Unit 2

Amplitude modulation

2.1 Introduction

Modulation is a process of varying one of the characteristics of high frequency sinusoidal (the carrier) in accordance with the instantaneous values of the modulating (the information) signal. The high frequency carrier signal is mathematically represented by the equation 2.1.

$$c(t) = A_c \cos(2\pi f_c t + \phi) \qquad \text{--- (2.1)}$$

Where $c(t)$--instantaneous values of the cosine wave

A_c--its maximum value

f_c--carrier frequency

ϕ --phase relation with respect to the reference

Any of the last three characteristics or parameters of the carrier can be varied by the modulating (message) signal, giving rise to amplitude, frequency or phase modulation respectively.

Need for modulation:

1. Practicability of antenna

In the audio frequency range, for efficient radiation and reception, the transmitting and receiving antennas must have sizes comparable to the wavelength of the frequency of the signal used. It is calculated using the relation $f\lambda = c$. The wavelength is 75 meters at 1MHz in the broadcast band, but at 1 KHz, the wavelength turns out to be 300 Kilometers. A practical antenna for this value of wavelength is unimaginable and impossible.

2. Modulation for ease of radiation

For efficient radiation of electromagnetic waves, the antenna dimension required is of the order of $\lambda/4$ to $\lambda/2$. It is possible to construct practical antennas only by increasing the frequency of the base band signal.

3. Modulation for multiplexing

The process of combining several signals for simultaneous transmission on a single channel is called multiplexing. In order to use a channel to transmit the different base band signals (information) at the same time, it becomes necessary to translate different signals so as to make them occupy different frequency slots or bands so that they do not interfere. This is a accomplished by using carrier of different frequencies.

4. Narrow banding:

Suppose that we want to transmit audio signal ranging from 50 - 10^4 Hz using suitable antenna. The ratio of highest to lowest frequency is 200. Therefore an antenna suitable for use at one end of the frequency range would be entirely too short or too long for the other end. Suppose that the audio spectrum is translated so that it occupies the range from $50+10^6$ to 10^4+10^6 Hz. Then the ratio of highest to lowest frequency becomes 1.01. Thus the process of frequency translation is useful to change wideband signals to narrow band signals.

At lower frequencies, the effects of flicker noise and burst noise are severe.

2.2 Amplitude modulation

In amplitude modulation, the amplitude of the carrier signal is varied by the modulating/message/information/base-band signal, in accordance with the instantaneous values of the message signal. That is amplitude of the carrier is made proportional to the instantaneous values (amplitude) of the modulating signal.

If *m(t)* is the information signal and $c(t) = A_c \cos(2\pi f_c t + \phi)$ is the carrier, the amplitude of the carrier signal is varied proportional to the $m(t)$.

The peak amplitude of carrier after modulation at any instant is given by [$A_c + m(t)$]. The carrier signal after modulation or the modulated signal is represented by the equation 2.2.

$$s(t) = [A_c + m(t)]\cos(2\pi fct + \phi) \quad \text{-- (2.2)}$$

$$s(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t + \phi) \quad \text{-- (2.3)}$$

where $k_a = \dfrac{1}{A_c}$ is called amplitude sensitivity of the modulator.

The equation (2.3) is the standard expression for Amplitude Modulated signal.

Let $m(t) = A_m \cos(2\pi f_m t)$ be the message signal of frequency f_m and peak amplitude A_m. Then single-tone modulated signal is given by the equation 2.4.

$$s(t) = Ac[1 + k_a A_m \cos(2\pi f_m t)]\cos(2\pi f_c t + \phi)$$

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

$$s(t) = Ac[1 + m\cos(2\pi f_m t)]\cos(2\pi f_c t + \phi) \qquad \text{--- (2.4)}$$

where $m = \frac{A_m}{A_c}$ is called **modulation index** or **depth of modulation.**

The modulation index m of AM system is defined as the ratio of peak amplitude of message signal to peak amplitude of carrier signal.

$$m = \frac{A_m}{A_c} \qquad \text{--- (2.5)}$$

The following figure 2.1 shows the message, carrier and amplitude modulated waveforms.

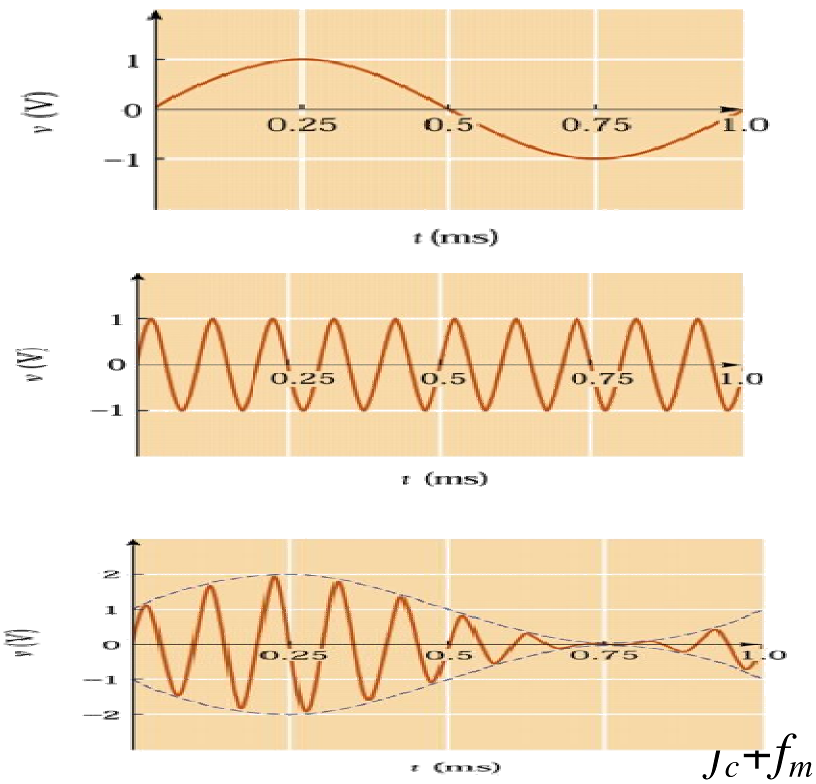


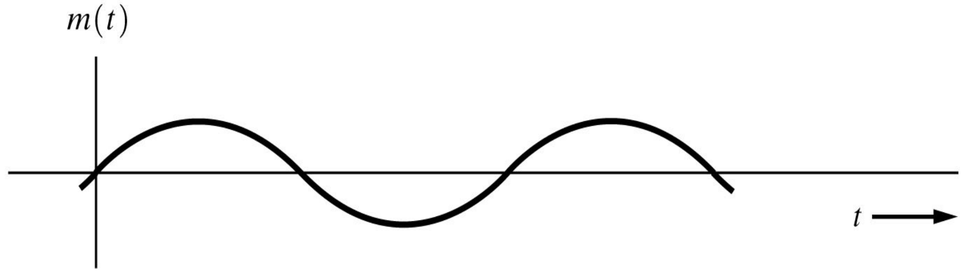
Figure 2.1: message, carrier and amplitude modulated signal

Note:

