the parameters of the carrier. At the receiver's side, these variations are detected through the demodulation process to recover the original baseband signal.

The following can be summarized with reference to modulation.

- The baseband signal is known as the modulating signal.
- The baseband signal is a low-frequency signal.
- The carrier signal is always a high frequency sinusoidal wave.
- During the modulation process, the modulating signal changes the frequency, amplitude, or phase of the carrier in accordance with its instantaneous amplitude.
- After modulation, the carrier is said to be modulated by the modulating signal.
- The output of the modulator is called the modulated signal.

The process of modulation in a communication system increases its cost and complexity. However, modulation is extensively used in most communication systems. There is a definite need for using modulation.

Need for modulation

(i) To separate signal from different transmitters

Audio frequencies are within the range of20 Hz to 20 kHz. Without modulation all signals of same frequencies from different transmitters would be mixed up. It will create an impossible situation to tune any one of them or separate them. In order to separate the various signals, the radio stations must broadcast at different frequencies. Each radio station must be given its own carrier frequency and frequency band. This can be achieved by a process called frequency translation. This frequency translation is the result of modulation process.

(ii) Size of the antenna: -

For efficient transmission, the transmitting antennas should have a length at least equal to a quarter of the wavelength of the signal to be transmitted. For an electromagnetic wave of frequency 15 kHz, the wavelength is 20 km and one-quarter of this will be equal to 5 km. Obviously a vertical antenna of this size is impractical. On the other hand, for a frequency of 1 MHz, the height is reduced to 75m. Thus, instead of transmitting the message signal directly using an antenna, we can give the modulated signal to the antenna which is a high frequency signal containing the message. In this way the size of the antenna can be reduced by modulation.

(iii) Power of radiation: -

Also, the power radiated by an antenna of length l is proportional to (l/λ^2) This shows that for the same antenna length, power radiated is large for shorter wavelength. Thus, the signal which is of low frequency, must be translated to the high frequency spectrum of the electromagnetic wave so that, more power can be effectively radiated. This is achieved through the process of modulation.

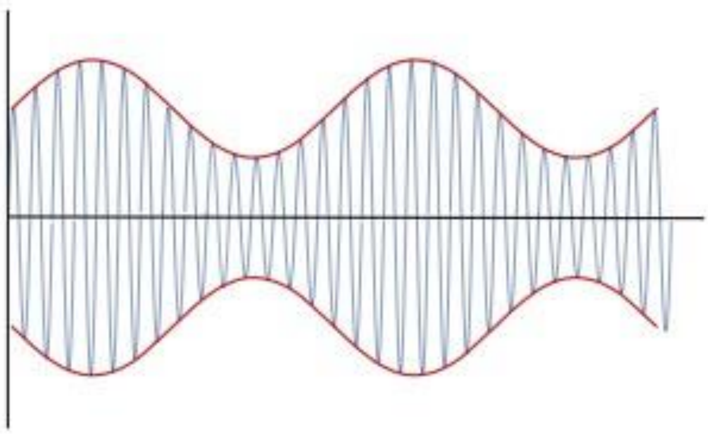
Amplitude Modulation, AM

Amplitude modulation or AM as it is often called, is a form of modulation used for radio transmissions for broadcasting and two-way radio communication applications.

The first amplitude modulated signal was transmitted in 1901 by a Canadian engineer named Reginald Fessenden. He took a continuous spark transmission and placed a carbon microphone in the antenna lead.

Although one of the earliest used forms of modulation it is still used today, mainly for long, medium and short-wave broadcasting and for some aeronautical point to point communications.

One of the key reasons for the use of amplitude modulation was its ease of use. The system simply required the carrier amplitude to be modulated, but more usefully the detector required in the receiver could be a simple diode-based circuit. This meant that AM radios did not need complicated demodulators and costs were reduced - a key requirement for widespread use of radio technology, especially in the early days of radio when ICs were not available.



With the introduction of continuous sine wave signals, transmissions improved significantly, and AM soon became the standard for voice transmissions. Nowadays, amplitude modulation, AM is used for audio broadcasting on the long medium and short wave bands, and for two way radio communication at VHF for aircraft.

