Angle Modulation (FM & PM)

Angle Modulation is the process in which the frequency or the phase of the carrier varies according to the message signal. This is further divided into frequency and phase modulation.

- Frequency Modulation is the process of varying the frequency of the carrier signal linearly with the message signal.
- Phase Modulation is the process of varying the phase of the carrier signal linearly with the message signal.

  Angle Modulated wave $S(t) = A_c \cos \odot_i(t)$

  Where A_c = Amplitude of the modulated wave or Amplitude of carrier wave.

  $\odot_i(t)$ = Angle of the modulated wave.

Let us now discuss these topics in greater detail.

Frequency Modulation (FM)

Frequency modulation (FM) is a technique used to encode data on an alternating digital or analog signal. The method includes varying the frequency of the carrier wave on which useful information is imposed or impressed upon. The signal on which data is imposed is known as the carrier signal and the resulting signal with variable frequency is called a frequency modulated signal.

Frequency modulation is widely used for radio transmission due to the fact that the signal-to-noise ratio (SNR) is large in this method of modulation and hence radio frequency interference is minimized. FM signals are used in technology such as radars, telemeters, Electroencephalogram (EEG- a test used to find the problem related to the brain), radio broadcasting, satellite communication and magnetic tape recording systems. The frequencies vary by up to 5 kHz in the case of wireless two-way communication and they vary up to several MHz in the case of wireless broadcasting.

In amplitude modulation, the amplitude of the carrier varies. But in Frequency Modulation (FM), the frequency of the carrier signal varies in accordance with the instantaneous amplitude of the modulating signal.

The amplitude and the phase of the carrier signal remains constant whereas the frequency of the carrier changes. This can be better understood by observing the following figures.

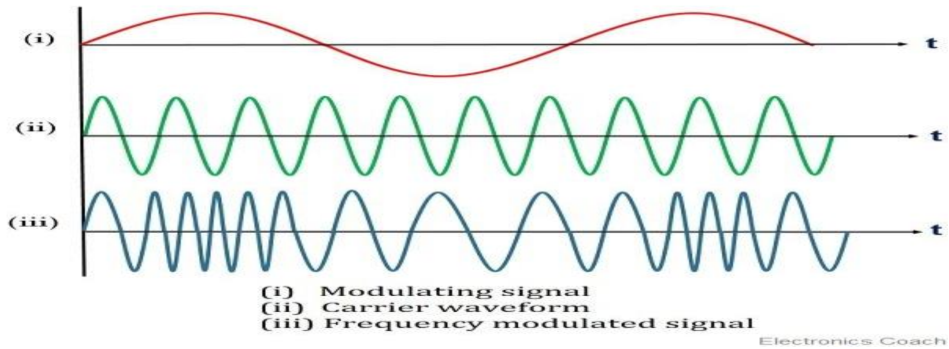


Fig: Frequency Modulation

The frequency of the modulated wave remains constant as the carrier wave frequency when the message signal is at zero. The frequency increases when the message signal reaches its maximum amplitude. That is, with the increase in amplitude of the modulating or message signal, the carrier frequency increases. Likewise, with the decrease in the amplitude of the modulating signal, the frequency also decreases.

Mathematical Representation

Let the carrier frequency be $\mathbf{f_c}$

The frequency at maximum amplitude of the message signal = $\mathbf{f_c + \Delta f}$

The frequency at minimum amplitude of the message signal = $\mathbf{f_c - \Delta f}$

The difference between FM modulated frequency and normal frequency is termed as **Frequency Deviation** and is denoted by $\mathbf{\Delta f}$.

The deviation of the frequency of the carrier signal from high to low or low to high can be termed as the **Carrier Swing**.

Carrier Swing = 2 × frequency deviation = $\mathbf{2 \times \Delta f}$

Equation for FM WAVE

The equation for FM wave is −

s(t) = A_ccos[W_ct + 2\pi k_fm(t)]

Where,

$\mathbf{A_c}$ = the amplitude of the carrier

$\mathbf{w_c}$ = angular frequency of the carrier = $\mathbf{2\pi f_c}$

m(t) = message signal

FM can be divided into **Narrowband FM** and **Wideband FM**.

Narrowband FM
The features of Narrowband FM are as follows −

- This frequency modulation has a small bandwidth.
- The modulation index is small.

- Its spectrum consists of carrier, USB, and LSB.i
- This is used in mobile communications such as police wireless, ambulances, taxicabs, etc.

Wideband FM

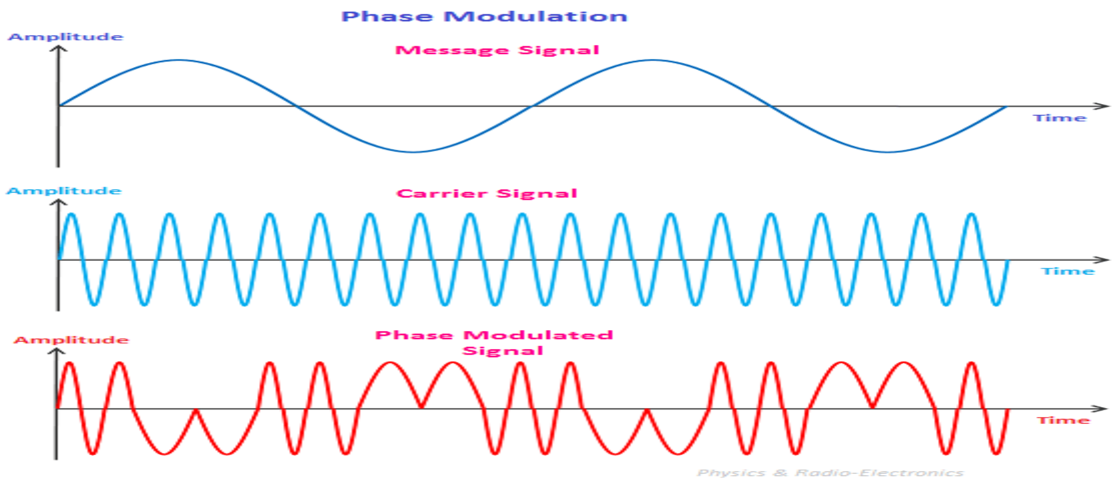
The features of Wideband FM are as follows −

- This frequency modulation has infinite bandwidth.
- The modulation index is large, i.e., higher than **1**.
- Its spectrum consists of a carrier and infinite number of sidebands, which are located around it.
- This is used in entertainment broadcasting applications such as FM radio, TV, etc.

Phase Modulation

In frequency modulation, the frequency of the carrier varies. But in **Phase Modulation (PM)**, the phase of the carrier signal varies in accordance with the instantaneous amplitude of the modulating signal.

The amplitude and the frequency of the carrier signal remains constant whereas the phase of the carrier changes. This can be better understood by observing the following figures.



The phase of the modulated wave has got infinite points where the phase shift in a wave can take place. The instantaneous amplitude of the modulating signal, changes the phase of the carrier. When the amplitude is positive, the phase changes in one direction and if the amplitude is negative, the phase changes in the opposite direction.

