1. Demodulation of AM

HTE - 21.01.2013

In general demodulation is established by frequency shifting of the modulated signal. In mathematical formulation, we use the following notation

a) Modulating (message) signal

$$m(t) = a\cos(2\pi f_m t) \quad , \quad \text{in practice } m(t) = \sum_i a_i \cos(2\pi f_i t + \phi_i) \qquad (1.1)$$

We will usually adopt the first expression for $m(t)$, since whatever happens to $m(t)$ will also be applicable to individual cosine terms of the summation on the right hand side.

b) Carrier signal

$$c(t) = A_c \cos(2\pi f_c t) \qquad (1.2)$$

The carrier is mostly a single sinusoidal signal except in multi carrier cases like orthogonal frequency division multiplexing. In most cases, $f_c \gg f_m$

c) Modulated signal

$$u(t) = m(t)c(t) \qquad \text{in DSB(SC) - AM} \qquad (1.3)$$

d) Received signal

$$r(t) = u(t) + n(t) \qquad \text{assuming AWGN channel,} \quad n(t)\text{: Noise signal}$$
$$r(t) = u(t) \qquad \text{if noise is to be ignored} \qquad (1.4)$$

The simplest form of modulation is multiplying modulating signal $m(t)$ directly by the carrier $c(t)$ as shown in (1.3). By setting the modulating signal to a single sinusoid, with this operation, we get

$$\begin{aligned} u(t) &= m(t)c(t) = aA_c \cos(2\pi f_m t)\cos(2\pi f_c t) \\ &= \frac{aA_c}{2}\left\{\cos\left[2\pi(f_c - f_m)t\right] + \cos\left[2\pi(f_c + f_m)t\right]\right\} \end{aligned} \qquad (1.5)$$

Corresponding frequency domain expression is

$$\begin{aligned} U(f) = F\left[u(t)\right] = \frac{aA_c}{4}\big[&\delta(f - f_c + f_m) + \delta(f + f_c - f_m) \\ &+ \delta(f - f_c - f_m) + \delta(f + f_c + f_m)\big] \end{aligned} \qquad (1.6)$$

The above is valid for double side band suppressed carrier, DSB(SC). Now if we add a normalized DC level of unity to the message signal, we get full amplitude modulation (AM), thus

$$u(t) = [1 + m(t)]c(t) = A_c[1 + a\cos(2\pi f_m t)]\cos(2\pi f_c t)$$

$$= \underbrace{A_c\cos(2\pi f_c t)}_{\text{carrier}} + \frac{aA_c}{2}\{\underbrace{\cos[2\pi(f_c - f_m)t]}_{\text{lower sideband}} + \underbrace{\cos[2\pi(f_c + f_m)t]}_{\text{upper sideband}}\} \qquad (1.7)$$

So the corresponding time waveforms and spectrums for (1.6) and (1.7) will be

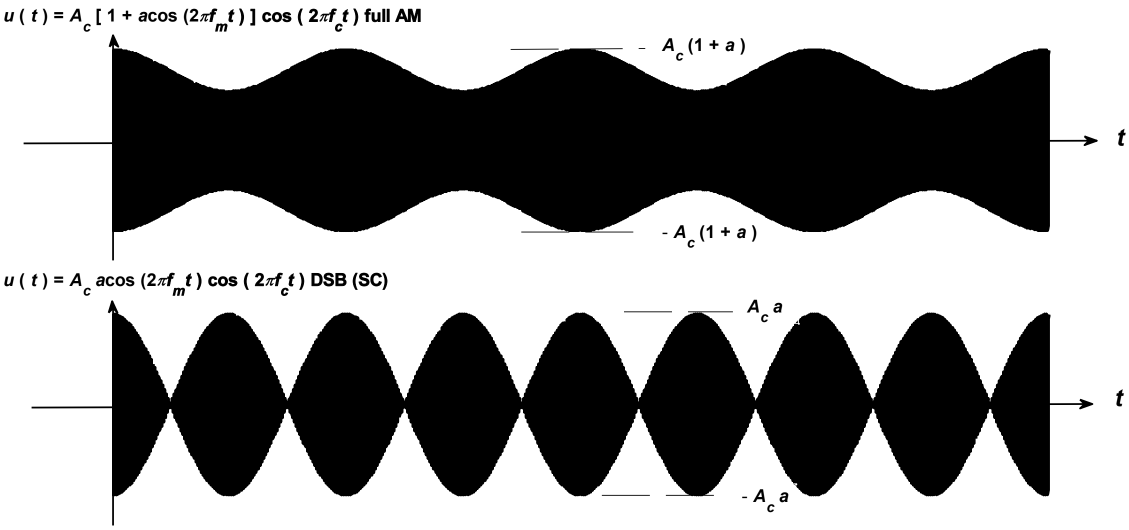

Fig. 1.1a Time waveforms of full AM and DSB(SC).

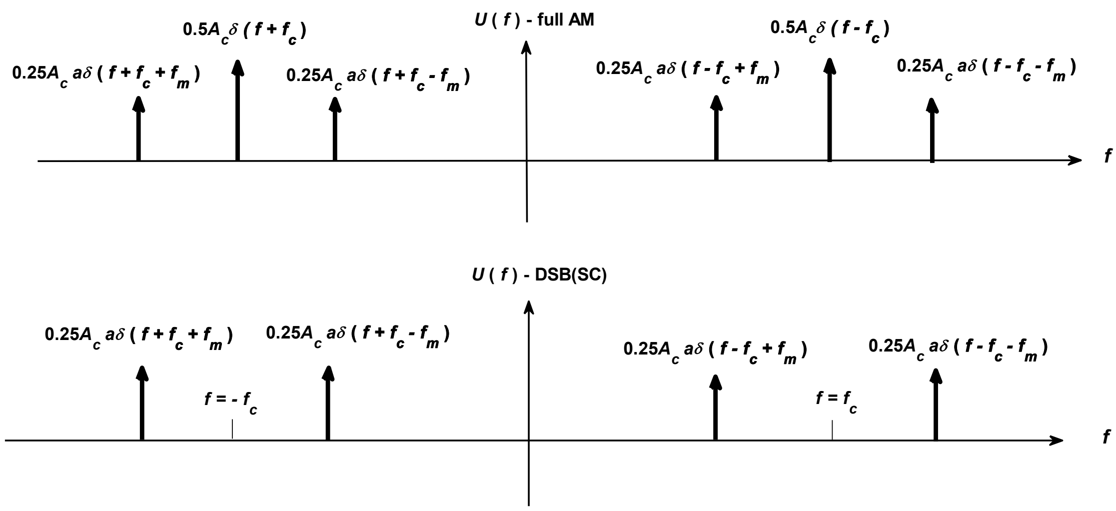

Fig. 1.1b Frequency spectrums of full AM and DSB(SC).

Fig. 1.1 Time waveforms and frequency spectrums for DSB(SC) and full AM.