However as there now are more efficient and convenient methods of modulating a signal, its use is declining, although it will still be very many years before it is no longer used.

Amplitude modulation applications

Amplitude modulation is used in a variety of applications. Even though it is not as widely used as it was in previous years in its basic format it can nevertheless still be found.

- Broadcast transmissions: AM is still widely used for broadcasting on the long, medium and short wave bands. It is simple to demodulate and this means that radio receivers capable of demodulating amplitude modulation are cheap and simple to manufacture. Nevertheless, many people are moving to high quality forms of transmission like frequency modulation, FM or digital transmissions.
- Air band radio: VHF transmissions for many airborne applications still use AM. It is used for ground to air radio communications as well as two way radio links for ground staff as well.
- Single sideband: Amplitude modulation in the form of single sideband is still used for HF radio links. Using a lower bandwidth and providing more effective use of the transmitted power this form of modulation is still used for many point to point HF links.
- Quadrature amplitude modulation: AM is widely used for the transmission of data in everything from short range wireless links such as Wi-Fi to cellular telecommunications and much more. Effectively it is formed by having two carriers 90° out of phase.

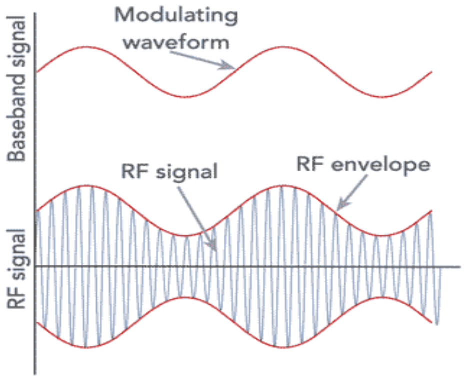
These form some of the main uses of amplitude modulation. However in its basic form, this form of modulation is being used less as a result of its inefficient use of both spectrum and power.

Amplitude modulation

In order that a radio signal can carry audio or other information for broadcasting or for two way radio communication, it must be modulated or changed in some way. Although there are a number of ways in which a radio signal may be modulated, one of the easiest is to change its amplitude in line with variations of the sound.

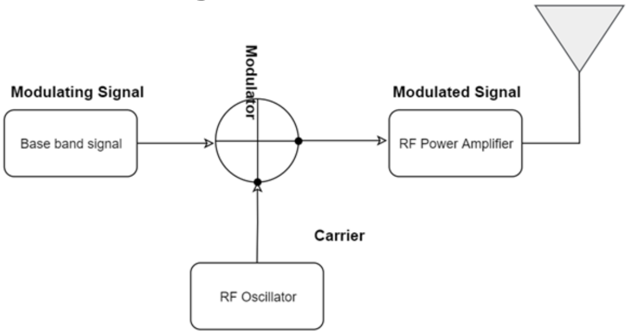
In this way the amplitude of the radio frequency signal varies in line with the instantaneous value of the intensity of the modulation. This means that the radio frequency signal has a representation of the sound wave superimposed in it.

In view of the way the basic signal "carries" the sound or modulation, the radio frequency signal is often termed the "carrier".