Relation between PM and FM

The change in phase, changes the frequency of the modulated wave. The frequency of the wave also changes the phase of the wave. Though they are related, their relationship is not linear. Phase modulation is an indirect method of producing FM. The amount of frequency shift, produced by a phase modulator increases with the modulating frequency. An audio equalizer is employed to compensate this.

Equation for PM Wave

The equation for PM wave is −

s(t) = A_ccos[W_ct + k_pm(t)]

Where,

$\mathbf{A_c}$ = the amplitude of the carrier

$\mathbf{w_c}$ = angular frequency of the carrier = $\mathbf{2\pi f_c}$

m(t) = message signal

Phase modulation is used in mobile communication systems, while frequency modulation is used mainly for FM broadcasting.

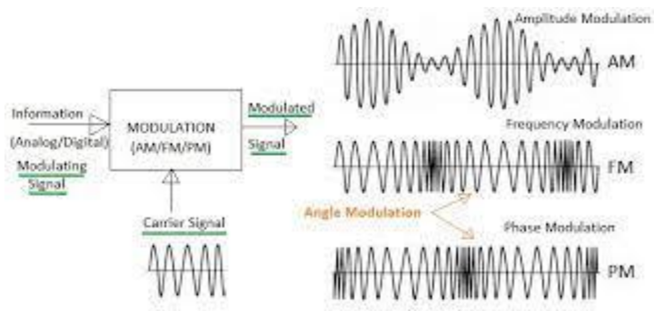


Fig. Comparison of AM, FM and PM diagrams

NOISE

In any communication system, during the transmission of the signal, or while receiving the signal, some unwanted signal gets introduced into the communication, making it unpleasant for the receiver, questioning the quality of the communication. Such a disturbance is called as **Noise**. Noise is an **unwanted signal** which interferes with the original message signal and corrupts the parameters of the message signal. This alteration in the communication process, leads to the message getting altered. It is most likely to be entered at the channel or the receiver.

The noise signal can be understood by taking a look at the following example.

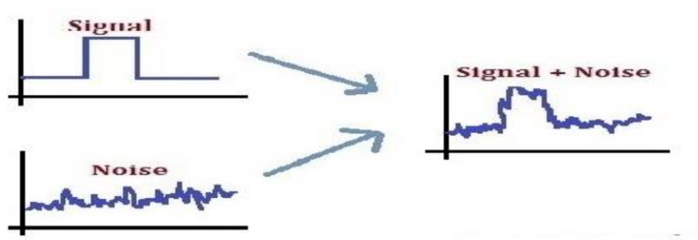
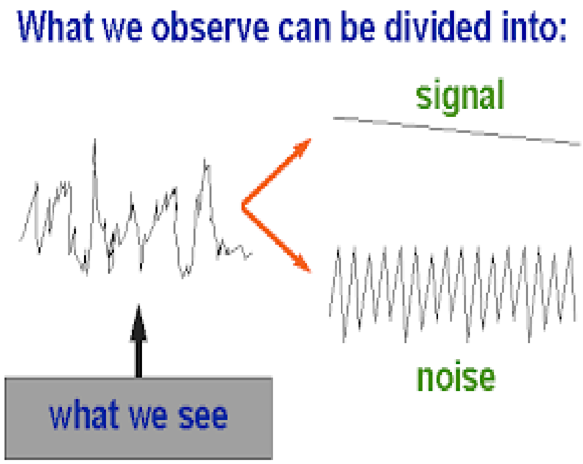


Fig: Noise signal



Hence, it is understood that noise is some signal which has no pattern and no constant frequency or amplitude. It is quite random and unpredictable. Measures are usually taken to reduce it, though it can't be completely eliminated.

Effects of Noise

1. Noise is an inconvenient feature which affects the system performance. Following are the effects of noise.
2. Noise limits the operating range of the systems
3. Noise indirectly places a limit on the weakest signal that can be amplified by an amplifier. The oscillator in the mixer circuit may limit its frequency because of noise. A system's operation depends on the operation of its circuits. Noise limits the smallest signal that a receiver is capable of processing.
4. Noise affects the sensitivity of receiver system.

-Sensitivity is the minimum amount of input signal necessary to obtain the specified quality output. Noise affects the sensitivity of a receiver system, which eventually affects the output.

Classifications of Noise

The classification of noise is done depending on the type of the source, the effect it shows or the relation it has with the receiver, etc. There are two main ways in which noise is produced. One is through some **external source** while the other is created by an **internal source**, within the receiver section.

External Source

This noise is produced by the external sources which may occur in the medium or channel of communication, usually. This noise cannot be completely eliminated. The best way is to avoid the noise from affecting the signal. Examples: Most common examples of this type of noise are −

- Atmospheric noise (due to irregularities in the atmosphere).
- Extra-terrestrial noise, such as solar noise and cosmic noise.
- Industrial noise.

Internal Source

This noise is produced by the receiver components while functioning. The components in the circuits, due to continuous functioning, may produce few types of noise. This noise is quantifiable. A proper receiver design may lower the effect of this internal noise.

Examples: Most common examples of this type of noise are −

- Thermal agitation noise (Johnson noise or Electrical noise).
- Shot noise (due to the random movement of electrons and holes).
- Transit-time noise (during transition).
- Miscellaneous noise is another type of noise which includes flicker, resistance effect and mixer generated noise, etc.

Different Types of Noises

Noise can be categorized in any sound, size, and shape. It is part of our day-to-day lives, but before one can accurately measure it, they need to know its various types. The types are as listed below:

- **Continuous noise: –** this is a noise produced in a continuous flow, for example, by a machine that runs without interruption.
- **Intermittent noise: –** this category refers to the noise whose levels keep increasing and decreasing. This noise is produced by such things like aircraft overhead, factory equipment operating in cycles, or trains passing by some place.
- **Impulsive noise: –** an impulsive noise is commonly associated with the construction and demolition fields. It is usually as a result of sudden bursts and can startle anyone nearby due to its fast and surprising nature.
- **Low-frequency noise: –** a low-frequency noise is associated with our daily soundscape. Such noise could be from power plants humming in the background or large diesel engines roaring in low frequencies.

Signal to Noise Ratio (SNR)

It is the **ratio of the signal power to the noise power**. The higher the value of SNR, the greater will be the quality of the received output. Signal-to-noise ratio is defined as the ratio of the power of a signal (input) to the power of background (unwanted input):

Signal-to-noise ratio at different points can be calculated by using the following formulae −

It is so because for a receiver, the channel is the input.

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}},$$ where P is average power.

Both signal and noise power must be measured at the same or equivalent points in a system, and within the same system bandwidth. Depending on whether the signal is a constant (s) or a random variable (S), the signal-to-noise ratio for random noise N becomes:

$$\mathrm{SNR} = \frac{s^2}{\mathrm{E}[N^2]}$$ where E refers to the expected value, i.e. in this case the mean square of N, or

$$\mathrm{SNR} = \frac{\mathrm{E}[S^2]}{\mathrm{E}[N^2]}$$

If the noise has expected value of zero, as is common, the denominator is its variance, the square of its standard deviation σ_N.