Demodulation of AM : The simplest form of demodulation in AM is to multiply the received signal by the same carrier generated at the receiver. Since the transmitter and receiver are far apart (as shown below), the carrier generated at the receiver will have a phase difference, thus

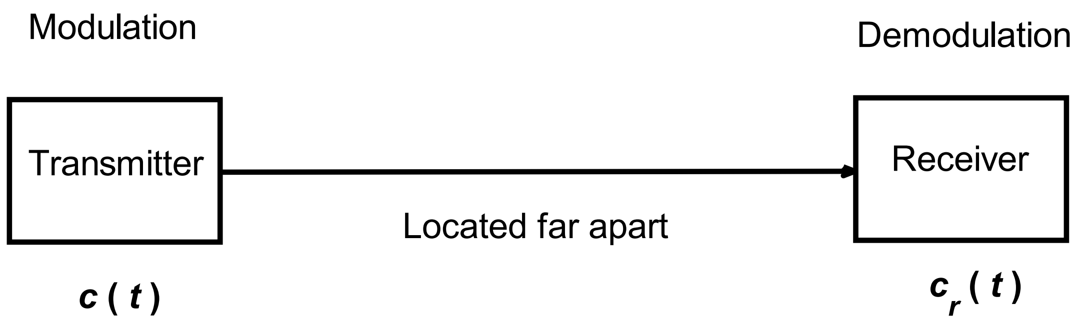

Fig. 1.2 Block diagram of transmitter and receiver locations.

$$c_r(t) = A_c \cos(2\pi f_c t + \phi) \qquad (1.8)$$

After multiplying the received signal by (1.8), for DSB(SC) we get (excluding noise)

$$\begin{aligned} r_r(t) &= A_c m(t)\cos(2\pi f_c t)\cos(2\pi f_c t + \phi) \\ &= \frac{A_c}{2} m(t)\cos(\phi) + \frac{A_c}{2} m(t)\cos(4\pi f_c t + \phi) \end{aligned} \qquad (1.10)$$

In terms of frequency spectrum, (1.10) will look like

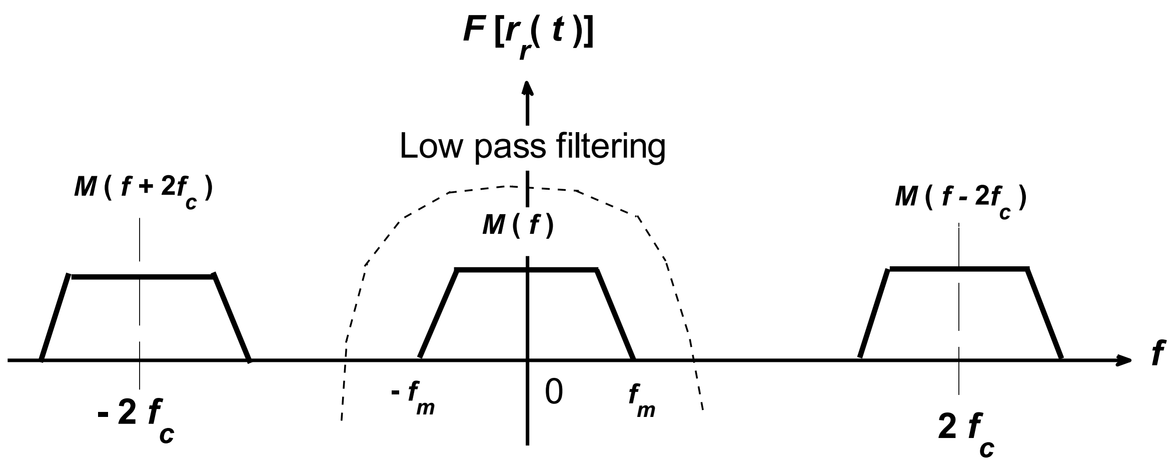

Fig. 1.3 Frequency spectrum of the time signal in (1.10)

From Fig. 1.3, we see that the recovery of the message signal $m(t)$ can be achieved by low pass filtering the first term on the second line of (1.10), after this operation, we will have

$$m_r(t) = \frac{A_c}{2} m(t)\cos(\phi) \qquad (1.11)$$

As seen (1.11) contains the message signal $m(t)$ multiplied by $\cos(\phi)$. Hence the message signal $m(t)$ can be recovered from (1.11) provided that $\phi \approx 0$. At other values of ϕ, the amplitude of (1.11) will be reduced. In particular, if $\phi = \pi/2$, then it will be impossible to recover the message signal $m(t)$ from (1.11). So we conclude that for demodulation to be successful, some synchronisation mechanism is required between the carrier used at transmitter (to perform modulation) and the carrier used at receiver (to perform demodulation). One way of achieving this is to add a small amount of carrier to the DSB signal before transmission. Then by some phase locked loop setup, it will be to phase synchronize the carrier generated locally at the receiver with the one sent from the transmitter. The effects of variations in ϕ can also be illustrated by the following MATLAB code, AMModDemod_Exp1.m. As can be seen from this MATLAB file, a setting of $\phi = \pi/2$ will make the demodulation impossible.

Exercise 1.1 : By using the above given formulation, predict if demodulation can be successful at $\phi = -\pi$. In general for a range of $0 \rightarrow 2\pi$, find which values of ϕ make the recovery of the message signal difficult or impossible and which values of ϕ are harmless for the demodulation.

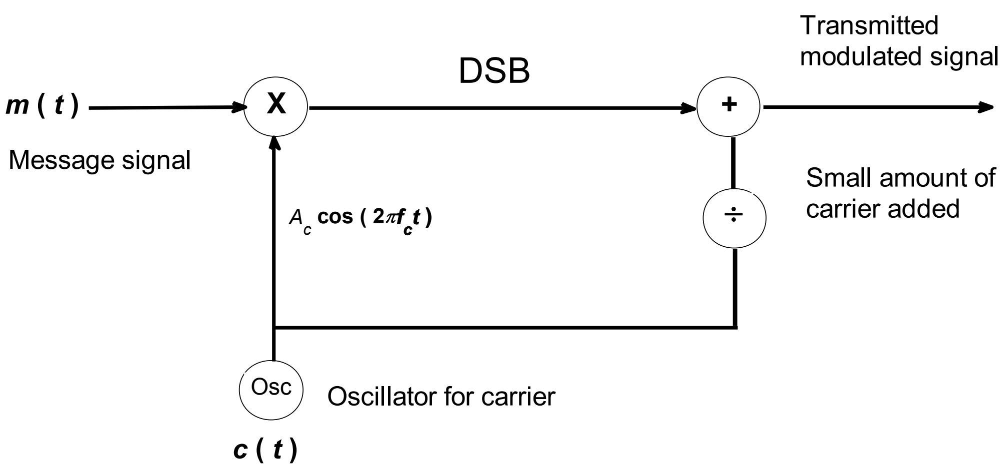

Fig. 1.4 Adding a small amount of carrier to DSB signal before transmission.

The above is called phase coherent or synchronous demodulation which is essential for the demodulation of DSB signal. In the case of full AM however, there is an alternative method called envelope detection (incoherent detection). This is based on the idea that in full AM, the envelope of the modulated signal follows (traces) the message signal $m(t)$ as shown below.

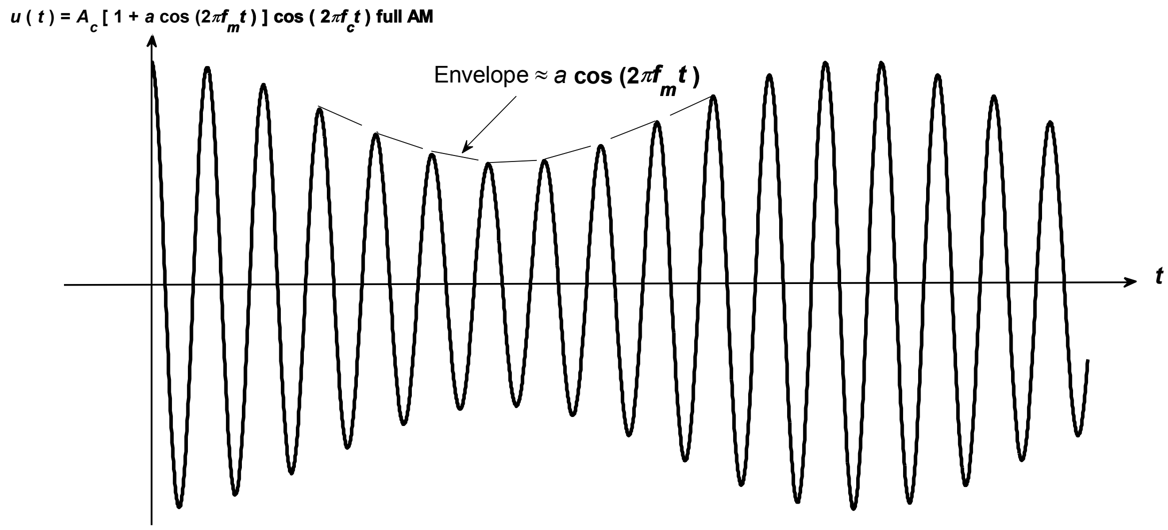

Fig. 1.5 Illustration of the envelope in full AM.