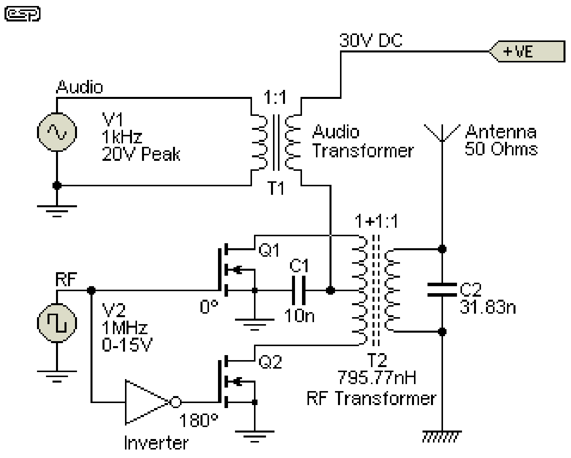


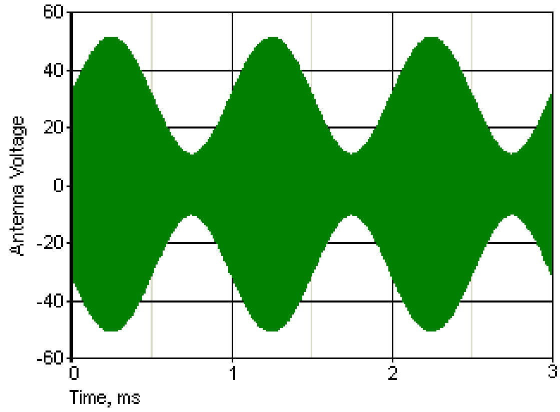
From the diagram, it can be seen that the envelope of the signal follows the contours of the modulating signal.

Amplitude Modulation Block Diagram



Here the modulating signals might be an audio or video signal. These are also called as baseband signals as these are modulated with the carrier signals. Carriers are extremely high-frequency radio signals, In general, carrier signals are received from the RF oscillators. These two signals are combined in a modulator. The modulator considers the instant amplitude of the modulating signal and modifies it as per the amplitude of the carrier signal. So, the resultant signal amplitude is the amplitude of the modulated signal. The modulated signal is passed through the amplifier for the amplitude modulation and then transmitted through an antenna or a co-axial cable.





Amplitude demodulation

Amplitude modulation, AM, is one of the most straightforward ways of modulating a radio signal or carrier. It can be achieved in a number of ways, but the simplest uses a single diode rectifier circuit.

Other methods of demodulating an AM signal use synchronous techniques and provide much lower levels of distortion and improved reception where selective fading is present.

One of the main reasons for the popularity of amplitude modulation has been the simplicity of the demodulation. It enables costs to be kept low - a significant advantage in producing vast quantities of very low cost AM radios.

Advantages & disadvantages of amplitude modulation, AM

As with any technology there are advantages and disadvantages to be considered. The summary below gives a highlight of the basic pros and cons.

Advantages

- It is simple to implement
- it can be demodulated using a circuit consisting of very few components
- AM receivers are very cheap as no specialised components are needed.

Disadvantages

- It is not efficient in terms of its power usage
- It is not efficient in terms of its use of bandwidth, requiring a bandwidth equal to twice that of the highest audio frequency
- It is prone to high levels of noise because most noise is amplitude based and obviously AM detectors are sensitive to it.

Although in the current technological climate, AM in its basic form is not nearly as effective as other modes that can be used, it is still retained in many areas like broadcasting, because of the number of users. However, it is likely that with time, its use will decrease still further and ultimately many AM transmissions will cease. However, its derivatives like quadrature amplitude modulation are widely used as they offer a very effective form of modulation, especially for data transmission.

Although the use of amplitude modulation is decreasing, it nevertheless forms the basis of other forms of modulation that are still being widely used, or their use is increasing.

Single sideband, SSB: Single sideband is widely used for HF communications. It is formed by taking a signal that has the carrier and one sideband removed. In this way it becomes far more efficient in terms of both spectrum and power.

Quadrature amplitude modulation, QAM: This form of modulation is essentially derived from two carriers that are 90° out of phase and adding information, either analogue or digital. Quadrature Amplitude Modulation is widely used for carrying many digital signals, everything from Wi-Fi to Mobile phone communications and very much more.

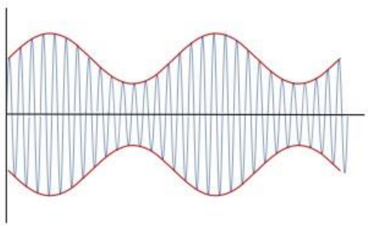
Amplitude Modulation Theory & Equations

The basic theory and equations behind amplitude modulation are relatively straightforward and can be handled using straightforward trigonometric calculations and manipulation.

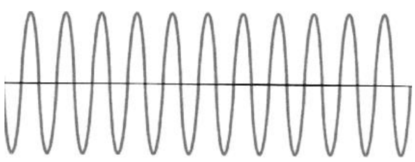
Essentially an amplitude modulated wave consists of a radio frequency carrier - a sine wave at one frequency, typically in the radio frequency portion of the spectrum. A modulating wave, which in theory could be another sine wave, typically at a lower audio frequency is superimposed upon the carrier.