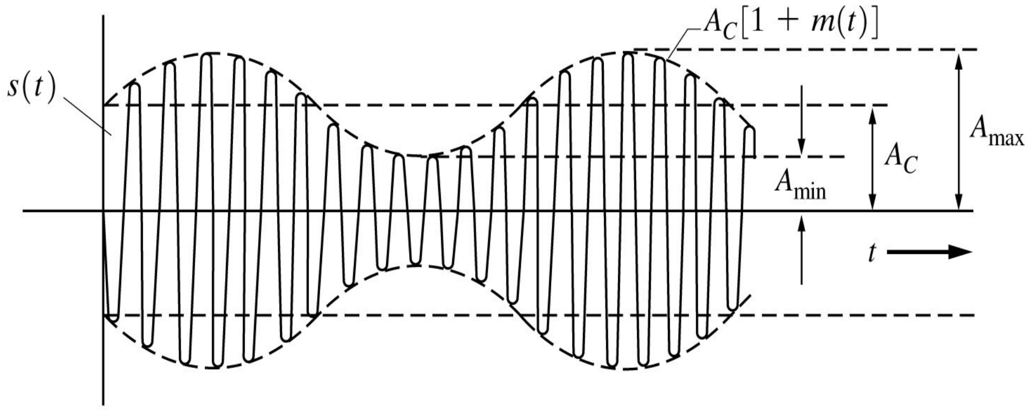
(1) m is also called depth of modulation.

(2) m specifies the system clarity. As m increases, the system clarity also increases.

Consider the Amplitude Modulated waveform shown in figure 2.2.



(a) Sinusoidal Modulating Wave



(b) Resulting AM Signal

Figure 2.2: Message and amplitude modulated signal

We have the modulation index given by

$$m = \frac{A_m}{A_c} \qquad \text{--- (2.6)}$$

From the figure 2.2, we get

$$A_m = \frac{A_{\max} - A_{\min}}{2} \qquad \text{--- (2.7)}$$

$$A_c = A_{\max} - A_m \qquad \text{--- (2.8)}$$

$$A_c = A_{\max} - \frac{A_{\max} - A_{\min}}{2}$$

$$A_c = \frac{A_{\max} + A_{\min}}{2} \qquad \text{--- (2.9)}$$

Dividing the equation (2.7) by (2.9), we get

$$m = \frac{A_m}{A_c} = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}} \quad \text{--- (2.10)}$$

Here $A_{\max}$ is the maximum amplitude and $A_{\min}$ is minimum amplitude of the modulated signal.

Modulation index m has to be governed such that it is always less than unity; otherwise it results in a situation known as 'over-modulation' ($m>1$). The over-modulation occurs, whenever the magnitude of the peak amplitude of the modulating signal exceeds the magnitude of the peak amplitude of the carrier signal. The signal gets distorted due to over modulation. Because of this limitation on 'm', the system clarity is also limited. The AM waveforms for different values of modulation index *m* are as shown in figure 2.3.

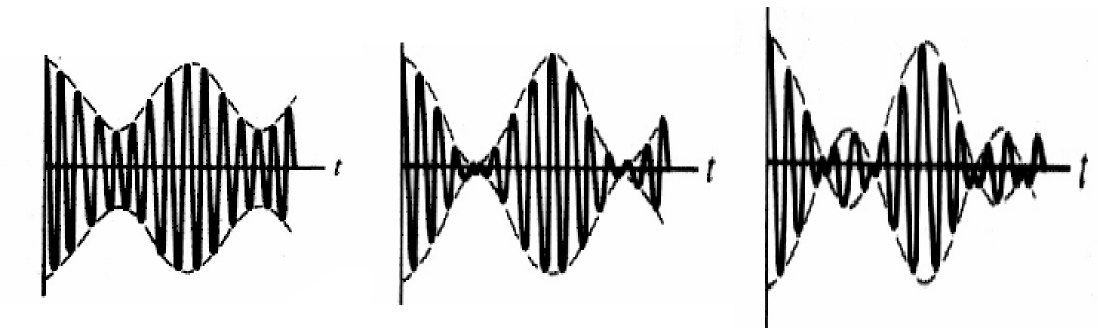


Figure 2.3: AM waveforms for different values of m

Note: If the modulation index exceeds unity the negative peak of the modulating waveform is clipped and $[A_c + m(t)]$ goes negative, which mathematically appears as a phase reversal rather than a clamped level.

Example 1.1

A modulating signal consists of a symmetrical triangular wave, which has zero dc component and peak-to-peak voltage 11v. It is used to amplitude modulate a carrier of peak voltage 10v. Calculate the modulation index?

The amplitude of the modulating signal is $\frac{11}{2} = 5.5 volts$

The modulation index is $m = \frac{A_m}{A_c} = \frac{5.5}{10} = 0.55$

1.3 Single tone Amplitude Modulation/ Sinusoidal AM

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

$$m(t) = A_m \cos(2\pi f_m t) \qquad \text{--- (2.11)}$$

where A_m is peak amplitude of the sinusoidal modulating wave

f_m is the frequency of the sinusoidal modulating wave

Let A_c be the peak amplitude and f_c be the frequency of the high frequency carrier signal. Then the corresponding single-tone AM wave is given by

$$s(t) = A_c[1 + m\cos(2\pi f_m t)]Cos(2\pi f_c t) \qquad \text{--- (2.12)}$$

Let $A_{\max}$ and $A_{\min}$ denote the maximum and minimum values of the envelope of the modulated wave. Then from the above equation (2.12), we get

$$\frac{A_{\max}}{A_{\min}} = \frac{A_c(1+m)}{A_c(1-m)}$$

$$m = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}}$$

Expanding the equation (2.12), we get

$$s(t) = A_c \cos(2\pi f_c t) + \frac{1}{2} m A_c \cos[2\pi(f_c + f_m)t] + \frac{1}{2} m A_c \cos[2\pi(f_c - f_m)t] \qquad \text{--- (2.13)}$$

The Fourier transform of $s(t)$ is obtained as follows.

$$s(f) = \frac{1}{2} A_c[\delta(f - f_c) + \delta(f + f_c)] + \frac{1}{4} m A_c[\delta(f - f_c - f_m) + \delta(f + f_c + f_m)]$$

$$+ \frac{1}{4} m A_c[\delta(f - f_c + f_m) + \delta(f + f_c - f_m)] \qquad \text{--- (2.14)}$$

Thus the spectrum of an AM wave, for the special case of sinusoidal modulation consists of delta functions at $\pm f_c$, $f_c \pm f_m$, and $-f_c \pm f_m$. The spectrum for positive frequencies is as shown in figure 2.4.