Mathematically, envelope detection functions as follows

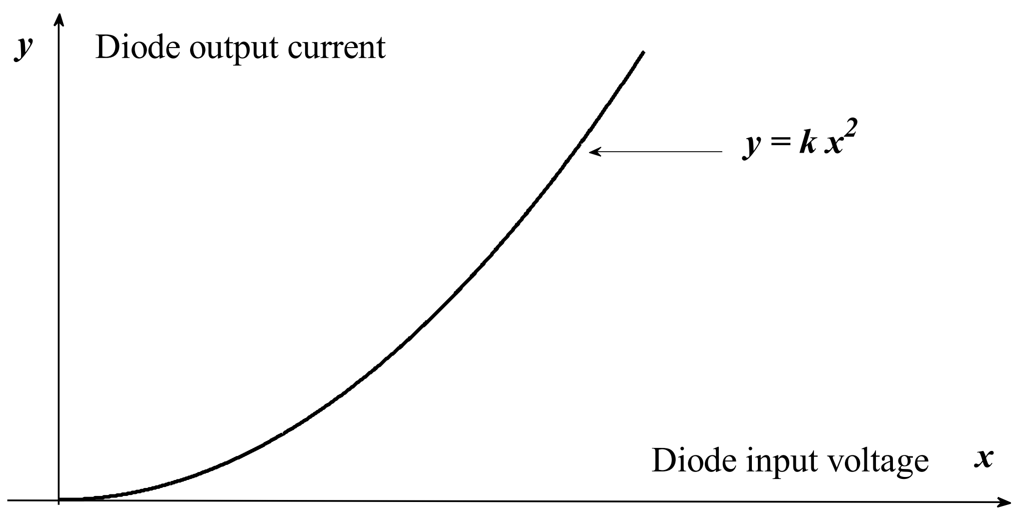

Fig. 1.6 Input and output characteristics of a diode (Germanium).

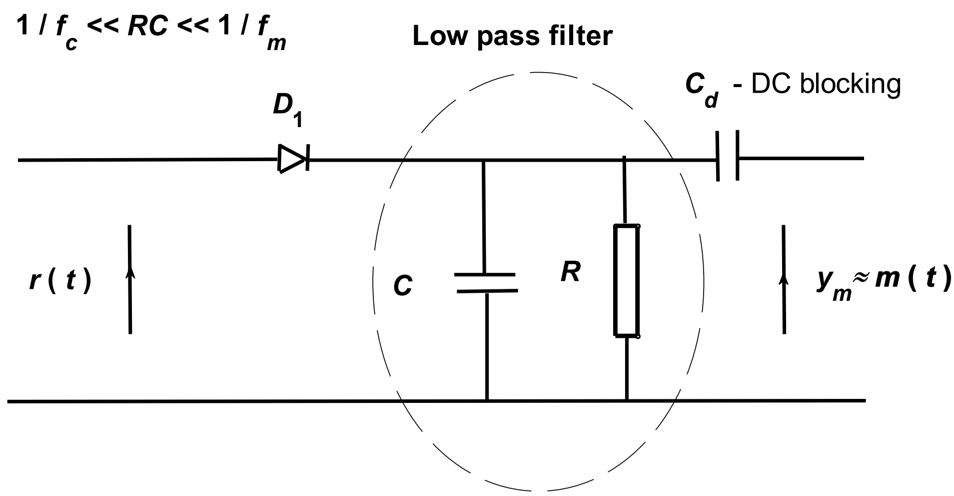

Fig. 1.7 Circuit diagram of envelope detector.

If we identify, x with $r(t)$, then the output (from the diode) will be

$$y = k\left\{A_c\left[1+m(t)\right]\cos\left(2\pi f_c t\right)\right\}^2 \qquad (1.12)$$

The squaring action in (1.12) will produce sinusoidal components around $2f_c$ which will be filtered by the RC filter in Fig. 1.5. The remaining terms will then be

$$y_m = kA_c^2\left[m(t)+\frac{1}{2}m^2(t)\right] \qquad (1.13)$$

We see here $m(t)$ as well as $m^2(t)$ are generated. It is easy to get the message signal $m(t)$ from (1.13) so long as $m(t) \gg \frac{1}{2}m^2(t)$ or simply $m(t) \ll 1$. In this case (1.13) will approximate to

$$y_m \simeq kA_c^2 m(t) \qquad (1.14)$$

Referring to Fig. 1.7, we can get another operational view of envelope detector, which is particularly applicable to silicon diode. Here we envisage that, the diode D_1 conducts only the positive cycles of the modulated signal, thus act as a half wave rectifier, then the obtained envelope is low pass filtered to get the message signal as shown below in Fig. 1.8.

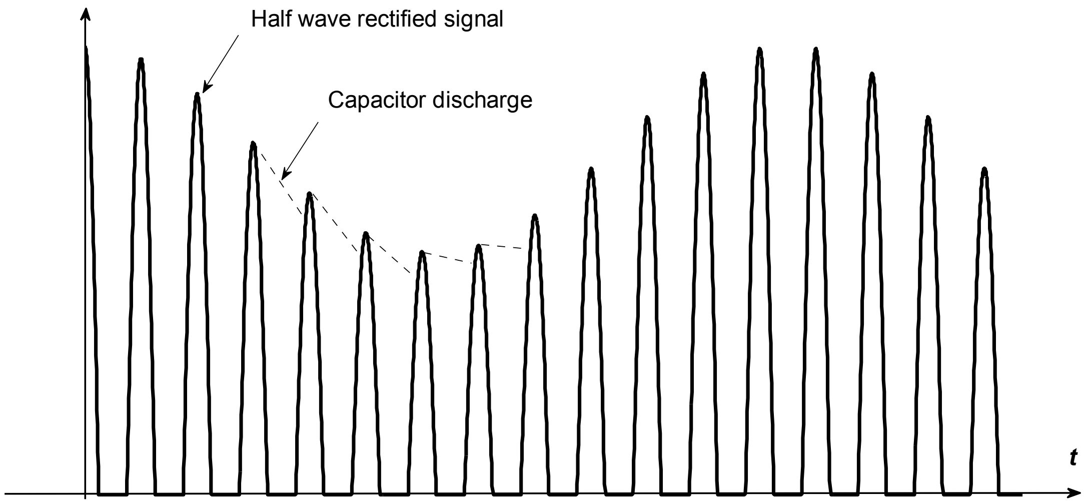

Fig. 1.8 Another operational view of envelope detector.