The two signals are multiplied together and the theory shows how they interact to create the carrier and two sidebands.

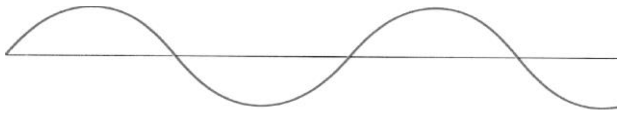
The equations for the simple example of a single tone used for modulation can be expanded to show how the signal will appear of a typical sound consisting of many frequencies is used to modulated the carrier.



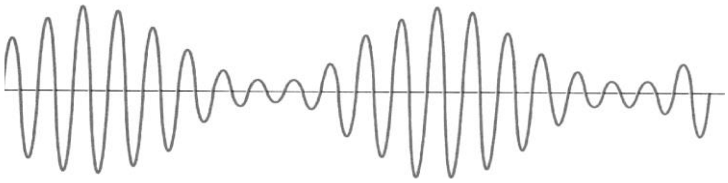
Amplitude modulated signal



Carrier Signal



Modulating Sine Wave Signal



Amplitude Modulated Signal

Amplitude modulation theory & equations

It is possible to look at the theory of the generation of an amplitude modulated signal in four steps:

1. Carrier signal
2. Modulating signal
3. Overall modulated signal for a single tone
4. Expansion to cover a typical audio signal

1. <u>Carrier signal equations</u>

Looking at the theory, it is possible to describe the carrier in terms of a sine wave as follows:

$C(t) = C \sin(\omega c + \varphi)$

Where:

carrier frequency in Hertz is equal to ωc / 2 π

C is the carrier amplitude

φ is the phase of the signal at the start of the reference time

Both C and φ can be omitted to simplify the equation by changing C to "1" and φ to "0".

2. **<u>Modulating signal equations</u>**

The modulating waveform can either be a single tone. This can be represented by a cosine waveform, or the modulating waveform could be a wide variety of frequencies - these can be represented by a series of cosine waveforms added together in a linear fashion.

For the initial look at how the signal is formed, it is easiest to look at the equation for a simple single tone waveform and then expand the concept to cover the more normal case. Take a single tone waveform:

$m(t) = M \sin(\omega m + \varphi)$

Where:

modulating signal frequency is equal to $\omega m / 2\pi$ Hertz

M is the carrier amplitude

φ is the phase of the signal at the start of the reference time

Both C and φ can be omitted to simplify the equation by changing C to "1" and φ to "0".

3. **Overall modulated signal for a single tone**

The equation for the overall modulated signal is obtained by multiplying the carrier and the modulating signal together.

$y(t) = [A + m(t)].c(t)$

The constant A is required as it represents the amplitude of the waveform.

Substituting in the individual relationships for the carrier and modulating signal, the overall signal becomes:

$y(t) = [A + M\cos(\omega mt+\varphi].\sin(\omega ct)$

The trigonometry can then be expanded out to give an equation that includes the components of the signal:

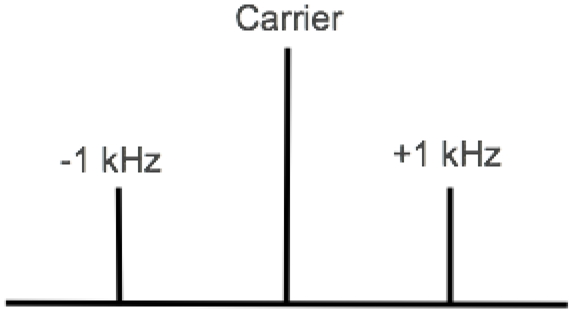
$y(t) = A.\sin(\omega ct)+\frac{AM}{2}[\sin((\omega c+\omega m)t+\varphi)]+\frac{AM}{2}[\sin((\omega c-\omega m)t-\varphi)]$

In this theory, three terms can be seen which represent the carrier, and upper and lower sidebands:

Carrier: $A . \sin(\omega c\, t)$

Upper sideband: $\frac{AM}{2}$ $[\sin((\omega c + \omega m)\, t + \varphi)$

Lower sideband: $\frac{AM}{2}$ $[\sin((\omega c - \omega m)\, t - \varphi)$



Spectrum (sidebands) arising from carrier
modulated by 1 kHz tone

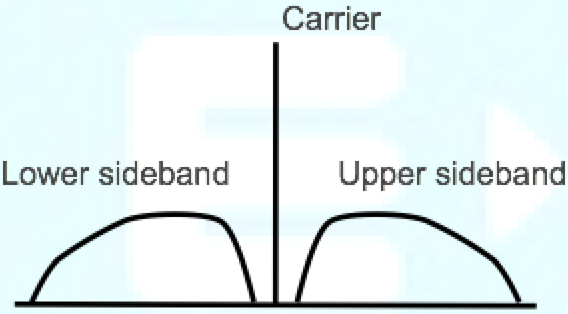
It can be seen that for a case where there is 100% modulation, i.e. M = 1, and where the carrier is not suppressed, i.e. A = 1, then the sidebands have half the value of the carrier, i.e. a quarter of the power each.

Expansion to cover a typical audio signal

With the basic concept of modulation and the resultant sidebands established, the same principles can be applied to the more complicated cases of modulation using speech, music or other audio sounds.

Theory can be used to break down a sound into a series of sinusoidal signals. These are linearly added to each other to form the audio spectrum of the modulating signal.

The spectrum of the modulating signal extends out either side from the carrier, one sideband is the mirror of the other, with the lowest frequencies closest to the carrier, and highest furthest away.



Sidebands on an amplitude modulated carrier when modulated with an audio signal containing many frequencies

It can be seen that the audio signal covers a band of frequencies either side of the main carrier. The theory and equations show that furthest extent of the sidebands from the carrier corresponds to the highest frequency of the modulating tone for the amplitude modulated signal.

Seeing a little of the theory and mathematics behind amplitude modulation gives a better understanding of how it works. This can then be applied to use this type of mode to its best, whether as amplitude modulation, single sideband, or even giving a better understanding of how QAM operates. Understanding how the modulating waveform not only generates undulations of the envelope, but also generates sidebands, etc enables the basic concepts behind AM to be understood.

Amplitude Modulation, AM Index & Modulation Depth

it is possible to vary the level of modulation applied to an amplitude modulated signal. This is an important factor for broadcast and two-way radio communications applications.

If little modulation is applied then the audio (assuming it us an audio transmission) will be difficult to hear. However, if too, much is applied, distortion can result and signals will not be easy to listen to and interference will increase and this could affect users on nearby frequencies or channels.

Modulation index

The amplitude modulation index describes the amount by which the modulated carrier envelope varies about the static level.

Modulation Index, m=M/A

Where:

A = the carrier amplitude.

M = the modulation amplitude and is the peak change in the RF amplitude from its un-modulated value.

Using the equation above it can be seen that a modulation index of 0.75 means that the signal will increase by a factor of 0.75 and decrease to 0.25 of its original level.

Types of AM

AM signal transmission is of different types. These types are developed on the basis of the power to be transmitted and the bandwidth availability. If one sideband or apart of the sideband is removed, we can save bandwidth and if the carrier is removed or reduced partly, we can save power. The various types are discussed in brief below.

Signal-to-Noise Ratio (SNR)

In analog and digital communications, a signal-to-noise ratio, often written S/N or SNR, is a measure of the strength of the desired signal relative to background noise (undesired signal). S/N can be determined by using a fixed formula that compares the two levels and returns the ratio, which shows whether the noise level is impacting the desired signal.

The ratio is typically expressed as a single numeric value in decibels (dB). The ratio can be zero, a positive number or a negative number. **A signal-to-noise ratio over 0 dB indicates that the signal level is greater than the noise level**. The higher the ratio, the better the signal quality.

Noise includes any unwanted disturbance that degrades the quality of the desired signal. It can include thermal, quantum, electronic, impulse or intermodulation noise, as well as other forms of noise. Environmental factors, such as temperature and humidity, can also affect noise levels.