The signal and the noise must be measured the same way, for example as voltages across the same impedance. The root mean squares can alternatively be used in the ratio:

$$\mathrm{SNR} = \frac{P_\mathrm{signal}}{P_\mathrm{noise}} = \left(\frac{A_\mathrm{signal}}{A_\mathrm{noise}}\right)^2,$$

where A is root mean square (RMS) amplitude (for example, RMS voltage).

Logarithm Decibels Scale

Because many signals have a very wide dynamic range, signals are often expressed using the logarithmic decibel scale. Based upon the definition of decibel, signal and noise may be expressed in decibels (dB) as

$$P_\mathrm{signal,dB} = 10\log_{10}(P_\mathrm{signal})$$

And

$$P_\mathrm{noise,dB} = 10\log_{10}(P_\mathrm{noise}).$$

In a similar manner, SNR may be expressed in decibels as

$$\mathrm{SNR_{dB}} = 10\log_{10}(\mathrm{SNR}).$$

Using the definition of SNR

$$\mathrm{SNR_{dB}} = 10\log_{10}\left(\frac{P_\mathrm{signal}}{P_\mathrm{noise}}\right).$$

Using the quotient rule for logarithms

$$10\log_{10}\left(\frac{P_\mathrm{signal}}{P_\mathrm{noise}}\right) = 10\log_{10}(P_\mathrm{signal}) - 10\log_{10}(P_\mathrm{noise}).$$

Substituting the definitions of SNR, signal, and noise in decibels into the above equation results in an important formula for calculating the signal to noise ratio in decibels, when the signal and noise are also in decibels:

$$\mathrm{SNR_{dB}} = P_\mathrm{signal,dB} - P_\mathrm{noise,dB}.$$

In the above formula, P is measured in units of power, such as watts (W) or milliwatts (mW), and the signal-to-noise ratio is a pure number.

However, when the signal and noise are measured in volts (V) or amperes (A), which are measures of amplitude,they must first be squared to obtain a quantity proportional to power, as shown below:

$$\mathrm{SNR_{dB}} = 10\log_{10}\left[\left(\frac{A_\mathrm{signal}}{A_\mathrm{noise}}\right)^2\right] = 20\log_{10}\left(\frac{A_\mathrm{signal}}{A_\mathrm{noise}}\right) = \left(A_\mathrm{signal,dB} - A_\mathrm{noise,dB}\right).$$

NOISE SUPPRESSION

Active noise control (ANC) also known as noise cancellation (NC) or active noise reduction (ANR) is a method of reducing unwanted sound by the addition of a second sound specifically designed to cancel the first.. The concept was first developed in the late 30's; later developmental work that began in 1950's eventually resulted in commercial airline headsets with the technology becoming available in the late 1980's. The technology is also used in road vehicles and mobile telephones.

Sound is a pressure wave, which consists of alternating periods of compression and rarefaction. A noise-cancellation speaker emits a sound wave with the same amplitude but with inverted phase (also known as antiphase) relative to the original sound. The waves combine to form a new wave, in a process called interference, and effectively cancel each other out – an effect which is called destructive interference.

Modern active noise control is generally achieved through the use of analog circuits or digital signal processing. Adaptive algorithms are designed to analyze the waveform of the background aural or non aural noise, then based on the specific algorithm generate a signal that will either phase shift or invert the polarity of the original signal. This inverted signal (in antiphase) is then amplified and a transducer creates a sound wave directly proportional to the amplitude of the original waveform, creating destructive interference. This effectively reduces the volume of the perceivable noise.

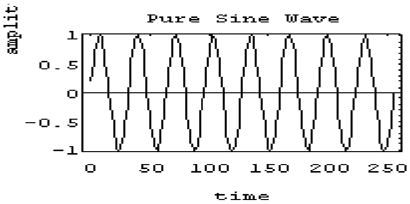
A noise-cancellation speaker may be co-located with the sound source to be attenuated. In this case it must have the same audio power level as the source of the unwanted sound in order to cancel the noise. Alternatively, the transducer emitting the cancellation signal may be located at the location where sound attenuation is wanted (e.g. the user's ear). This requires a much lower power level for cancellation but is effective only for a single user. Noise cancellation at other locations is more difficult as the three-dimensional wavefronts of the unwanted sound and the cancellation signal could match and create alternating zones of constructive and destructive interference, reducing noise in some spots while doubling noise in others. In small enclosed spaces (e.g. the passenger compartment of a car) global noise reduction can be achieved via multiple speakers and feedback microphones, and measurement of the modal responses of the enclosure.

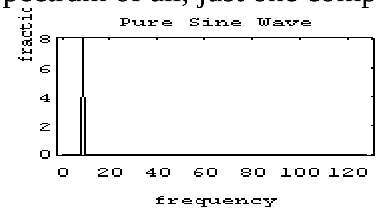
Noise control is an active or passive means of reducing sound emissions, often for personal comfort, environmental considerations or legal compliance. *Active* noise control is sound reduction using a power source. *Passive* noise control is sound reduction by noise-isolating materials such as insulation, sound-absorbing tiles, or a muffler rather than a power source.

Active noise cancelling is best suited for low frequencies. For higher frequencies, the spacing requirements for free space and zone of silence techniques become prohibitive. In acoustic cavity and duct based systems, the number of nodes grows rapidly with increasing frequency, which quickly makes active noise control techniques unmanageable. Passive treatments become more effective at higher frequencies and often provide an adequate solution without the need for active control.

FM Spectrum

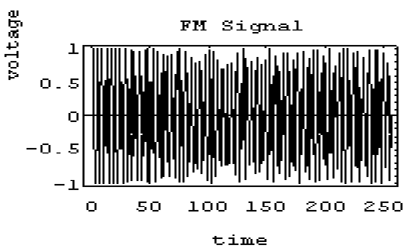
A spectrum represents the relative amounts of different frequency components in any signal. Its like the display on the graphic-equalizer in your stereo which has leds showing the relative amounts of bass, midrange and treble. These correspond directly to increasing frequencies (treble being the high frequency components). It is a well-know fact of mathematics, that any function (signal) can be decomposed into purely sinusoidal components (with a few pathological exceptions) . In technical terms, the sines and cosines form a complete set of functions, also known as a basis in the infinite-dimensional vector space of real-valued functions (gag reflex). Given that any signal can be thought to be made up of sinusoidal signals, the spectrum then represents the "recipe card" of how to make the signal from sinusoids. Like: 1 part of 50 Hz and 2 parts of 200 Hz. Pure sinusoids have the simplest spectrum of all, just one component:

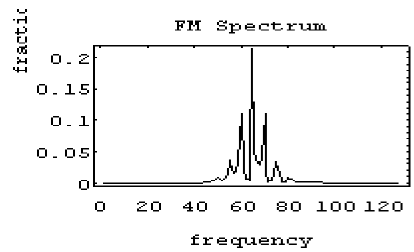




In this example, the carrier has 8 Hz and so the spectrum has a single component with value 1.0 at 8 Hz

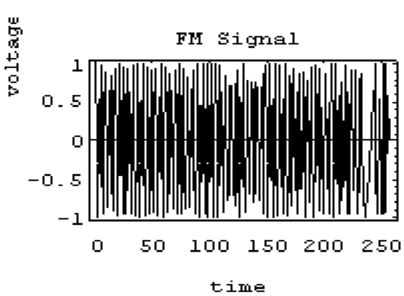
The FM spectrum is considerably more complicated. The spectrum of a simple FM signal looks like:

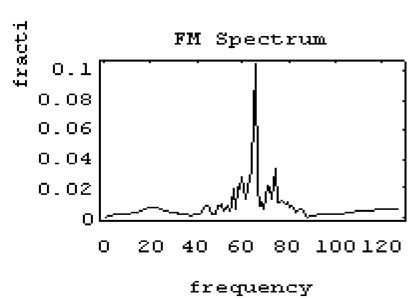




The carrier is now 65 Hz, the modulating signal is a pure 5 Hz tone, and the modulation index is 2. What we see are multiple side-bands (spikes at other than the carrier frequency) separated by the modulating frequency, 5 Hz. There are roughly 3 side-bands on either side of the carrier. The shape of the spectrum may be explained using a simple heterodyne argument: when you mix the three frequencies (f_c, f_m and Δf) together you get the sum and difference frequencies. The largest combination is $f_c + f_m + \Delta f$, and the smallest is $f_c - f_m - \Delta f$. Since $\Delta f = \beta f_m$, the frequency varies $(\beta + 1) f_m$ above and below the carrier.

A more realistic example is to use an audio spectrum to provide the modulation:





In this example, the information signal varies between 1 and 11 Hz. The carrier is at 65 Hz and the modulation index is 2. The individual side-band spikes are replaced by a more-or-less continuous spectrum. However, the extent of the side-bands is limited (approximately) to (β + 1) f_m above and below. Here, that would be 33 Hz above and below, making the bandwidth about 66 Hz. We see the side-bands extend from 35 to 90 Hz, so out observed bandwidth is 65 Hz.

You may have wondered why we ignored the smooth humps at the extreme ends of the spectrum. The truth is that they are in fact a by-product of frequency modulation (there is no random noise in this example). However, they may be safely ignored because they are have only a minute fraction of the total power. In practice, the random noise would obscure them anyway.

Example: FM Radio

FM radio uses frequency modulation, of course. The frequency band for FM radio is about 88 to 108 MHz. The information signal is music and voice which falls in the audio spectrum. The full audio spectrum ranges from 20 to 20,000 Hz, but FM radio limits the upper modulating frequency to 15 kHz (cf. AM radio which limits the upper frequency to 5 kHz). Although, some of the signal may be lost above 15 kHz, most people can't hear it anyway, so there is little loss of fidelity. FM radio maybe appropriately referred to as "high-fidelity."

If FM transmitters use a maximum modulation index of about 5.0, so the resulting bandwidth is 180 kHz (roughly 0.2 MHz). The FCC assigns stations) 0.2 MHz apart to prevent overlapping signals (coincidence? I think not!). If you were to fill up the FM band with stations, you could get 108 - 88 / .2 = 100 stations, about the same number as AM radio (107). This sounds convincing, but is actually more complicated.

FM radio is broadcast in stereo, meaning two channels of information. In practice, they generate three signals prior to applying the modulation:

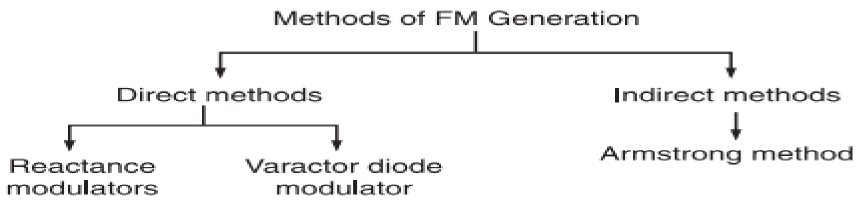
- the L + R (left + right) signal in the range of 50 to 15,000 Hz.
- a 19 kHz pilot carrier.
- the L-R signal centered on a 38 kHz pilot carrier (which is suppressed) that ranges from 23 to 53 kHz .

So, the information signal actually has a maximum modulating frequency of 53 kHz, requiring a reduction in the modulation index to about 1.0 to keep the total signal bandwidth about 200 kHz.

Generation of FM Wave

The FM modulator circuits used for generating FM signals can be divided into two categories such as: (i) The direct method or parameter variation method
(ii) The Indirect method or the Armstrong method

The classification of FM generation methods is shown below :



The Direct Method or Parameter Variation Method

In direct method or parameter variation method, the baseband or modulating signal directly modulates the carrier. The carrier signal is generated with the help of an oscillator circuit. This oscillator circuit uses a parallel tuned L-C circuit. Thus the frequency of oscillation of the carrier generation is governed by the expression:

$$\omega_c = \frac{1}{\sqrt{LC}} \tag{1}$$

Now, we can make the carrier frequency ωc to vary in accordance with the baseband or modulating signal x(t) if L or C is varied according to x(t). An oscillator circuit whose frequency is controlled by a modulating voltage is called voltage controlled oscillator (VCO). The frequency of VCO is varied according to the modulating signal simply by putting a shunt voltage variable capacitor with its tuned circuit. This voltage variable capacitor is called varactor or varicap. This type of property is exhibited by reverse biased semiconductor diodes. Also the capacitance of bipolar junction transistors (BJT) and field-effect transistors (FET) is varied by the Miller-effect. This miller capacitance may be utilized for frequency modulation. In addition to this, the electron tubes may also provide variable reactance (either it is inductive or capacitive) which is proportional to modulating or baseband signal. These type of tubes are called reactance tubes and may be used for FM generation. The inductance L of the tuned circuit may also be varied in accordance with the baseband or modulating signal x(t). The FM circuit using such inductors is called saturable reactor modulator. Frequency modulation can also be achieved from voltage controlled devices such as PIN diode, Klystron oscillators and multivibrators.

Reactance Modulator

In direct FM generation shown below, the instantaneous frequency of the carrier is changed directly in proportion with the message signal. For this, a device called voltage controlled oscillator (VCO) is used. A VCO can be implemented by using a sinusoidal oscillator with a tuned circuit having a high value of Q.

The frequency of this oscillator is changed by changing the reactive components involved in the tuned circuit. If L or C of a tuned circuit of an oscillator is changed in accordance with the amplitude of modulating signal then FM can be obtained across the tuned circuit as shown in figure 2.4.1 below.