2. Generation and Demodulation of PM and FM

PM and FM are angle modulations. In this case, the modulated signal can be represented as

$$u(t) = A_c \cos\left[\theta(t)\right] \qquad (2.1)$$

Here the angle, $\theta(t)$ is the quantity to be modulated by the message signal. The (instantaneous) frequency of $\theta(t)$ can be retrieved from

$$f_i(t) = \frac{1}{2\pi}\frac{d}{dt}\theta(t) \qquad (2.2)$$

Assume that $\theta(t)$ consists of two parts such that

$$\theta(t) = 2\pi f_c t + \phi(t) \qquad (2.3)$$

With this arrangement, (2.1) and (2.2) will become

$$u(t) = A_c \cos\left[2\pi f_c t + \phi(t)\right] \quad , \quad f_i(t) = f_c + \frac{1}{2\pi}\frac{d}{dt}\phi(t) \qquad (2.4)$$

where $f_i(t)$ is known as instantaneous frequency. Now it is possible to arrive at PM or FM, depending on how $\phi(t)$ is related to the message signal

$$\phi(t) = \begin{cases} k_p m(t) & \text{PM} \quad , \quad k_p = \text{ Phase deviation related parameter} \\ 2\pi k_f \int_{-\infty}^{t} m(\tau)\,d\tau & \text{FM} \quad , \quad k_f = \text{ Frequency deviation related parameter} \end{cases} \qquad (2.5)$$

(2.5) can alternatively be written as

$$\frac{d}{dt}\phi(t)=\begin{cases} k_p \dfrac{d}{dt}m(t) & \text{PM} \\ 2\pi k_f m(t) & \text{FM} \end{cases} \qquad (2.6)$$

Now we define modulation indices (an indication of the depth of modulation) for PM and FM as

$$\beta_p = k_p a \quad , \quad \beta_p = k_p \max\left[\left|m(t)\right|\right] \qquad \text{PM}$$
$$\beta_f = \frac{k_f a}{f_m} \quad , \quad \beta_f = \frac{k_f \max\left[\left|m(t)\right|\right]}{f_m} \qquad \text{FM} \qquad (2.7)$$

where f_m corresponds to the highest frequency in $m(t)$. The first definition on the two line refer to the message signal being a single sinusoid, while the second expressions are valid for the case of the modulating signal, $m(t)$ being in the form of a summation as given on the right hand side of (1.1). The complete PM and FM expressions and the instantaneous frequencies for a general $m(t)$ then become

$$u_{PM}(t)=A_c\cos\left[2\pi f_c t + k_p m(t)\right] \quad , \quad f_{iPM}(t)=f_c+\frac{k_f}{2\pi}\frac{d}{dt}m(t) \qquad \text{PM}$$
$$u_{FM}(t)=A_c\cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\tau)d\tau\right] \quad , \quad f_{iFM}(t)=f_c+k_f m(t) \qquad \text{FM} \qquad (2.8)$$

In order to highlight the implications of (2.6) and (2.7), we show the following waveforms (copied directly from Proakis 2002)

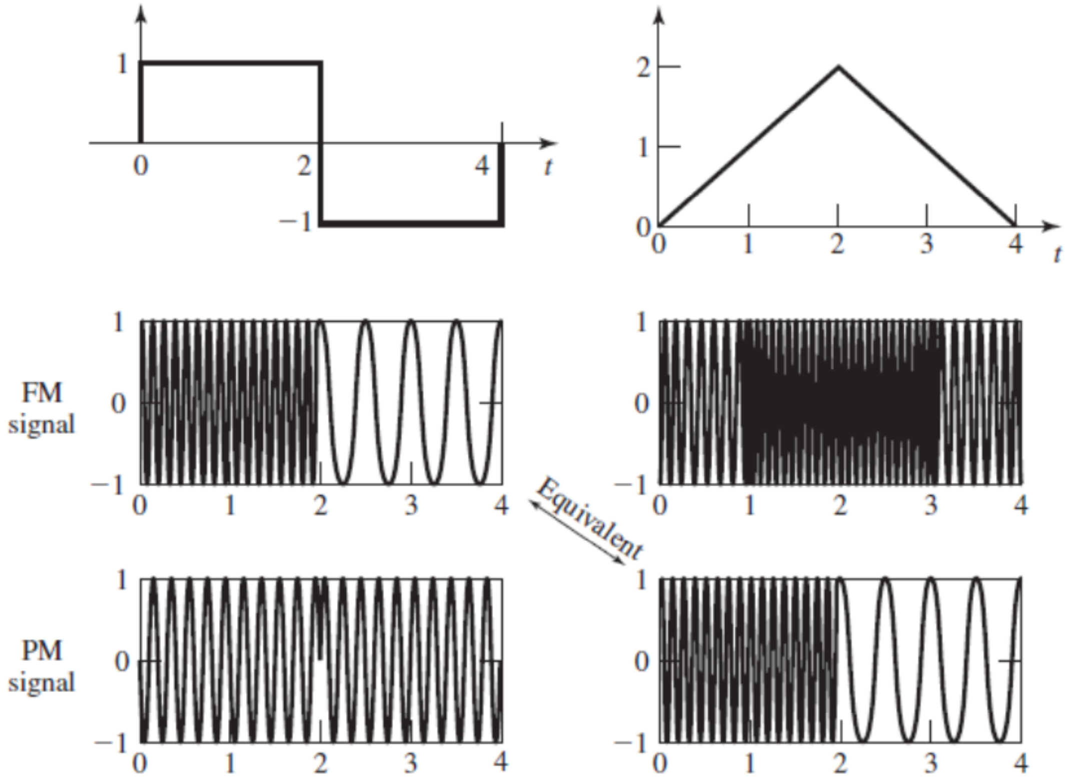

Fig. 2.1 FM and PM waveforms for square and triangular modulating signals.

By using (2.4), (2.6) and (2.8), we obtain the following instantaneous frequency expressions for the FM and PM signals of Fig. 2.1.

For FM signal modulated with $m_1(t)$ − Square waveform

$$f_i(t) = \begin{cases} f_c + \frac{1}{2\pi}\frac{d}{dt}\phi(t) = f_c + k_f m_1(t) = f_c + k_f = \text{constant} = f_{high} & 0 \le t \le 2 \\ f_c + \frac{1}{2\pi}\frac{d}{dt}\phi(t) = f_c + k_f m_1(t) = f_c - k_f = \text{constant} = f_{low} & 2 \le t \le 4 \end{cases}$$

For FM signal modulated with $m_2(t)$ − Triangular waveform

$$f_i(t) = \begin{cases} f_c + \frac{1}{2\pi}\frac{d}{dt}\phi(t) = f_c + k_f m_2(t) = f_c + k_f t = \text{increasing with time} & 0 \le t \le 2 \\ f_c + \frac{1}{2\pi}\frac{d}{dt}\phi(t) = f_c + k_f m_2(t) = f_c - k_f t = \text{decreasing with time} & 2 \le t \le 4 \end{cases} \quad (2.9)$$

For PM signal modulated with $m_1(t)$ − Square waveform

$$f_i(t)=\begin{cases} f_c+\dfrac{1}{2\pi}\dfrac{d}{dt}\phi(t)=f_c+\dfrac{k_p}{2\pi}\dfrac{d}{dt}m_1(t)=f_c=\text{constant} & 0\le t\le 2 \\ f_c+\dfrac{1}{2\pi}\dfrac{d}{dt}\phi(t)=f_c+\dfrac{k_p}{2\pi}\dfrac{d}{dt}m_1(t)=f_c=\text{constant} & 2\le t\le 4 \\ \text{Discontiniuty at } t=0, \text{ which makes} \quad \phi(t=2-)=\phi(t=2+) & \end{cases}$$

For PM signal modulated with $m_2(t)$ − Triangular waveform

$$f_i(t)=\begin{cases} f_c+\dfrac{1}{2\pi}\dfrac{d}{dt}\phi(t)=f_c+\dfrac{k_p}{2\pi}\dfrac{d}{dt}m_2(t)=f_c+\dfrac{k_p}{2\pi}=\text{constant}=f_{high} & 0\le t\le 2 \\ f_c+\dfrac{1}{2\pi}\dfrac{d}{dt}\phi(t)=f_c+\dfrac{k_p}{2\pi}\dfrac{d}{dt}m_1(t)=f_c-\dfrac{k_p}{2\pi}=\text{constant}=f_{low} & 2\le t\le 4 \end{cases} \qquad (2.10)$$

As understood from (2.9) and (2.10), provided that $k_f=\dfrac{k_p}{2\pi}$, then we will obtain equivalent waveforms, for FM signal modulated with $m_1(t)$ and the PM signal modulated with $m_2(t)$. From this point onwards, we concentrate on FM and set $\beta_f=\beta$.