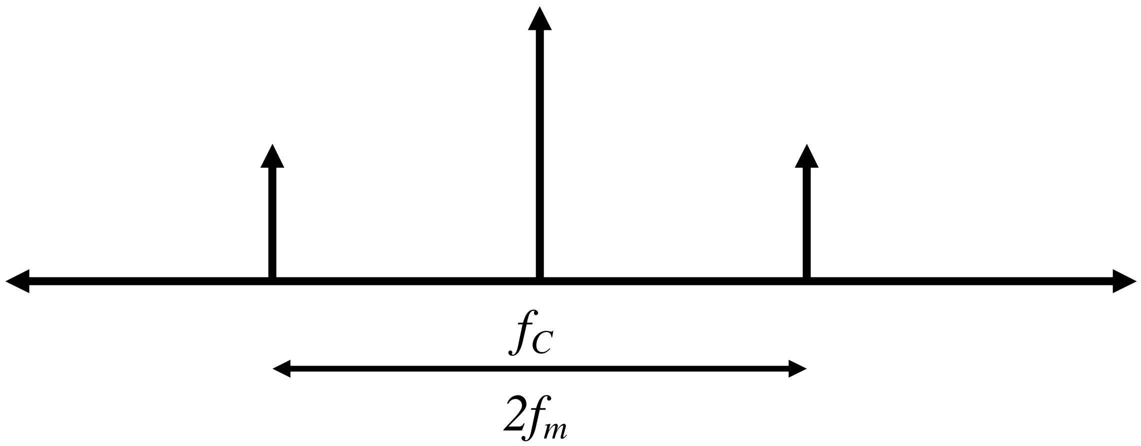


Figure 2.4: Frequency-domain characteristics of single tone AM

1.4 Frequency spectrum of AM wave:

Consider the standard expression for AM wave $s(t) = A_c[1 + k_a A_m \cos(2\pi f_m t)]\cos(2\pi f_c t)$. The carrier frequency f_c is much greater than the highest frequency component W of the message signal.

$$i.\ e., \quad fc >> W$$

W is called the message bandwidth.

The Fourier transform $S(f)$ of AM wave $s(t)$ is given by

$$s(f) = \frac{A_c}{2}[\delta(f - f_c) + \delta(f + f_c)] + \frac{k_a A_c}{2}[M(f - f_c) + M(f + f_c)] \quad \text{--- (2.15)}$$

Suppose that the base band signal $m(t)$ is band limited to the interval $-W \le f \le W$ as shown in figure 2.5.

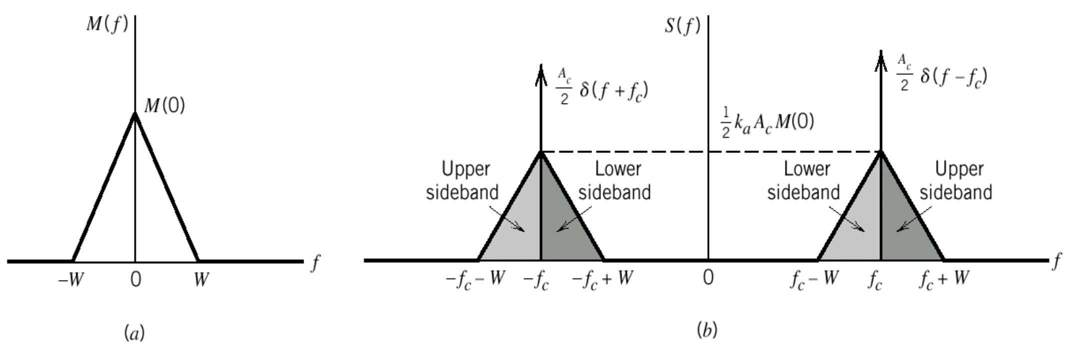


Figure 2.5: Spectrum of message and AM waves

From the equation (2.15), the spectrum of AM wave obtained is as shown in figure 2.5, for $f_c > W$. This spectrum consists of two delta functions weighted by the factor $A_c/2$, and occurring at $\pm f_c$, and two versions of the base band spectrum translated in frequency by $\pm f_c$. From the spectrum, the following points are noted.

(i) For positive frequencies, the highest frequency component of the AM wave is $f_c + W$, and the lowest frequency component is $f_c - W$. The difference between these two frequencies defines the transmission bandwidth B_T for an AM wave, which is exactly twice the message bandwidth *W*.

$$\therefore B_T = 2W \qquad \text{---(2.16)}$$

(ii) For positive frequencies, the portion of the spectrum of an AM wave lying above the carrier frequency f_c, is referred to as the Upper Side Band (USB), where as the symmetric portion below f_c, is called the Lower Side Band (LSB). For negative frequencies, the USB is the portion of the spectrum below $-f_c$ and the LSB is the portion above $-f_c$. The condition $f_c > W$ ensures that the side bands do not overlap.

The AM wave $s(t)$ is a voltage or current wave. In either case, the average power delivered to 1Ω resistor by $s(t)$ is comprised of three components.

$$\text{Carrier power} = \frac{A^2{}_c}{2}$$

$$\text{Upper side-frequency power} = \frac{m^2 A^2{}_c}{8}$$

$$\text{Lower side-frequency power} = \frac{m^2 A^2{}_c}{8}$$

Example 1.2 A carrier wave of frequency 10 MHz and peak value 10V is amplitude modulated by a 5 KHz sine wave of amplitude 6V. Determine the modulation index and amplitude of the side frequencies.

$$m = \frac{A_m}{A_c} = \frac{6}{10} = 0.6$$

The side frequencies are 10.005 MHz and 9.995 MHz.

The amplitude of side frequencies is given by

$$\frac{mAc}{2} = \frac{0.6*10}{2} = 3volts$$

1.5 Average power for sinusoidal AM (Power relations in AM)

Consider the expression for single tone/sinusoidal AM wave

$$s(t) = A_c Cos(2\pi f_c t) + \frac{1}{2} mA_c Cos[2\pi(f_c + f_m)t] + \frac{1}{2} mA_c Cos[2\pi(f_c - f_m)t] \text{--- (2.17)}$$

This expression contains three components. They are carrier component, upper side band and lower side band. Therefore Average power of the AM wave is sum of these three components.

Therefore the total power in the amplitude modulated wave is given by

$$Pt = \frac{V^2{}_{car}}{R} + \frac{V^2{}_{LSB}}{R} + \frac{V^2{}_{USB}}{R} \quad \text{--- (2.18)}$$

Where all the voltages are rms values and R is the resistance, in which the power is dissipated.

$$P_C = \frac{V^2{}_{car}}{R} = \frac{\left(A_c/\sqrt{2}\right)^2}{R} = \frac{A_c^2}{2R}$$

$$P_{LSB} = \frac{V_{LSB}^2}{R} = \left(\frac{mA_c}{2\sqrt{2}}\right)^2 \frac{1}{R} = \frac{m^2 A_c^2}{8R} = \frac{m^2}{4} P_c$$

$$P_{USB} = \frac{V_{USB}^2}{R} = \left(\frac{mA_c}{2\sqrt{2}}\right)^2 \frac{1}{R} = \frac{m^2 A_c^2}{8R} = \frac{m^2}{4} P_c$$

Therefore total average power is given by

$$P_t = P_c + P_{LSB} + P_{USB}$$

$$P_t = P_c + \frac{m^2}{4} P_c + \frac{m^2}{4} P_c$$

$$P_t = P_c\left(1 + \frac{m^2}{4} + \frac{m^2}{4}\right)$$

$$P_t = P_c\left(1 + \frac{m^2}{2}\right) \quad \text{--- (2.19)}$$

The ratio of total side band power to the total power in the modulated wave is given by

$$\frac{P_{SB}}{P_t} = \frac{P_c\left(m^2/2\right)}{P_c\left(1+m^2/2\right)}$$

$$\frac{P_{SB}}{P_t} = \frac{m^2}{2+m^2} \qquad \text{--- (2.20)}$$

The ratio is called the efficiency of AM system and it takes maximum value of *33%* at *m=1*.