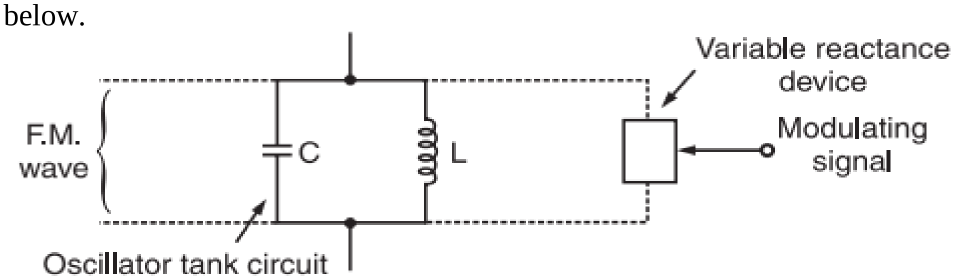
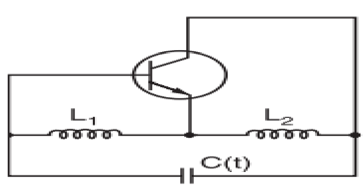


Figure: Principle of Reactance Modulator

A two or three terminal device is placed across the tuned circuit. The reactance of the device is varied proportional to modulating signal voltage. This will vary the frequency of the oscillator to produce FM. The devices used are FET, transistor or varactor diode. An example of direct FM is shown in figure 1 which uses a Hartley oscillator along with a varactor diode. The varactor diode is reverse biased. Its capacitance is dependent on the reverse voltage applied across it. This capacitance is shown by the capacitor C(t) in the figure below.

Figure: Hartley Oscillator



The FM transmitter has three basic sections.

1. The exciter section contains the carrier oscillator, reactance modulator and the buffer amplifier.

2. The frequency multiplier section, which features several frequency multipliers.

3. The power output section, which includes a low- level power amplifier, the final power amplifier, and the impedance matching network to properly load the power section with the antenna impedance.

The essential function of each circuit in the FM transmitter may be described as follows.

The Exciter

1. The function of the carrier oscillator is to generate a stable sine wave signal at the rest frequency, when no modulation is applied. It must be able to linearly change frequency when fully modulated, with no measurable change in amplitude.

2. The buffer amplifier acts as a constant high-impedance load on the oscillator to help stabilize the oscillator frequency. The buffer amplifier may have a small gain.

3. The modulator acts to change the carrier oscillator frequency by application of the message signal. The positive peak of the message signal generally lowers the oscillator's frequency to a point below the rest frequency, and the negative message peak raises the oscillator frequency to a value above the rest frequency. The greater the peak-to-peak message signal, the larger the oscillator deviation.

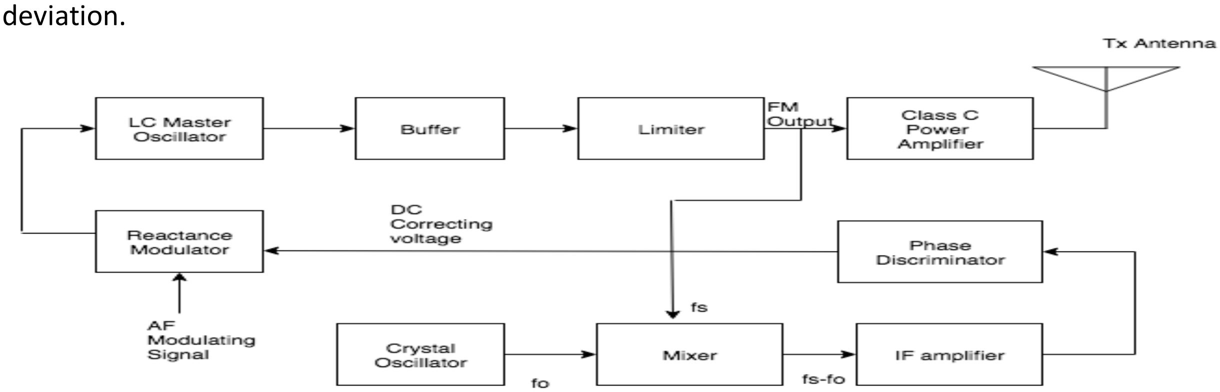


Figure: Reactance Modulator

Frequency multipliers are tuned-input, tuned-output RF amplifiers in which the output resonant circuit is tuned to a multiple of the input frequency in the above figure 2.4.3. Common frequency multipliers are 2x, 3x and 4x multiplication. A 5x Frequency multiplier is sometimes seen, but its extreme low efficiency forbids widespread usage. Note that multiplication is by whole numbers only. There can not a 1.5x multiplier, for instance. The final power section develops the carrier power, to be transmitted and often has a low-power amplifier driven the final power amplifier. The impedance matching network is the same as for the AM transmitter and matches the antenna impedance to the correct load on the final over amplifier.

Frequency Multiplier

A special form of class C amplifier is the frequency multiplier. Any class C amplifier is capable of performing frequency multiplication if the tuned circuit in the collector resonates at some integer multiple of the input frequency.

For example a frequency doubler can be constructed by simply connecting a parallel tuned circuit in the collector of a class C amplifier that resonates at twice the input frequency. When the collector

current pulse occurs, it excites or rings the tuned circuit at twice the input frequency. A current pulse flows for every other cycle of the input.
A Tripler circuit is constructed in the same way except that the tuned circuit resonates at 3 times the input - frequency. In this way, the tuned circuit receives one input pulse for every three cycles of oscillation it produces Multipliers can be constructed to increase the input frequency by any integer factor up to approximately 10. As' the multiplication factor gets higher, the power output of the multiplier decreases. For most practical applications, the best result is obtained with multipliers of 2 and 3.
Another way to look the operation of class C multipliers is .to .remember that the non-sinusoidal current pulse is rich in harmonics. Each time the pulse occurs, the second, third, fourth, fifth, and higher harmonics are generated. The purpose of the tuned circuit in the collector is to act as a filter to select the desired harmonics.
In many applications a multiplication factor greater than that achievable with a single multiplier stage is required. In such cases two or more multipliers are cascaded to produce an overall multiplication of 6. In the second example, three multipliers provide an overall multiplication of 30. The total multiplication factor is simply the product of individual stage multiplication factors.

Reactance Modulator
The reactance modulator takes its name from the fact that the impedance of the circuit acts as a reactance (capacitive or inductive) that is connected in parallel with the resonant circuit of the Oscillator. The varicap can only appear as a capacitance that becomes part of the frequency determining branch of the oscillator circuit. However, other discrete devices can appear as a capacitor or as an inductor to the oscillator, depending on how the circuit is arranged. A colpitts oscillator uses a capacitive voltage divider as the phase-reversing feedback path and would most likely tapped coil as the phase-reversing element in the feedback loop and most commonly uses a modulator that appears inductive.