2.1 Spectral Components of FM

For a single sinusoidal message signal

$$m(t)=a\cos(2\pi f_m t) \qquad (2.11)$$

The FM expression will become

$$u(t)=A_c\cos[2\pi f_c t+\beta\sin(2\pi f_m t)]=\text{Re}\{A_c\exp(j2\pi f_c t)\exp[j\beta\sin(2\pi f_m t)]\} \qquad (2.12)$$

We know that

$$\exp[j\beta\sin(2\pi f_m t)]=\sum_{n=-\infty}^{\infty}J_n(\beta)\exp(j2\pi n f_m t) \qquad (2.13)$$

By using (2.13) in (2.12) we get the following

$$u(t)=A_c\sum_{n=-\infty}^{\infty}J_n(\beta)\cos[2\pi(f_c+nf_m)t] \qquad (2.14)$$

We conclude from (2.14) that FM frequency spectrum extends from minus infinity to plus infinity in theory and the spectral components are placed around f_c (carrier) at frequency intervals of $\pm nf_m$

where the respective amplitudes are determined by the values of Bessel function $J_n(\beta)$. A typical (one sided) FM spectrum is shown in Fig. 2.2.

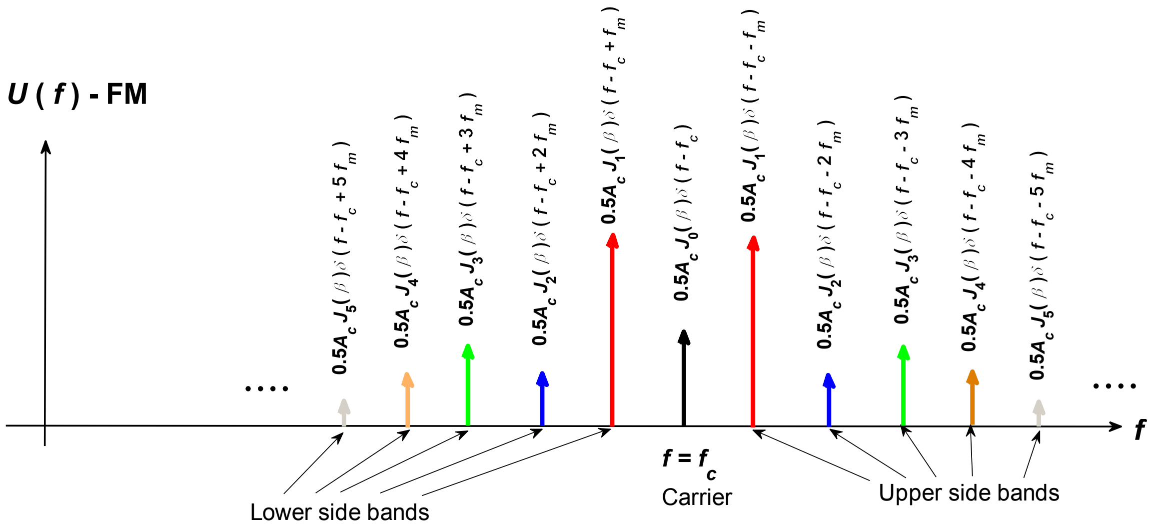

Fig. 2.2 Typical FM spectrum for a single sinusoidal modulating signal.

Of course we cannot tolerate an single message signal to occupy an infinite bandwidth, besides at higher orders of the Bessel function, i.e. higher n , the magnitudes of $J_n(\beta)$ start to become smaller. A reasonable estimate of bandwidth of FM that will accommodate 98 % of the power is given by the following formulation

$$B_{FM} = 2(\beta + 1) f_m \qquad (2.15)$$

Example 2.1.1 : We are given a carrier of $c(t) = 10\cos(2\pi f_c t)$ and a message signal of $m(t) = \cos(20\pi t)$. To generate FM from these two signals, we set $k_f = 50$. Find the related FM expression, modulation index and the required bandwidth to transmit this signal. Plot the resulting FM waveform and the frequency spectrum.

Solution : From (2.4) and (2.5) or (2.6) and (2.8), we have

$$u(t) = 10\cos\left[2\pi f_c t + \phi(t)\right] = 10\cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} \cos(20\pi\tau)\, d\tau\right]$$

$$= 10\cos\left[2\pi f_c t + \frac{50}{10}\sin(20\pi t)\right] \qquad (2.16)$$

By taking $k_f = 50,\ a = 1,\ f_m = 10$ Hz, we get from (2.7)

$$\beta = \frac{k_f a}{f_m} = \frac{50 \times 1}{10} = 5 \qquad (2.17)$$

Finally using (2.15), we find the required bandwidth as

$$B_{FM} = 2(\beta+1) f_m = 120 \text{ Hz} \quad (2.17)$$

Note that this is six times the bandwidth of an AM signal, since in AM, $B_{AM} = 2f_m = 20 \text{ Hz}$. Time waveform and the frequency spectrum of the FM signal are left as class exercise.

2.2 Generation of FM

The easiest way to generate FM is to use a circuit element whose reactance will change with the applied voltage or current. These circuit elements are capacitor and inductor, since

$$X_C = \frac{1}{j2\pi fC} \text{ for capacitor }, X_L = j2\pi fL \text{ for inductor} \quad (2.18)$$

One method is based on varactor diode whose (junction) capacitance changes with the voltage applied, whose circuit diagram is given in Fig. 2.3 (circuit diagram copied directly from Proakis 2002)

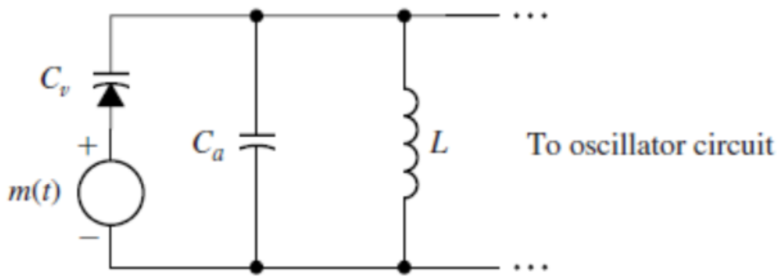

Fig. 2.3 FM generation using varactor diode.

When the message signal is set to zero, that is $m(t) = 0$, we have the carrier frequency generated by the tuned circuit of C_a and L, so that

$$f_c = \frac{1}{2\pi\sqrt{C_a L}} \quad (2.19)$$

When $m(t) \neq 0$, the instantaneous frequency will change as follows

$$f_i(t) = \frac{1}{2\pi\sqrt{L\left[C_a + km(t)\right]}} \quad (2.20)$$

where $km(t)$ is the amount of parallel (time variable) capacitance, C_v added by the varactor. After rearrangement, (2.20) will become

$$f_i(t) = \frac{1}{2\pi\sqrt{LC_a}} \frac{1}{\sqrt{1+\frac{k}{C_a} m(t)}} = f_c \frac{1}{\sqrt{1+\frac{k}{C_a} m(t)}} \quad (2.21)$$

Provided $\frac{k}{C_a}m(t) \ll 1$, we make and expansion of the denominator of the right hand side of (2.21) to arrive at

$$f_i(t) \approx f_c\left[1-\frac{k}{2C_a}m(t)\right] \qquad (2.22)$$

(2.22) is exactly in the form of the instantaneous frequency definition given on the second line of (2.8).