Example 1.3 A broadcast radio transmitter radiates 10KW, when the modulation percentage is 60. How much of this is carrier power.

$$P_c = \frac{P_t}{1+m^2/2}$$

$$P_c = \frac{10}{1+0.6^2/2} = 8.47KW$$

Example 1.4 A radio transmitter radiates 10 KW and carrier power is 8.5 KW. Calculate modulation index.

$$m = \left[\left(\frac{P_t}{P_c}-1\right)2\right]^{1/2}$$

$$m = \left[\left(\frac{10\times10^3}{8.5\times10^3}-1\right)2\right]^{1/2}$$

$$m = 0.59$$

1.6 Effective voltage and current for sinusoidal AM

In AM systems, the modulated and unmodulated currents are necessary to calculate the modulation index from them.

The effective or rms value of voltage E_t of the modulated wave is defined by the equation $\frac{E_t^{\ 2}}{R} = P_t$.

Similarly the effective or root mean square voltage E_c of carrier component is defined by $\frac{E_c^{\,2}}{R} = P_c$.

Now using the relation, $P_t = P_c\left(1+\frac{m^2}{2}\right)$

We get, $E_t = E_c\left(\sqrt{1+\frac{m^2}{2}}\right)$

A similar argument applied to currents, yields $I_t = I_c\left(\sqrt{1+\frac{m^2}{2}}\right)$

Where I_t is the rms current of modulated wave and I_c is the rms current of unmodulated carrier.

Note: The maximum power in the AM wave is $P_t = 1.5P_c$, when $m=1$. This is important, because it is the maximum power that relevant amplifiers must be capable of handling without distortion.

Example 1.5 A 400 W carrier is modulated to a depth of 7.5 %. Calculate total power in the modulated wave. (**Ans:** $\boldsymbol{P_t}$**=512.5w**)

Example 1.6 The antenna current of an AM transmitter is 8 Amps, when only the carrier is sent, but it increases to 8.93A, when the carrier is modulated by a single sine wave. Find percentage modulation. Determine the antenna current when the percent modulation changes to 0.8. (**Ans:** $\boldsymbol{m}$**=70.1%,** $\boldsymbol{I_t}$**=9.19A**)

1.7 Nonsinusoidal Modulation

When a sinusoidal carrier signal is modulated by a non-sinusoidal modulating signal, the process is called Non-sinusoidal modulation. Consider a high frequency sinusoidal signal $c(t) = A_c\cos(2\pi f_c t)$ and the non-sinusoidal message signal *m(t)* as shown in figure 2.6. The non-sinusoidal modulating signal has a line spectrum that is many frequency components of different amplitudes.

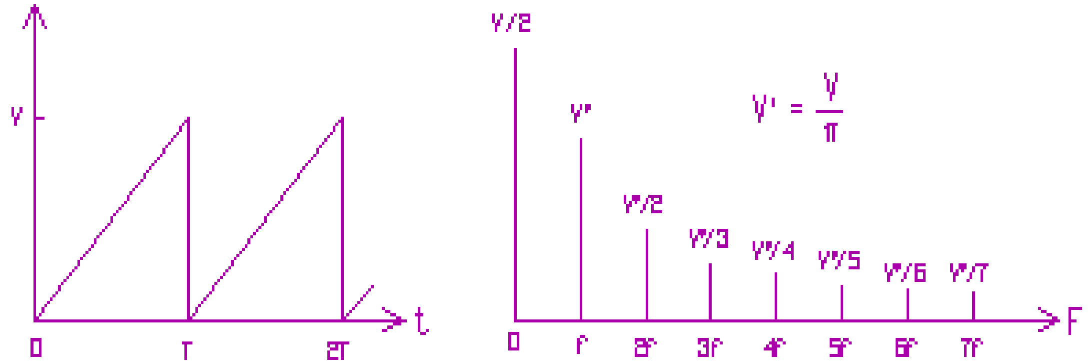


Figure 2.6: Non-sinusoidal message signal and spectrum

The expression for the non-sinusoidal AM is given by

$$S(t) = A_c[1 + k_a A_1 \cos(2\pi f_1 t) + k_a A_2 \cos(2\pi f_2 t) + ------]\cos(2\pi f_c t)$$

The total average power can be obtained by adding the average power for each component,

$$P_t = P_c\left(1 + \frac{m_1^2}{2} + \frac{m_2^2}{2} + -----\right)$$

Hence the effective modulation index can be defined as $m_{eff} = \sqrt{m_1^2 + m_2^2 + ---}$

Amplitude Modulators

Two basic amplitude modulation principles are discussed. They are square law modulation and switching modulation.

Square law modulator

When the output of a device is not directly proportional to input throughout the operation, the device is said to be non-linear. The Input-Output relation of a non-linear device can be expressed as

$$V_0 = a_0 + a_1 V_{in} + a_2 V_{in}^2 + a_3 V_{in}^3 + a_4 V_{in}^4 + ...$$

When the in put is very small, the higher power terms can be neglected. Hence the output is approximately given by $V_0 = a_O + a_1 V_{in} + a_2 V_{in}^2$

When the output is considered up to square of the in put, the device is called a square law device and the square law modulator is as shown in the figure 2.7.

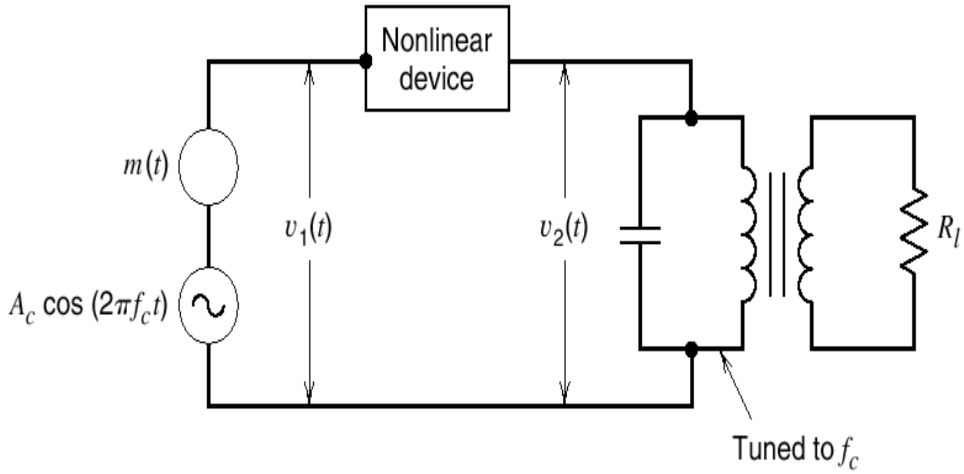


Figure 2.7: Square law modulator

Consider a non linear device to which a carrier *c(t)*= $A_c \cos(2\pi f_c t)$ and an information signal *m(t)* are fed simultaneously as shown in figure 2.7. The total input to the device at any instant is

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

As the level of the input is very small, the output can be considered up to square of the input, *i.e.*, $V_0 = a_O + a_1 V_{in} + a_2 V_{in}^2$

$$V_0 = a_0 + a_1[A_c \cos 2\pi f_c t + m(t)] + a_2[A_c \cos 2\pi f_c t + m(t)]^2$$

$$V_0 = a_0 + a_1 A_c \cos 2\pi f_c t + a_1 m(t) + \frac{a_2 A_c^2}{2}(1 + \cos 4\pi f_c t) + a_2[m(t)]^2 + 2a_2 m(t) A_c \cos 2\pi f_c t$$

$$V_0 = a_0 + a_1 A_c \cos 2\pi f_c t + a_1 m(t) + \frac{a_2 A_c^2}{2}\cos 4\pi f_c t + a_2 m^2(t) + 2a_2 m(t) A_c \cos 2\pi f_c t$$

Taking Fourier transform on both sides, we get

$$V_0(f) = (a_0 + \frac{a_2 A_c^2}{2})\delta(f) + \frac{a_1 A_c}{2}\left[\delta(f - f_c) + \delta(f + f_c)\right] + a_1 M(f) +$$
$$\frac{a_2 A_c^2}{4}\left[\delta(f - 2f_c) + \delta(f + 2f_c)\right] + a_2 M(f) + a_2 A_c\left[M(f - f_c) + M(f + f_c)\right]$$

Therefore the square law device output V_0 consists of

The dc component at *f = 0.*

The information signal ranging from 0 to W Hz and its second harmonics.