If the noise is significant enough in comparison to the desired signal, that is, S/N is low. It can disrupt a wide range of data transfers, including text files, graphics, telemetry, applications, and audio and video streams.

Communications engineers always strive to maximize S/N. Traditionally, this has been done by using the narrowest possible receiving-system bandwidth consistent with the data speed desired. However, there are other methods. For example, engineers might use spread spectrum techniques to improve system performance, or they might boost the signal output power to increase S/N.

In some high-level systems, such as radio telescopes, internal noise is minimized by lowering the temperature of the receiving circuitry to near-absolute zero (-273 degrees

Celsius or -459 degrees Fahrenheit). In wireless systems, it is always important to optimize the performance of the transmitting and receiving antennas.

Calculation

The signal-to-noise ratio is typically measured in decibels and can be calculated by using a base 10 logarithm. The exact formula depends on how the signal and noise levels are measured, though.

For example, if they're measured in microvolts, the following formula can be used:

$S/N = 20 \log_{10}(Ps/Pn)$

Ps is the signal in microvolts, and Pn is the noise in microvolts.

However, if the signal and noise are measured in watts, the formula is slightly different:

$S/N = 10 \log_{10}(Ps/Pn)$

The letter P is often used in these formulas to indicate power.

When Ps equals Pn, S/N will be 0. A ratio of 0 dB indicates that the signal is competing directly with the noise level, resulting in a signal that borders on unreadable. In digital communications, this can cause a reduction in data speed because of frequent errors that require the transmitting system to resend data packets.

When Ps is greater than Pn, S/N will be positive. Ideally, Ps should be much greater than Pn to minimize noise interference. As an example, suppose that Ps equals 10 microvolts and Pn equals 1 microvolt. Because 10 divided by 1 equal 10, the following formula can be used to calculate S/N:

$S/N = 20 \log_{10}(10) = 20$ dB

A ratio of 20 dB means that the signal is clearly readable. If the signal is much weaker but still above the noise level, say, 1.3 microvolts then S/N is much lower, in this case, only 2.28 dB:

$S/N = 20 \log_{10}(1.3) = 2.28$ dB

This is a marginal situation that could impact network performance, although it's not the worst possible situation. When Ps is less than Pn, S/N is negative, a low signal-to-noise ratio. In this type of situation, reliable communication is nearly impossible, and steps should be taken to increase the signal level, decrease the noise level or implement a combination of both.

Pulse Code Modulation

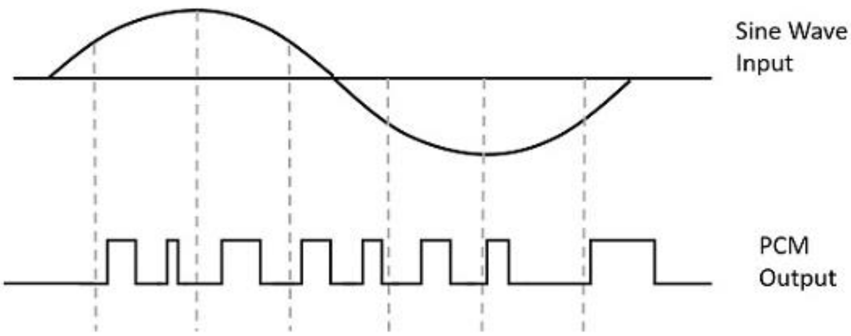
Modulation is the process of varying one or more parameters of a carrier signal in accordance with the instantaneous values of the message signal.

The message signal is the signal which is being transmitted for communication and the carrier signal is a high frequency signal which has no data, but is used for long distance transmission.

There are many modulation techniques, which are classified according to the type of modulation employed. Of them all, the digital modulation technique used is **Pulse Code Modulation (**PCM).

Pulse-code modulation (PCM) is a method used to digitally represent sampled analog signals. It is the standard form of digital audio in computers, compact discs, digital telephony and other digital audio applications. In a PCM stream, the amplitude of the analog signal is sampled regularly at uniform intervals, and each sample is quantized to the nearest value within a range of digital steps.

A signal is pulse code modulated to convert its analog information into a binary sequence, i.e., **1s** and **0s**. The output of a PCM will resemble a binary sequence. The following figure shows an example of PCM output with respect to instantaneous values of a given sine wave.