2.3 Demodulation of FM

It is easy to see from (2.8) that the frequency modulations in the FM signal can be converted into amplitude modulations by time differentiating the FM signal given on the second line of (2.8), thus

$$\begin{aligned}\frac{d}{dt}u_{FM}(t) &= \frac{d}{dt}\left\{A_c\cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\tau)d\tau\right]\right\} \\ &= -A_c\left[2\pi f_c + 2\pi k_f m(t)\right]\cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\tau)d\tau\right] \qquad (2.23)\end{aligned}$$

Upon multiplying the term on the second line of (2.23) by a phase synchronized carrier, i.e. by $\cos(2\pi f_c t)$, it will be possible to fully demodulate the message signal as it was done in the case of coherent demodulation of AM. Alternatively envelope detection can be used. This way it is important to realize that FM demodulation involves two distinct stages.

A device to perform this combined task is the phase locked loop (PLL) which is shown in Fig. 2.4

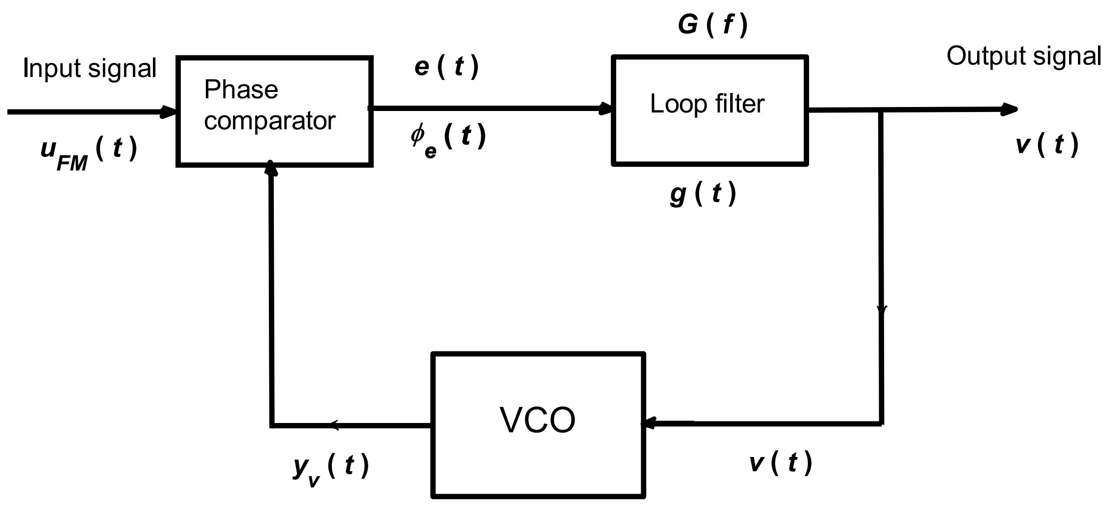

Fig. 2.4 The block diagram of PLL used in the demodulation of FM signal.

The input to PLL is the FM signal, hence

$$u_{FM}(t) = A_c \cos[2\pi f_c t + \phi(t)] = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\tau)\, d\tau\right] \qquad (2.24)$$

Voltage controlled oscillator (VCO) of PLL also acts a FM generator in the following manner

$$y_v(t) = A_v \sin[2\pi f_c t + \phi_v(t)] = A_c \sin\left[2\pi f_c t + 2\pi k_v \int_{-\infty}^{t} v(\tau)\, d\tau\right] \qquad (2.25)$$

while the instantaneous frequency of the VCO is

$$f_v(t) = f_c + \frac{1}{2\pi}\frac{d}{dt}\phi_v(t) = f_c + k_v v(t) \qquad (2.26)$$

After feeding $y_v(t)$ to phase comparator, that acts as a multiplier plus a rejection filter for frequency components around $2f_c$, hence

$$\begin{aligned}
&\text{Input to phase comparator : } u_{FM}(t)\, y_v(t) \\
&\text{Output from phase comparator : } e(t) = \frac{A_c A_v}{2}\sin[\phi(t) - \phi_v(t)] \\
&\text{For small } e(t) \text{ : } \sin[\phi(t) - \phi_v(t)] \approx \phi(t) - \phi_v(t) \quad , \quad \text{set } \phi(t) - \phi_v(t) = \phi_e(t) \qquad (2.27)
\end{aligned}$$

If we substitute for $\phi_v(t)$ on the third line of (2.27) from (2.25), we get

$$\phi_e(t) = \phi(t) - \phi_v(t) = \phi(t) - 2\pi k_v \int_{-\infty}^{t} v(\tau)\, d\tau \qquad (2.28)$$

In (2.28), we differentiate both side with respect to time and rearrange as follows

$$\frac{d}{dt}\phi_e(t) + 2\pi k_v v(t) = \frac{d}{dt}\phi(t) \qquad (2.29)$$

Now $v(t)$ is the output of the loop filter, while $\phi_e(t)$ is the input, hence they will be related by the convolution integral such that

$$v(t) = \int_{-\infty}^{\infty} \phi_e(\tau)\, g(t-\tau)\, d\tau \qquad (2.30)$$

By substituting for $v(t)$in (2.29) from (2.30), we get

$$\frac{d}{dt}\phi_e(t) + 2\pi k_v \int_{-\infty}^{\infty} \phi_e(\tau)\, g(t-\tau)\, d\tau = \frac{d}{dt}\phi(t) \qquad (2.31)$$

Frequency domain equivalent of (2.31) is

$$j2\pi f \phi_e(f) + 2\pi k_v \phi_e(f) G(f) = j2\pi f \phi(f) \qquad (2.32)$$

Deriving $\phi_e(f)$ from (2.32)

$$\phi_e(f) = \frac{\phi(f)}{1 + \frac{k_v}{jf}G(f)} \qquad (2.33)$$

On the other hand

$$V(f) = \phi_e(f)G(f) = \frac{\phi(f)G(f)}{1 + \frac{k_v}{jf}G(f)} \qquad (2.34)$$

In (2.34), if the condition $\frac{k_v}{jf}G(f) \gg 1$ is satisfied, then

$$V(f) \approx \frac{j2\pi f}{2\pi k_v}\phi(f) \qquad (2.35)$$

By using (2.24), we get the time domain equivalent of (2.35),

$$v(t) = \frac{1}{2\pi k_v}\frac{d}{dt}\phi(t) = \frac{1}{2\pi k_v}\frac{d}{dt}\left[2\pi k_f \int_{-\infty}^{t} m(\tau)d\tau\right] = \frac{k_f}{k_v}m(t) \qquad (2.36)$$

In summary we can say that if an FM signal is supplied to the input of PLL, then demodulation is performed by PLL so that we get the message signal, $m(t)$ back at the output.