Signal at f_c and $2f_c$.

Frequency band centered at f_c with a deviation of $\pm W$, Hz.

The required AM signal with a carrier frequency f_c can be separated using a band pass filter at the out put of the square law device. The filter should have a lower cut-off frequency ranging between 2W and (f_c -W) and upper cut-off frequency between (f_c +W) and 2 f_c

Therefore the filter out put is

$$s(t) = a_1 A_c \cos 2\pi f_c t + 2a_2 A_c m(t) \cos 2\pi f_c t$$

$$s(t) = a_1 A_c \left[1 + 2\frac{a_2}{a_1} m(t) \right] \cos 2\pi f_c t$$

If $m(t) = A_m \cos 2\pi f_m t$, we get

$$s(t) = a_1 A_c \left[1 + 2\frac{a_2}{a_1} A_m \cos 2\pi f_m t \right] \cos 2\pi f_c t$$

Comparing this with the standard representation of AM signal,

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

Therefore modulation index of the output signal is given by

$$m = 2\frac{a_2}{a_1} A_m$$

The output AM signal is free from distortion and attenuation only when ($f_c - W$)> 2W or $f_c > 3W$.

Switching modulator

Consider a semiconductor diode used as an ideal switch to which the carrier signal $c(t) = A_c \cos(2\pi f_c t)$ and information signal $m(t)$ are applied simultaneously as shown figure 2.8.

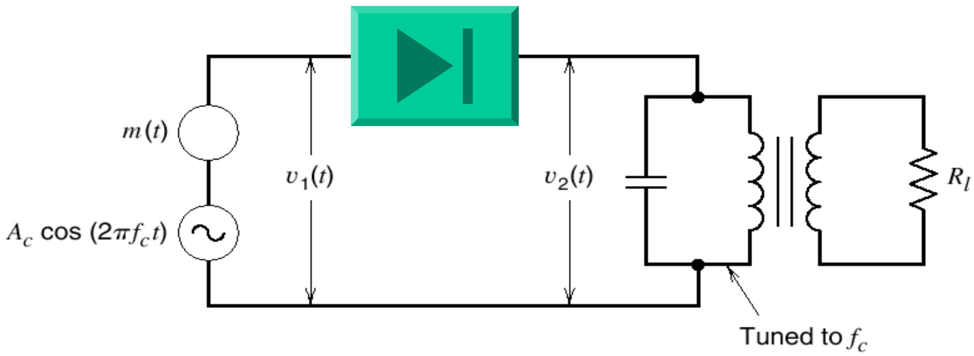


Figure 2.8: Switching modulator

The total input for the diode at any instant is given by

$$v_1 = c(t) + m(t)$$
$$v_1 = A_c \cos 2\pi f_c t + m(t)$$

When the peak amplitude of *c(t)* is maintained more than that of information signal, the operation is assumed to be dependent on only c(t) irrespective of *m(t)*. When *c(t)* is positive, $v_2=v_1$since the diode is forward biased. Similarly, when *c(t)* is negative, $v_2=0$ since diode is reverse biased. Based upon above operation, switching response of the diode is periodic rectangular wave with an amplitude unity and is given by

$$p(t) = \frac{1}{2} + \frac{1}{\pi}\sum_{n=-\infty}^{\infty} \frac{(-1)^{n-1}}{2n-1} \cos(2\pi f_c t(2n-1))$$

$$p(t) = \frac{1}{2} + \underset{n=0,1}{\frac{2}{\pi}} \cos(2\pi f_c t) - \underset{n=-1,+2}{\frac{2}{3\pi}} \cos(6\pi f_c t) + -$$

Therefore the diode response V_o is a product of switching response *p(t)* and input v_1.

$$v_2=v_1 * p(t)$$

$$V_2 = [A_c \cos 2\pi f_c t + m(t)]\left[\frac{1}{2} + \frac{2}{\pi}\cos 2\pi f_c t - \frac{2}{3\pi}\cos 6\pi f_c t + - + -\right]$$

Applying the Fourier Transform, we get

$$V_2(f) = \frac{A_c}{4}\left[\delta(f-f_c) + \delta(f+f_{c)}\right] + \frac{M(f)}{2} + \frac{A_c}{\pi}\delta(f)$$
$$+\frac{A_c}{2\pi}[\delta(f-2f_c) + \delta(f+2f_c)] + \frac{1}{\pi}[M(f-f_c) + M(f+f_c)]$$
$$-\frac{A_c}{6\pi}[\delta(f-4f_c) + \delta(f+4f_c)] - \frac{A_c}{3\pi}[\delta(f-2f_c) + \delta(f+2f_c)]$$
$$-\frac{1}{3\pi}[M(f-3f_c) + M(f+f_c)]$$

The diode output v_2 consists of

a dc component at $f=0$.

Information signal ranging from 0 to *w* Hz and infinite number of frequency bands centered at f, $2f_c$, $3f_c$, $4f_c$, ---------

The required AM signal centered at f_c can be separated using band pass filter. The lower cutoff-frequency for the band pass filter should be between w and f_c-w and the upper cut-off frequency between f_c+w and $2f_c$. The filter output is given by the equation

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

For a single tone information, let $m(t) = A_m \cos(2\pi f_m t)$

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

Therefore modulation index, $m = \frac{4}{\pi}\frac{A_m}{A_c}$

The output AM signal is free from distortions and attenuations only when fc-w>w or $fc>2w$.

Demodulation of AM: -

Demodulation is the process of recovering the information signal (base band) from the incoming modulated signal at the receiver. There are two methods.

Square law demodulator

Consider a non-linear device to which the AM signal s(t) is applied. When the level of s(t) is very small, output can be considered upto square of the input.