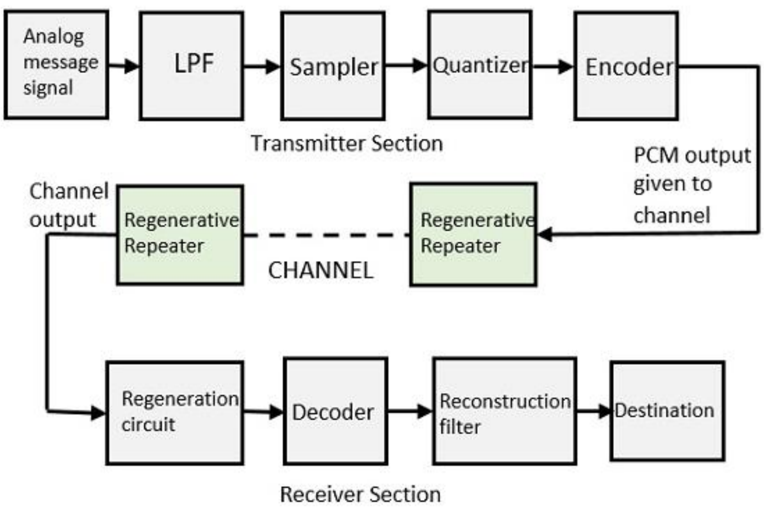
Instead of a pulse train, PCM produces a series of numbers or digits, and hence this process is called as **digital**. Each one of these digits, though in binary code, represent the approximate amplitude of the signal sample at that instant.

In Pulse Code Modulation, the message signal is represented by a sequence of coded pulses. This message signal is achieved by representing the signal in discrete form in both time and amplitude.

Basic Elements of PCM

The transmitter section of a Pulse Code Modulator circuit consists of **Sampling, Quantizing** and **Encoding**, which are performed in the analog-to-digital converter section. The low pass filter prior to sampling prevents aliasing of the message signal.

The basic operations in the receiver section are **regeneration of impaired signals, decoding,** and **reconstruction** of the quantized pulse train. Following is the block diagram of PCM which represents the basic elements of both the transmitter and the receiver sections.



Low Pass Filter

This filter eliminates the high frequency components present in the input analog signal which is greater than the highest frequency of the message signal, to avoid aliasing of the message signal.

Sampler

This is the technique which helps to collect the sample data at instantaneous values of message signal, so as to reconstruct the original signal. The sampling rate must be greater than twice the highest frequency component **W** of the message signal, in accordance with the sampling theorem.

Quantizer

Quantizing is a process of reducing the excessive bits and confining the data. The sampled output when given to Quantizer, reduces the redundant bits and compresses the value.

Encoder

The digitization of analog signal is done by the encoder. It designates each quantized level by a binary code. The sampling done here is the sample-and-hold process. These three sections LPF, Sampler and Quantizer will act as an analog to digital converter. Encoding minimizes the bandwidth used.

Regenerative Repeater

This section increases the signal strength. The output of the channel also has one regenerative repeater circuit, to compensate the signal loss and reconstruct the signal, and also to increase its strength.

Decoder

The decoder circuit decodes the pulse coded waveform to reproduce the original signal. This circuit acts as the demodulator.

Reconstruction Filter

After the digital-to-analog conversion is done by the regenerative circuit and the decoder, a low-pass filter is employed, called as the reconstruction filter to get back the original signal.

Hence, the Pulse Code Modulator circuit digitizes the given analog signal, codes it and samples it, and then transmits it in an analog form. This whole process is repeated in a reverse pattern to obtain the original signal.

Advantages and Disadvantages

Advantages

- Pulse Code Modulation is used in long-distance communication.
- The efficiency of the transmitter in PCM is high.
- Higher noise immunity is seen.
- Efficient method.

Disadvantages

- The bandwidth requirement is high.
- PCM is a complex process, since it involves encoding, decoding and quantisation of the circuit.

Applications of Pulse Code Modulation

- It is used in telephony and compact discs.
- Pulse Code Modulation is used in satellite transmission systems and space communications.