Frequency of oscillations of the Hartley oscillator is given by:

$$f_i(t) = \frac{1}{2\pi\sqrt{(L_1 + L_2)C(t)}}$$

where C(t) = C + Cvarector
This means that C(t) is the effective capacitance of the fixed tuned circuit capacitance C and the varactor diode capacitance C varector.
Let the relation between the modulating voltage x(t) = 0 and the capacitance C(t) be represented as under:

$C(t) = C - k_c\, x(t)$ Where C = total capacitance when x(t)

kc is the sensitivity of the varactor capacitance to change in voltage
Substituting expression for C(t) in equation(1) , we get

$$f_i(t) = \frac{1}{2\pi\sqrt{(L_1 + L_2)(C - k_c\, x(t))}} = \frac{1}{2\pi\left[\sqrt{(L_1 + L_2)\,C - (L_1 + L_2)k_c\, x(t)}\right]}$$

or $$f_i(t) = \frac{1}{2\pi\sqrt{(L_1 + L_2)\,C}\left[1 - \frac{k_c\, x(t)}{C}\right]^{1/2}}$$

But, let

$$\frac{1}{2\pi\sqrt{(L_1 + L_2)\,C}} = f_0$$

which is the oscillator frequency in absence of the modulating signal [x(t) = 0]. Therefore, we have,

$$f_i(t) = f_0 \left[1 - \frac{k_c}{C} x(t)\right]^{-1/2}$$

If the maximum change in the capacitance corresponding to the modulating wave is assumed to be small as compared to the unmodulated capacitance C then equation (2) for fi (t) can be approximated as under:

$$f_i(t) = f_0 \left[1 + \frac{k_c}{2C} x(t)\right]$$

or $$f_i(t) = f_0 + \frac{f_0\, k_c}{2\, C} \,.\, x(t)$$

Now, let us define

$$\frac{f_0\, k_c}{2C} = k_f$$

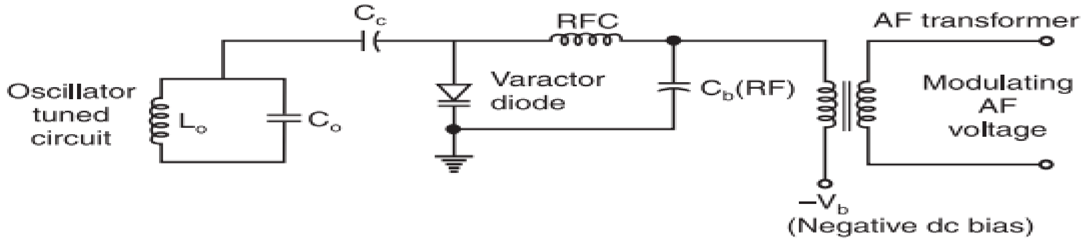
Therefore, we have: $$f_i(t) = f_0 + k_f\, x(t)$$

Where kf is known as the frequency sensitivity of the modulator.

Varactor Diode Modulator

- Varactor diode modulator is the direct method of FM generation wherein the carrier frequency is directly varied by the modulating signal.
- A varactor diode is a semiconductor diode whose junction capacitance varies linearly with applied voltage when the diode is reverse biased.

- Varactor diodes are used along with reactance modulator to provide automatic frequency correction for an FM transmitter. The varactor diode modulator circuit is shown in Figure below for generation of FM wave.



A varactor diode is a semiconductor diode whose junction capacitance varies linearly with the applied bias and the varactor diode must be reverse biased.

Working Operation

Varactor diode is arranged in reverse bias to offer junction capacitance effect. The modulating voltage which is in series with the varactor diode will vary the bias and hence the junction capacitance, resulting the oscillator frequency to change accordingly. The external modulating AF voltage adds to and subtracts from the dc bias, which changes the capacitance of the diode and thus the frequency of oscillation. Positive alternations of the modulating signal increase the reverse bias on the varactor diode, which decreases its capacitance and increases the frequency of oscillation.

Conversely, negative alternations of the modulating signal decrease the frequency of oscillation. The RFC and capacitor C b act as a filter which transmits only the AF variations to the varactor diode and blocks high frequency RF voltage from reaching the AF stage. The varactor diode FM modulators are widely accepted because they are simple to use, reliable and have the stability of a crystal oscillator. This method of FM generation is direct because the oscillator frequency is varied directly by the modulating signal, and the magnitude of frequency change is proportional to the amplitude of the modulating signal voltage.

Varactor diode modulator is used for automatic frequency control and remote tuning. The drawback of varactor diode modulator is that since it uses a crystal, the peak frequency deviation is limited to relatively small values. Thus they are used mostly for low index applications such as two way mobile radio. Also since they are a two terminal device, the applications are quite limited. The varactor diode is reverse biased by the negative dc source –Vb.

The modulating AF voltage appears in series with the negative supply voltage. Hence, the voltage applied across the varactor diode varies in proportion with the modulating voltage .This will vary the junction capacitance of the varactor diode. The varactor diode appears in parallel with the oscillator tuned circuit. Hence the oscillator frequency will change with change in varactor diode capacitance and FM wave is produced. The RFC will connect the dc and modulating signal to the varactor diode but it offers very high impedance at high oscillator frequency. Therefore, the oscillator circuit is isolated from the dc bias and modulating signal.

Indirect Method of WBFM Generation

This method is called as Indirect Method because we are generating a wide band FM wave indirectly. This means, first we will generate NBFM wave and then with the help of frequency multipliers we will get WBFM wave. The block diagram of generation of WBFM wave is shown in the following figure.

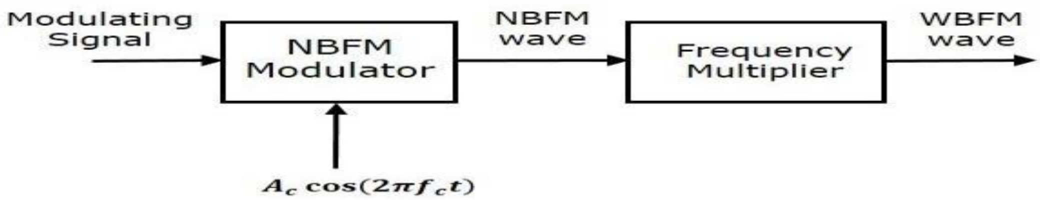


Figure: Varactor Diode Modulator

This block diagram shown in figure 2.4.5 contains mainly two stages. In the first stage, the NBFM wave will be generated using NBFM modulator. We have seen the block diagram of NBFM modulator at the beginning of this chapter. We know that the modulation index of NBFM wave is less than one. Hence, in order to get the required modulation index (greater than one) of FM wave, choose the frequency multiplier value properly. Frequency multiplier is a non-linear device, which produces an output signal whose frequency is 'n' times the input signal frequency. Where, 'n' is the multiplication factor. If NBFM wave whose modulation index ββ is less than 1 is applied as the input of frequency multiplier, then the frequency multiplier produces an output signal, whose modulation index is 'n' times ββ and the frequency also 'n' times the frequency of WBFM wave. Sometimes, we may require multiple stages of frequency multiplier and mixers in order to increase the frequency deviation and modulation index of FM wave.
The part of the Armstrong FM transmitter (Armstrong phase modulator) which is expressed in dotted lines describes the principle of operation of an Armstrong phase modulator. It should be noted, first that the output signal from the carrier oscillator is supplied to circuits that perform the task of modulating the carrier signal. The oscillator does not change frequency, as is the case of direct FM. These points out the major advantage of phase modulation (PM), or indirect FM, over direct FM. That is the phase modulator is crystal controlled for frequency.