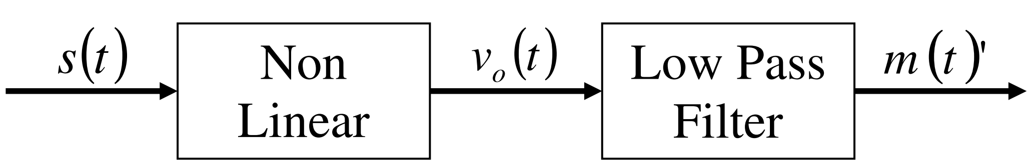


Figure: Demodulation of AM using square law device

Therefore $V_o = a_o + a_1 V_{in} + a_2 V_{in}^{\ 2}$

If $m(t)$ is the information signal (0-wHz) and $c(t) = A_c \cos(2\pi f_c t)$ is the carrier, input AM signal to the non-linear device is given by

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

$$V_o = a_o + a_1 s(t) + a_2 [s(t)]^2$$

$$V_o = a_o + a_1 A_c \cos 2\pi f_c t + a_1 A_c K_a m(t) \cos 2\pi f_c t + a_2 [A_c \cos 2\pi f_c t + A_c k_a m(t) \cos 2\pi f_c t]^2$$

Applying Fourier transform on both sides, we get

$$V_o(f) = \left[a_o + \frac{a_2 A_c^{\,2}}{2} \right] \delta(f) + \frac{a_1 A_c}{2} [\delta(f - f_c) + \delta(f + f_c)]$$

$$+ \frac{a_1 A_c K_a}{2} [M(f - f_c) + M(f + f_c)] + \frac{a_2 A_c^{\,2} K_a^{\,2}}{4} [M(f - 2f_c) + M(f + 2f_c)]$$

$$+ \frac{a_2 A_c^{\,2} K_a^{\,2}}{2} \left[\underset{\pm 2W}{M(f)} \right] + \frac{a_2 A_c^{\,2} K_a^{\,2}}{2} [M(f - 2f_c) + M(f + 2f_c)]$$

$$+ \frac{a_2 A_c^{\,2}}{4} [\delta(f - 2f_c) + \delta(f + 2f_c)] + a_2 A_c^{\,2} K_a [M(f)]$$

The device output consists of a dc component at $f = 0$, information signal ranging from 0-W Hz and its second harmonics and frequency bands centered at f_c and $2f_c$.

The required information can be separated using low pass filter with cut off frequency ranging between W and f_c-w.

The filter output is given by

$$m'(t) = \left(a_o + \frac{a_2 A_c^{\,2}}{2} \right) + a_2 A_c^{\,2} K_a m(t) + \frac{a_2 A_c^{\,2} K_a^{\,2} m^2(t)}{2}$$

DC component + message signal + second harmonic

The dc component (first term) can be eliminated using a coupling capacitor or a transformer. The effect of second harmonics of information signal can be reduced by maintaining its level very low. When $m(t)$ is very low, the filter output is given by

$$m^1(t) = a_2 A_c^{\,2} K_a m(t)$$

When the information level is very low, the noise effect increases at the receiver, hence the system clarity is very low using square law demodulator.

Envelop detector

It is a simple and highly effective system. This method is used in most of the commercial AM radio receivers. An envelop detector is as shown below.

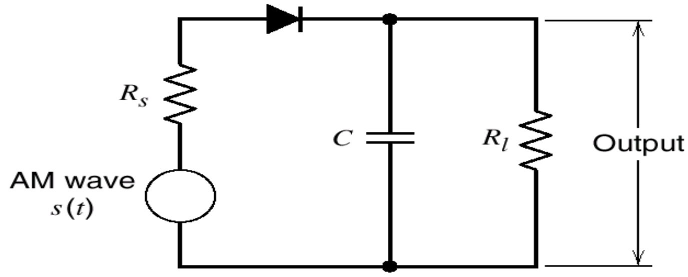


Figure: Envelope detector

During the positive half cycles of the input signals, the diode D is forward biased and the capacitor C charges up rapidly to the peak of the input signal. When the input signal falls below this value, the diode becomes reverse biased and the capacitor C discharges through the load resistor R_L.

The discharge process continues until the next positive half cycle. When the input signal becomes greater than the voltage across the capacitor, the diode conducts again and the process is repeated.

The charge time constant $(r_f+R_s)C$ must be short compared with the carrier period, the capacitor charges rapidly and there by follows the applied voltage up to the positive peak when the diode is conducting.

That is the charging time constant shall satisfy the condition,

$$(r_f + R_s)C << \frac{1}{f_c}$$

On the other hand, the discharging time-constant R_LC must be long enough to ensure that the capacitor discharges slowly through the load resistor R_L between the positive peaks of the carrier wave, but not so long that the capacitor voltage will not discharge at the maximum rate of change of the modulating wave.

That is the discharge time constant shall satisfy the condition,

$$\frac{1}{f_c} << R_L C << \frac{1}{W}$$

where 'W' is band width of the message signal.

The result is that the capacitor voltage or detector output is nearly the same as the envelope of AM wave.

Advantages of AM: Generation and demodulation of AM wave are easy. AM systems are cost effective and easy to build.

Disadvantages: AM contains unwanted carrier component, hence it requires more transmission power. The transmission bandwidth is equal to twice the message bandwidth.

To overcome these limitations, the conventional AM system is modified at the cost of increased system complexity. Therefore, three types of modified AM systems are discussed.

DSBSC (Double Side Band Suppressed Carrier) modulation:

In DSBC modulation, the modulated wave consists of only the upper and lower side bands. Transmitted power is saved through the suppression of the carrier wave, but the channel bandwidth requirement is the same as before.

SSBSC (Single Side Band Suppressed Carrier) modulation:

The SSBSC modulated wave consists of only the upper side band or lower side band. SSBSC is suited for transmission of voice signals. It is an optimum form of modulation in that it requires the minimum transmission power and minimum channel band width. Disadvantage is increased cost and complexity.

VSB (Vestigial Side Band) modulation:

In VSB, one side band is completely passed and just a trace or vestige of the other side band is retained. The required channel bandwidth is therefore in excess of the message bandwidth by an amount equal to the width of the vestigial side band. This method is suitable for the transmission of wide band signals.

Double Side Band Suppressed Carrier Modulation

DSBSC modulators make use of the multiplying action in which the modulating signal multiplies the carrier wave. In this system, the carrier component is eliminated and both upper and lower side bands are transmitted. As the carrier component is suppressed, the power required for transmission is less than that of AM.

If $m(t)$ is the message signal and $c(t) = A_c \cos(2\pi f_c t)$ is the carrier signal, then DSBSC modulated wave $s(t)$ is given by

$$s(t) = c(t)\, m(t)$$

$$s(t) = A_c \cos(2\pi f_c t) m(t)$$

Consequently, the modulated signal *s(t)* under goes a phase reversal , whenever the message signal *m(t)* crosses zero as shown below.

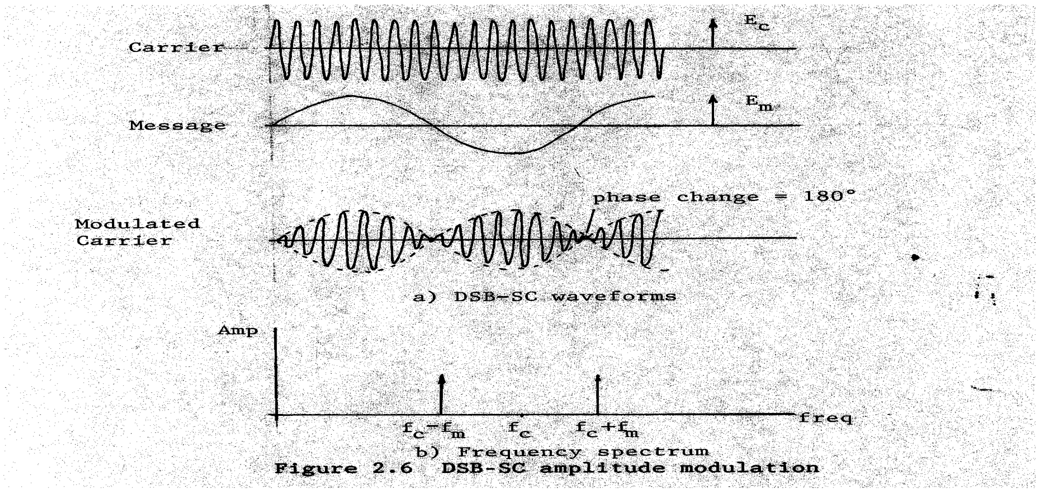


Figure: Carrier, message and DSBSC wave forms

The envelope of a DSBSC modulated signal is therefore different from the message signal and the Fourier transform of *s(t)* is given by

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

For the case when base band signal *m(t)* is limited to the interval $-W<f<W$ as shown in figure below, we find that the spectrum *S(f)* of the DSBSC wave *s(t)* is as illustrated below. Except for a change in scaling factor, the modulation process simply translates the spectrum of the base band signal by f_c. The transmission bandwidth required by DSBSC modulation is the same as that for AM.