It is worth pointing out that, in PLL of Fig. 2.4, we have first performed multiplication of $u_{FM}(t)$ the by locally generated carrier (coming from VCO) and the elimination of high frequency component around $2f_c$ in the phase comparator, then carried out the differentiation the loop filter. This way in the PLL of Fig. 2.4, we have reversed the demodulation operation in (2.23) in the following manner

Multiplying by local carrier, $\cos(2\pi f_c t)$ assuming phase locked case

$$u_{FM}(t)\cos(2\pi f_c t) = \frac{A_c}{2}\cos\left[4\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\tau)d\tau\right] + \frac{A_c}{2}\cos\left[2\pi k_f \int_{-\infty}^{t} m(\tau)d\tau\right]$$

After removing the first term at $2f_c$, differentiating the second term and taking the envelope

$$\text{Envelope}\left(\frac{d}{dt}\left\{\frac{A_c}{2}\cos\left[2\pi k_f \int_{-\infty}^{t} m(\tau)d\tau\right]\right\}\right) \rightarrow m_r(t) = A_c\pi k_f m(t) \qquad (2.37)$$

3. Noise Analysis

3.1 Noise Analysis in AM

Since AM systems are known as narrow band systems, we model the noise as narrow band and represent it by

$$n(t) = n_c(t)\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t) \qquad (3.1)$$

where $n_c(t)$ and $n_s(t)$ are known as in phase and quadrature noise components. Being white Gaussian noise, $n(t)$, $n_c(t)$ and $n_s(t)$ have the **flat** frequency spectral density functions of $S_n(f)$, $S_{nc}(f)$ and $S_{ns}(f)$ and their spectral views are shown in Fig. 3.1, where $N_0 = kT$ with $k = 1.38\times10^{-23}$, Boltzman constant and T being the absolute temperature in ${}^0K$.

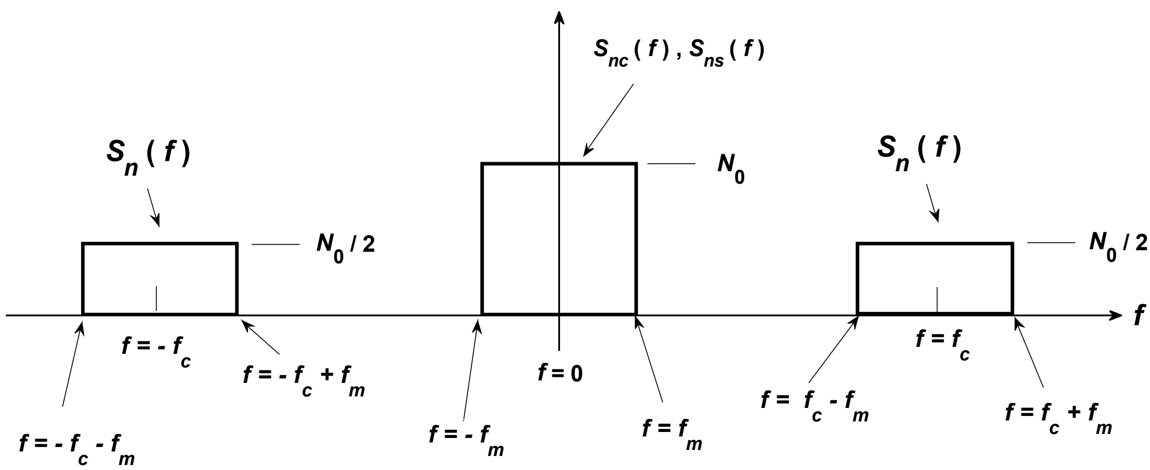

Fig. 3.1 Frequency spectral density appearance of $n(t)$, $n_c(t)$ and $n_s(t)$

From Fig. 3.1, it is clear that the power of $n(t)$ will be equal to power of $n_c(t)$ and $n_s(t)$ so that

$$P_n = \int_{-fc-fm}^{-fc+fm} S_n(f)\,df + \int_{fc-fm}^{fc+fm} S_n(f)\,df = \int_{-fc-fm}^{-fc+fm} \frac{N_0}{2}\,df + \int_{fc-fm}^{fc+fm} \frac{N_0}{2}\,df = 2f_m N_0$$

$$P_{nc} = \int_{-fm}^{fm} S_{nc}(f)\,df = \int_{-fm}^{fm} N_0\,df = 2f_m N_0$$

$$P_{ns} = \int_{-fm}^{fm} S_{ns}(f)\,df = \int_{-fm}^{fm} N_0\,df = 2f_m N_0 \qquad (3.2)$$

On the other hand the right hand side of (3.1) will also have noise power of $2f_m N_0$, since the powers of cosine and sine terms will be equal to 0.5.

Now the received signal will be

$$r(t)=u(t)+n(t) \qquad (3.3)$$

where $n(t)$ will be as given in (3.1) and for $u(t)$, we assume DSB modulation, thus

$$r(t)=A_c m(t)\cos(2\pi f_c t)+n_c(t)\cos(2\pi f_c t)-n_s(t)\sin(2\pi f_c t) \qquad (3.4)$$

In the demodulation process, we will multiply $r(t)$ by the locally generated carrier $\cos(2\pi f_c t+\phi)$ at the receiver, thus

$$\begin{aligned} r(t)\cos(2\pi f_c t+\phi) &= \frac{A_c}{2}m(t)\cos(\phi)+\frac{A_c}{2}m(t)\cos(4\pi f_c t+\phi) \\ &+\frac{1}{2}\left[n_c(t)\cos(\phi)+n_s(t)\sin(\phi)\right] \\ &+\frac{1}{2}\left[n_c(t)\cos(4\pi f_c t+\phi)-n_s(t)\sin(4\pi f_c t+\phi)\right] \qquad (3.5) \end{aligned}$$

Low pass filtering will reject the frequency components around $2f_c$, then we are left with

$$y(t)=\frac{A_c}{2}m(t)\cos(\phi)+\frac{1}{2}\left[n_c(t)\cos(\phi)+n_s(t)\sin(\phi)\right] \qquad (3.6)$$

Assuming the phase difference between the transmitted carrier and the one locally generated at the receiver via PLL, that is $\phi=0$. This way (3.6) becomes

$$y_m(t)=\frac{1}{2}\left[\underbrace{A_c m(t)}_{\text{message signal}}+\overbrace{n_c(t)}^{\text{noise}}\right] \qquad (3.7)$$

Now we estimate the power in the (demodulated) message signal and noise

$$\text{Signal power}: P_s=\frac{A_c^2 P_m}{4} \quad , \quad P_m: \text{Power in } m(t)$$

$$\text{Noise power}: P_{nc2}=\frac{P_{nc}}{4}=\frac{f_m N_0}{2} \qquad (3.8)$$

So the signal to noise ratio (SNR) at the end of the demodulation process will

$$\text{SNR after demodulation}: \frac{\text{Signal power}}{\text{Noise power}}=\frac{P_s}{P_{nc2}}=\frac{A_c^2 P_m}{2 f_m N_0} \qquad (3.9)$$

To make a comparison, let's consider the SNR in an equivalent baseband system where there is no modulation. In this case, the signal and noise power (from 3.2)) would be

$$\text{Signal power}: P_{sb} = A_c^2 P_m$$
$$\text{Noise power}: P_{nb} = P_n = \int_{-f_m}^{f_m} S_n(f)\, df = \int_{-f_m}^{f_m} N_0\, df = 2 f_m N_0 \qquad (3.10)$$

Subsequently SNR if no modulation is applied is found as

$$\text{SNR if no modulation}: \frac{\text{Signal power}}{\text{Noise power}} = \frac{P_{sb}}{P_{nb}} = \frac{A_c^2 P_m}{2 f_m N_0} \qquad (3.11)$$

Comparing (3.11) to (3.9) we see that the SNR in both case is the same. Thus we conclude that DSB process contributes neither positively or negatively to the demodulation process. In other words, the SNR at input to the system is the same as the SNR at output from the system.