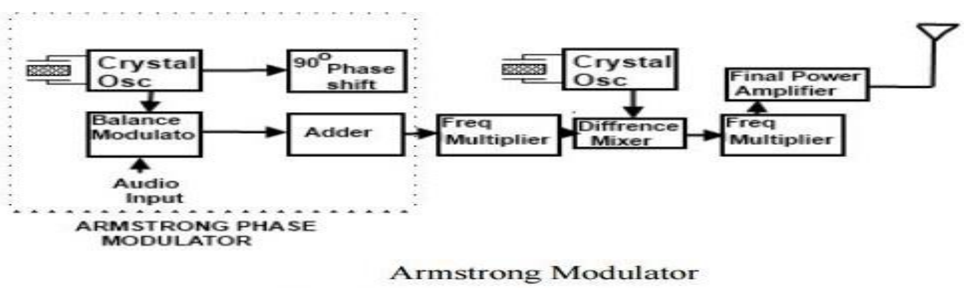


Figure: Armstrong Modulator

The crystal-controlled carrier oscillator signal is directed to two circuits in parallel. This signal (usually a sine wave) is established as the reference past carrier signal and is assigned a value 0°.The balanced modulator is an amplitude modulator used to form an envelope of double side-bands and to suppress the carrier signal (DSSC). This requires two input signals, the carrier signal and the modulating message signal shown in te figure 2.4.6. The output of the modulator is connected to the adder circuit; here the 90° phase-delayed carriers signal will be added back to replace the suppressed carrier. The act of delaying the carrier phase by 90° does not change the carrier frequency or its wave-shape. This signal identified as the 90° carrier signal.

The carrier frequency change at the adder output is a function of the output phase shift and is found by. fc = Δθfs (in hertz) .When θ is the phase change in radians and fs is the lowest audio modulating frequency. In most FM radio bands, the lowest audio frequency is 50Hz. Therefore, the carrier frequency change at the adder output is 0.6125 x 50Hz = ± 30Hz since 10% AM represents the upper limit of carrier voltage change, then ± 30Hz is the maximum deviation from the modulator for PM. The 90° phase shift network does not change the signal frequency because the components and resulting phase change are constant with time. However, the phase of the adder output voltage is in a continual state of change brought about by the cyclical variations of the message signal, and during the time of a phase change, there will also be a frequency change.

AM DETECTORS

The diode detector is the simplest and most basic form of amplitude modulation, AM signal detector and it detects the envelope of the AM signal.
The AM diode detector can be built from just a diode and a few other components and as a result it is a very low cost circuit block within an overall receiver. In the early days of radio, these signal detectors were made using discrete components, but modern radios will use integrated circuits with inbuilt detectors.
As a result of its cost and convenience, the AM diode envelope detector has been widely used for many years in transistor portable radios. Although its simplicity has been the main reason for its widespread use, its performance is not as good as other types of AM detector / demodulator, particularly with respect to the distortion levels.
Not only is the basic AM diode signal detector used for AM envelope detection, but are also widely used n RF circuits in general for signal level detection.

AM diode detector basics

The AM diode detector is an envelope detector – it provides an output of the envelope of the signal. As such the diode detector or demodulator is able to provide an output proportional to the amplitude of the envelope of the amplitude modulated signal.
As the name implies, the main component within the AM diode detector is a semiconductor diode, although in the days of valve / tube technology, diodes using this form of technology were also used.
The signal diode detector consists of two main elements to the circuit:

- ***Diode / rectifier:*** The diode in the detector serves to that enhances one half of the received signal over the other. In many instances Schottky diodes are used for this form of detector, because signal levels may be low, and Schottky diodes have a much lower turn on voltage (typically around 0.2 V) than standard silicon diodes (typically around 0.7 or 0.7 V).
- ***Low pass filter:*** The low pass filter is required to remove the high frequency elements that remain within the signal after detection / demodulation. The filter usually consists of a very simple RC network but in some cases It can be provided simply by relying on the limited frequency response of the circuitry following the rectifier. As the capacitor in the circuit stores the voltage, the output voltage reflects the peak of the waveform. Sometimes these circuits are used as peak detectors.

When selecting the value of the capacitor used int he circuit, it should be large enough to hold the peak of the RF waveform, but not so large that it attenuates any modulation on the signal, i.e. it should act as a filter for the RF carrier and not the audio modulation.

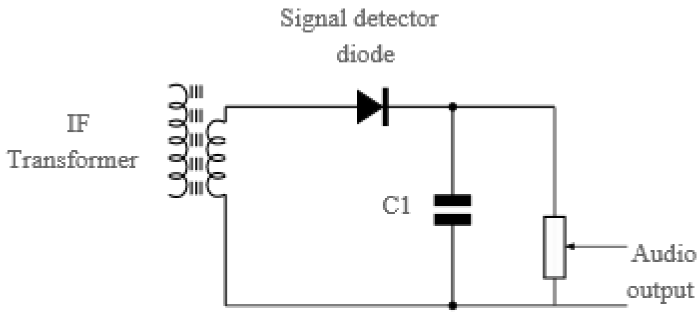


Fig: Circuit of an envelope detector as used in an AM radio receiver.

The circuit typically has relatively high source impedance. When linking the circuit to a following stage of the circuit, care should be taken not to land the detector too much otherwise the operation will be impaired.
Normally a resistor is placed across the capacitor - this may either be the load of the next stage, a volume control, or resistor in the circuit. This level of this should be determined by calculating the time constant of the capacitor and the load. This should be between the RF signal and audio modulation so that the RF is satisfactorily removed, but the audio modulation is left untouched.
It is worth noting in this circuit that the secondary of the transformer provides a DC return to ground. Sometimes when the AM signal detector is used using a capacitor connection to the previous stage, then a resistor or choke (inductor) to ground must be used at the input so that a DC return path is provided. If not the circuit will not operate correctly.

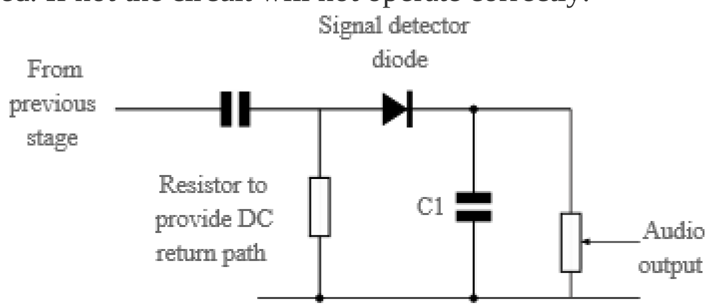


Fig: Capacitor coupled envelope signal detector showing resistor providing DC return path.