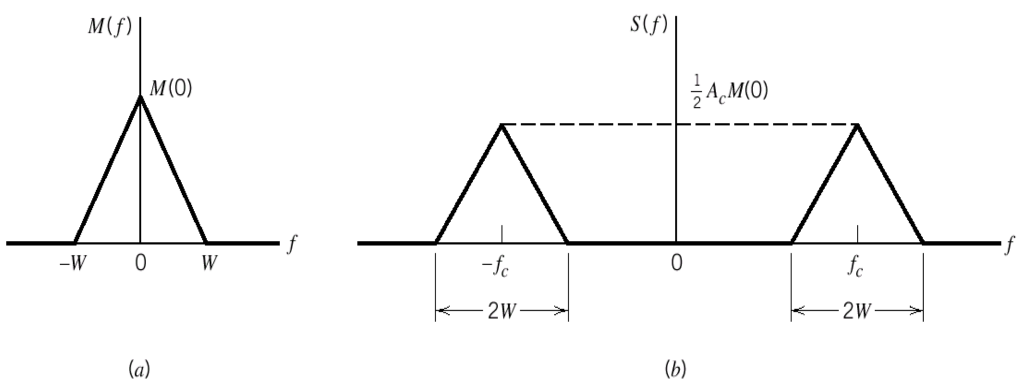


Figure: Message and the corresponding DSBSC spectrum

Ring modulator: -

Ring modulator is the most widely used product modulator for generating DSBSC wave and is shown below.

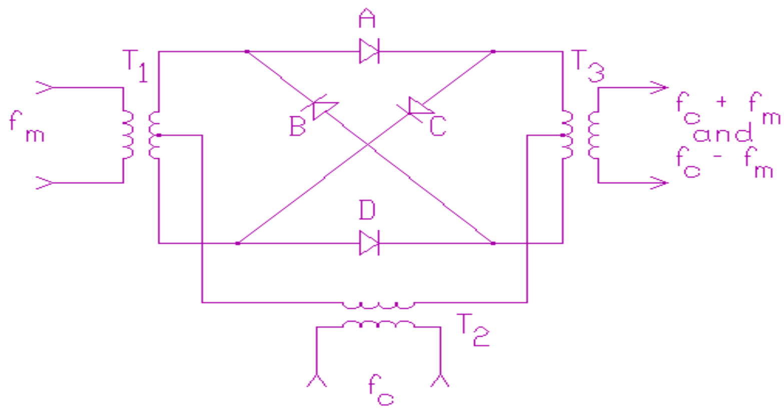


Figure: Ring modulator

The four diodes form a ring in which they all point in the same direction. The diodes are controlled by square wave carrier *c(t)* of frequency f_c, which is applied longitudinally by means of two center-tapped transformers. Assuming the diodes are ideal, when the carrier is positive, the outer diodes D1 and D2 are forward biased where as the inner diodes D3 and D4 are reverse biased, so that the modulator multiplies the base band signal *m(t)* by *c(t)*. When the carrier is negative, the diodes D1 and D2 are reverse biased and D3 and D4 are forward, and the modulator multiplies the base band signal –*m(t)* by *c(t)*. Thus the ring modulator in its ideal form is a product modulator for square wave carrier and the base band signal *m(t)*. The square wave carrier can be expanded using Fourier series as

$$c(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}\cos(2\pi f_c t(2n-1))$$

Therefore the ring modulator out put is given by

$$s(t) = m(t)c(t)$$

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

From the above equation it is clear that output from the modulator consists entirely of modulation products. If the message signal *m(t)* is band limited to the frequency band $-w < f < w$, the output spectrum consists of side bands centered at f_c.

Balance modulator (Product modulator)

A balanced modulator consists of two standard amplitude modulators arranged in a balanced configuration so as to suppress the carrier wave as shown in the following block diagram. It is assumed that the AM modulators are identical, except for the sign reversal of the modulating wave applied to the input of one of them. Thus, the output of the two modulators may be expressed as,

$$s_1(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t)$$
$$s_2(t) = A_c[1 - k_a m(t)]\cos(2\pi f_c t)$$

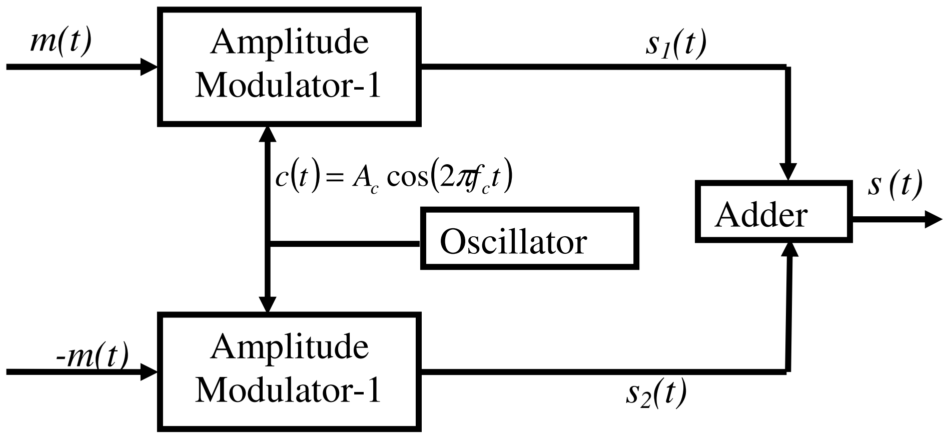


Figure: Balanced modulator

Subtracting $s_2(t)$ from $s_1(t)$,

$$s(t) = s_1(t) - s_2(t)$$
$$s(t) = 2k_a m(t) A_c \cos(2\pi f_c t)$$

Hence, except for the scaling factor $2k_a$, the balanced modulator output is equal to the product of the modulating wave and the carrier.

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.

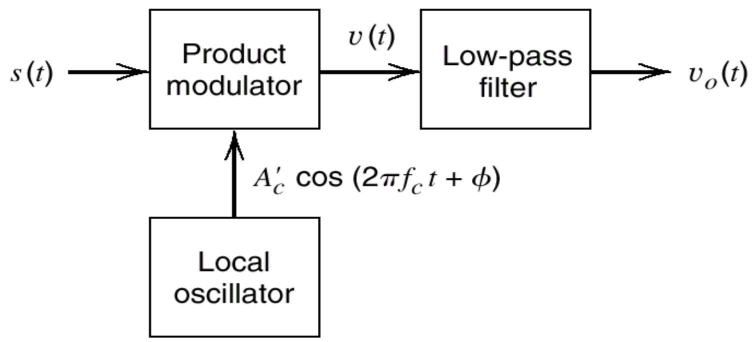


Figure: Coherent detector

It is assumed that the local oscillator signal is exactly coherent or synchronized, in both frequency and phase, with the 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.