3.1 Noise Analysis in FM

Similar to DSB, we again utilize narrow band noise, then the received signal is

$$r(t) = u(t) + n(t) = u(t) + n_c(t)\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t) \qquad (3.12)$$

For this analysis, we ignore the modulation part, hence set $u(t)$ to carrier, then

$$r(t) = c(t) + n(t) = \left[A_c + n_c(t)\right]\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t) \qquad (3.13)$$

(3.13) can be rearranged as

$$r(t) = R(t)\cos\left[2\pi f_c t + \theta(t)\right]$$
$$R(t) = \left\{\left[A_c + n_c(t)\right]^2 + n_s^2(t)\right\}^{0.5} \quad , \quad \theta(t) = \tan^{-1}\left[\frac{n_s(t)}{A_c + n_c(t)}\right] \qquad (3.14)$$

The arrangement of (3.14) is illustrated in Fig. 3.2.in the form of a phasor diagram.

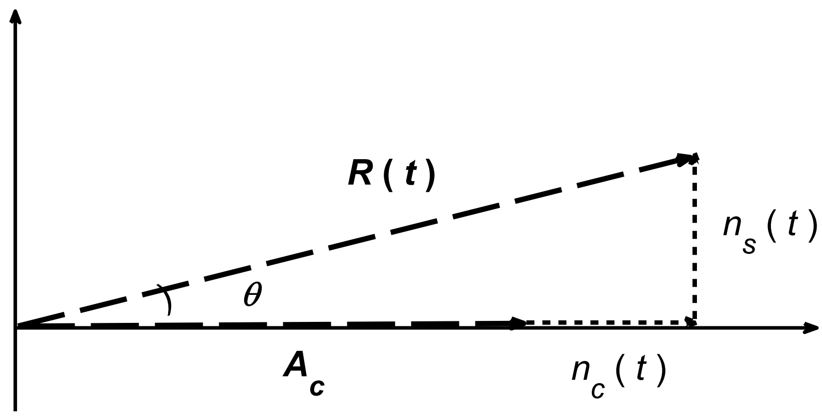

Fig. 3.2 The phasor diagram for the expression in (3.14)

If we are operating under high SNR conditions (which is usually the case), then

$$|n_c(t)| \ll |A_c| \quad , \quad |n_s(t)| \ll |A_c| \quad , \quad \tan^{-1}(x) \approx x \quad , \quad R(t) \approx A_c \qquad (3.15)$$

Under these conditions $r(t)$ of (3.14) will become

$$r(t) = A_c \cos\left[2\pi f_c t + \frac{n_s(t)}{A_c}\right] \qquad (3.16)$$

As explained above the message signal will be extracted from (3.16) by differentiating with respect to time and then taking the envelope of the resulting expression, thus

$$\frac{d}{dt} r(t) = -A_c \left[2\pi f_c + \frac{1}{A_c}\frac{d}{dt} n_s(t)\right] \sin\left[2\pi f_c t + \frac{n_s(t)}{A_c}\right]$$

$$\text{Envelope}\left[\frac{d}{dt} r(t)\right] = A_c \left[2\pi f_c + \frac{1}{A_c}\frac{d}{dt} n_s(t)\right] \qquad (3.17)$$

On the second line of (3.17), the first term of the right hand side corresponds to DC, after dropping this DC term, we get

$$r_m(t) = \frac{d}{dt} n_s(t) \qquad (3.18)$$

As stated when describing the operational aspects of PLL, the time derivative operator, $\frac{d}{dt}$ corresponds to a frequency response of $H(f) = j2\pi f$. This way after the application of (3.18) to the noise component $n_s(t)$, its frequency spectral density will have become

$$S_{ncm}(f)=|H(f)|^2 S_{nc}(f)=4\pi^2 f^2 N_0 \qquad (3.19)$$

The implication of (3.19) is illustrated in Fig. 3.3.

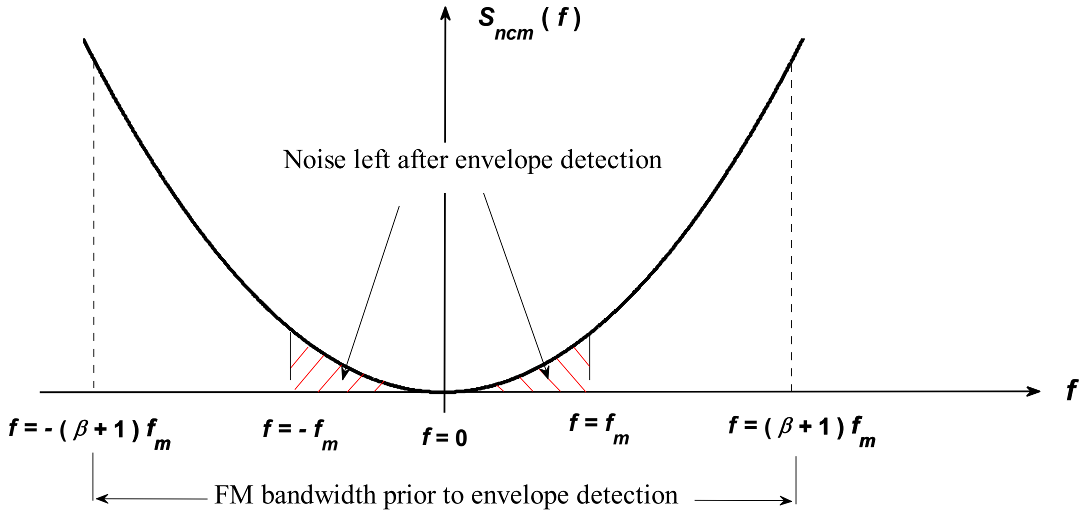

Fig. 3.3 Noise spectral density during FM demodulation stages.

We see both from (3.19) and Fig. 3.3. because of the act of (time) differentiation, the noise spectral density is converted from flat spectral density into parabolic type. After this differentiation, AM demodulation is to be applied. This means that the original FM (baseband) bandwidth has been reduced from $2(\beta+1)f_m$ to $2f_m$. In this process, message signal sidebands at $\pm nf_m$ will fold back into f_m, since they are correlated, but the same thing is not valid for noise, because noise spectral components are uncorrelated. Since AM demodulation will only cover a bandwidth of $2f_m$, we will leave out an important amount of noise outside. This will be particularly so due to the parabolic nature of noise spectral density function of $S_{ncm}(f)$. In the end we can safely claim that there is a considerable SNR improvement from the input of FM demodulator to its output. That is the SNR at the output of FM demodulator is larger than the SNR at its input. It is important to realize that this SNR improvement of FM is achieved at the expense of expanding the message signal bandwidth as illustrated in Fig. 2.2.

To estimate what SNR improvement we have gained, we assume that the message signal power has remained the same during this demodulation process, so it is sufficient to take the ratio of noise power when the bandwidth is $2(\beta+1)f_m$ to that of the bandwidth being reduced to $2f_m$, thus

$$\text{SNR improvement due to noise reduction } = \frac{\int_{-(\beta+1)f_m}^{(\beta+1)f_m} S_{ncm}(f)\, df}{\int_{-f_m}^{f_m} S_{ncm}(f)\, df} = \frac{f^3\Big|_{-(\beta+1)f_m}^{(\beta+1)f_m}}{f^3\Big|_{-f_m}^{f_m}} = (\beta+1)^3 \qquad (3.20)$$

Hence the larger the modulation index (effectively meaning the utilization of larger bandwidth), the more SNR improvement we get.

The above text is based on

1) John G. Proakis, Masoud Salehi, "Communication Systems Engineering" 2nd Ed. 2002, ISBN : 0-13-061793-8.
2) Bernard Sklar, "Digital Communications Fundamentals and Applications", 2nd Ed. Prentice Hall 2002, ISBN : 0-13-084788-7.
3) My own lecture notes.