The value of the resistor on the input providing the DC return path is normally critical, but it can help provide the require match without absorbing too much signal.

AM diode detection process

In rectifying the RF signal, the AM diode detector provides an output equivalent to the envelope of one half of the signal, i.e. it is an envelope detector.
In view of the operation of the diode detector, it may sometimes be referred to as an envelope detector.
The incoming amplitude modulated RF signal consists of a waveform of both positive and negative going voltages as shown. Any audio transducer would not respond to this.

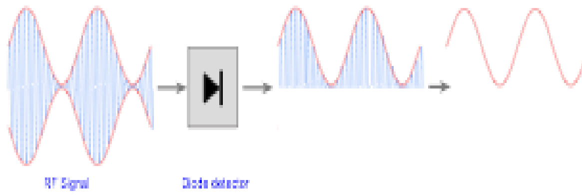


Fig: AM diode envelope detection process

The diode envelope detector rectifies the waveform leaving only the positive or negative half of the waveform.
The high frequency element of this is then filtered out, typically using a capacitor which forms the low pass filter and effectively 'fills in' the high frequency elements, leaving a waveform to which a transducer like a pair of earphones or a loudspeaker could respond to and convert into sound waves.

Vestigial Sideband (VSB) Demodulation
In case of SSB modulation, when a sideband is passed through the filters, the band pass filter may not work perfectly in practice. As a result of which, some of the information may get lost. Hence to avoid this loss, a technique is chosen, which is a compromise between **DSB-SC** and **SSB**, called as **Vestigial Sideband (VSB)** technique. The word vestige which means "a part" from which the name is derived.
Both of the sidebands are not required for the transmission, as it is a waste. But a single band if transmitted, leads to loss of information. Hence, this technique has evolved.
Vestigial Sideband Modulation or **VSB Modulation** is the process where a part of the signal called as **vestige** is modulated, along with one sideband. A VSB signal can be plotted as shown in the following figure.

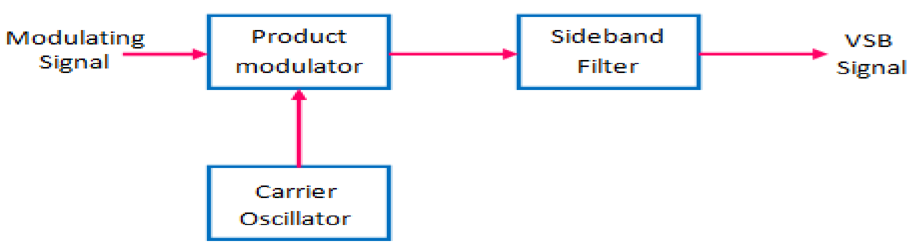


Fig: Modulation/Generation of VSB signal

Along with the upper sideband, a part of the lower sideband is also being transmitted in this technique. A guard band of very small width is laid on either side of VSB in order to avoid the interferences. VSB modulation is mostly used in television transmissions.

The transmission bandwidth of VSB modulated wave is represented as −

B=(f{m}+ f{v}) Hz,   Where, $\mathbf{f_m}$ = Message bandwidth and $\mathbf{f_v}$ = Width of the vestigial sideband

VSB Modulation − Advantages

- Highly efficient.
- Reduction in bandwidth.

- Filter design is easy as high accuracy is not needed.
- The transmission of low frequency components is possible, without difficulty.
- Possesses good phase characteristics.

VSB Modulation − Disadvantages

- Bandwidth when compared to SSB is greater.
- Demodulation is complex.

VSB Modulation − Application

The most prominent and standard application of VSB is for the transmission of **television signals**. Also, this is the most convenient and efficient technique when bandwidth usage is considered.

Demodulation of VSB Wave

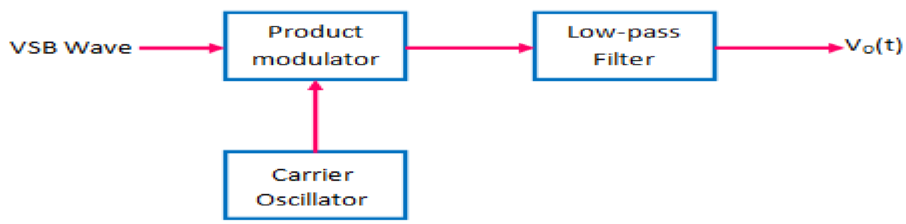


Fig.4 : VSB demodulator

VSB Demodulation Application

The VSB modulated wave is passed through a product modulator where it is multiplied with the locally generated synchronous carrier.

Hence, the output of the product modulator is given by:

$$m(t) = s(t) \times c(t) = s(t)V_c \cos(2\pi f_c t)$$

Taking the Fourier transform of both sides, we get

$$M(f) = S(f) \times \left[\frac{1}{2}\delta(f+f_c) + \frac{1}{2}\delta(f-f_c)\right] = \frac{1}{2}S(f+f_c) + \frac{1}{2}S(f-f_c)$$

But

$$S(f) = \frac{V_c}{2}[X(f-f_c) + X(f+f_c)]H(f)$$

Hence, we have

$$M(f) = \frac{V_c}{2}[X(f-2f_c)H(f-f_c) + X((f+2f_c)H(f+f_c)] + \frac{V_c}{4}[X(f)[H(f-f_c) + H(f+f_c)]$$

The first term in the above expression represents the VSB modulated wave, corresponding to a carrier frequency of $2f_c$.This term will be eliminated by the filter to produce output $v_o(t)$.
The second term in the above expression for M(f) represents the spectrum of demodulated VSB output .

Therefore ,

$$V_O(f) = \frac{V_c}{4}[X(f)[H(f - f_c) + H(f + f_c)]$$

This spectrum is shown in fig.5 .

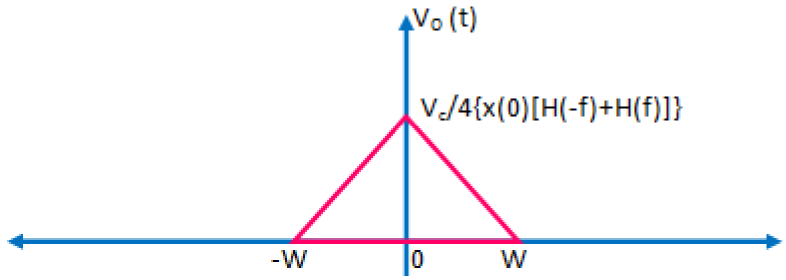


fig 5: Spectrum of VSB Demodulator

In order to obtain the undistorted message signal x(t) at the output of the demodulator, $V_o(f)$ should be a scaled version of X(f) .
For this the transfer function H(f) should satisfy the following conditions :

$$H(f - f_c) + H(f + f_c) = 2H(f + f_c)$$

Where $H(f_c)$ is constant .