Let $A_c^{\ 1}\cos(2\pi f_c t + \phi)$ be the local oscillator signal, and $s(t) = A_c \cos(2\pi f_c t)m(t)$ be the DSBSC wave. Then the product modulator output *v(t)* is given by

$$v(t) = A_c A_c^{\ 1} \cos(2\pi f_c t)\cos(2\pi f_c t + \phi)m(t)$$

$$v(t) = \frac{A_c A_c^{\ 1}}{4}\cos(4\pi f_c t + \phi)m(t) + \frac{A_c A_c^{\ 1}}{2}\cos(\phi)m(t)$$

The first term in the above expression 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 interval $-w < f < w$, the spectrum of *v(t)* is plotted as shown below.

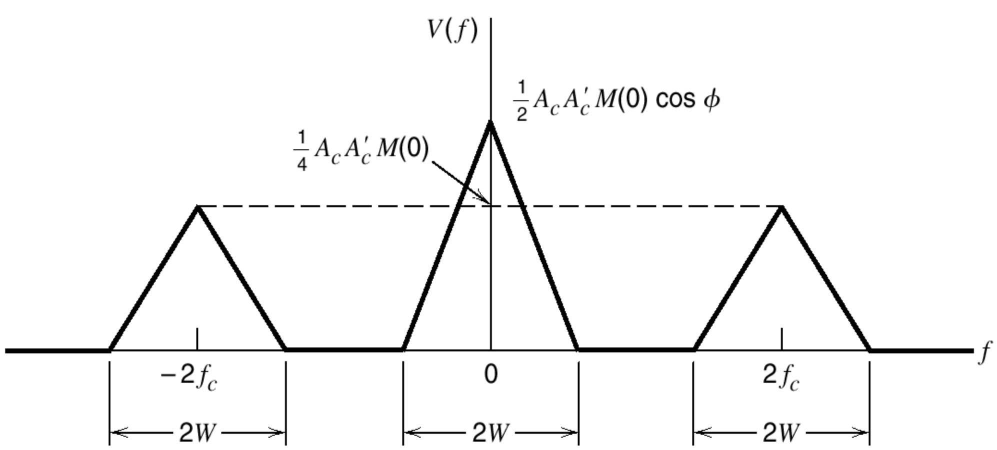


Figure: Spectrum of output of the product modulator

From the spectrum, it is clear that the unwanted component (first term in the expression) can be removed by the low-pass filter, provided that the cut-off frequency of the filter is greater than W but less than $2f_c$-W. The filter output is given by

$$v_o(t) = \frac{A_c A_c^1}{2}\cos(\phi)m(t)$$

The demodulated signal $v_o(t)$ is therefore proportional to $m(t)$ when the phase error ϕ is constant.

Single tone DSBSC modulation

Consider a sinusoidal modulating signal $m(t) = A_m \cos(2\pi f_m t)$ of single frequency and the carrier signal $c(t) = A_c \cos(2\pi f_c t)$.

The corresponding DSBSC modulated wave is given by

$$s(t) = A_c A_m \cos(2\pi f_c t)\cos(2\pi f_m t)$$

$$s(t) = \frac{1}{2}A_c A_m \cos(2\pi(f_c - f_m)t) + \frac{1}{2}A_c A_m \cos(2\pi(fc + f_m)t)$$

Thus the spectrum of the DSBSC modulated wave, for the case of sinusoidal modulating wave, consists of delta functions located at $-f_c \pm f_m$ and $f_c \pm f_m$.

Assuming perfect synchronism between the local oscillator and carrier wave in a coherent detector, the product modulator output contains the high frequency components and scaled version of original information signal. The Low Pass Filter is used to separate the desired message signal.

Costas Receiver (Costas loop)

Costas receiver is a synchronous receiver system, suitable for demodulating DSBSC waves. It consists of two coherent detectors supplied with the same input signal, that is the incoming DSBSC wave $s(t) = A_c \cos(2\pi f_c t)m(t)$ but with individual local oscillator signals that are in phase quadrature with respect to each other as shown below.

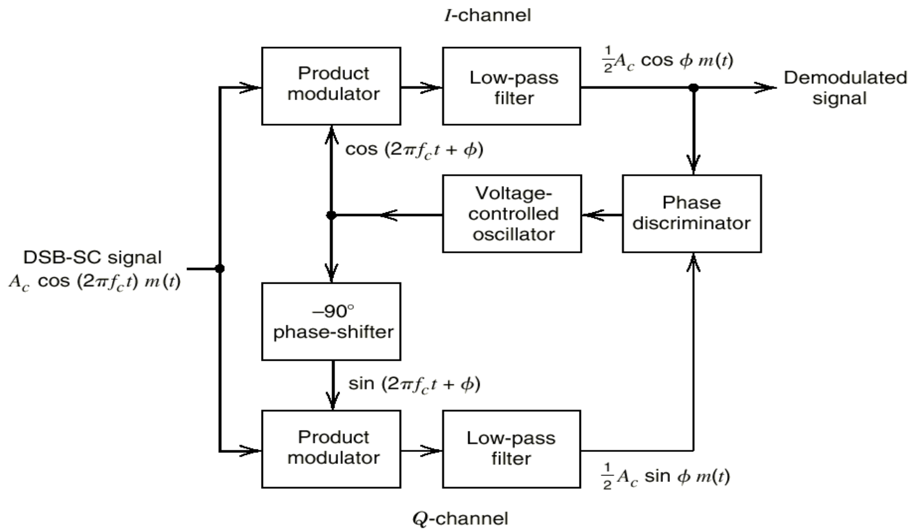


Figure: Costas receiver

The frequency of the local oscillator is adjusted to be the same as the carrier frequency f_c. The detector in the upper path is referred to as the in-phase coherent detector or I-channel, and that in the lower path is referred to as the quadrature-phase coherent detector or Q-channel. These two detector are coupled together to form a negative feed back system designed in such a way as to maintain the local oscillator synchronous with the carrier wave. Suppose the local oscillator signal is of the same phase as the carrier wave $c(t) = A_c \cos(2\pi f_c t)$ used to generate the incoming DSBSC wave. Then we find that the I-channel output contains the desired demodulated signal *m(t)*, where as the Q-channel output is zero due to quadrature null effect of the Q-channel. Suppose that the local oscillator phase drifts from its proper value by a small angle ϕ radiations. The I-channel output will remain essentially unchanged, but there will be some signal appearing at the Q-channel output, which is proportional to $\sin(\phi) \cong \phi$ for small ϕ. This Q-channel output will have same polarity as the I-channel output for one direction of local oscillator phase drift and opposite polarity for the opposite direction of local oscillator phase drift. Thus by combining the I-channel and Q-channel outputs in a phase discriminator (which consists of a multiplier followed by a LPF), a dc control signal is obtained that automatically corrects for the local phase errors in the voltage-controlled oscillator.