CHAPTER 5

Angle Modulation

5.1 Introduction

Consider a sinusoid, $A_c \cos(2\pi f_c t + \phi_0)$, where A_c is the (constant) amplitude, f_c is the (constant) frequency in Hz and ϕ_0 is the initial phase angle. Let the sinusoid be written as $A_c \cos[\theta(t)]$ where $\theta(t) = 2\pi f_c t + \phi_0$. In chapter 4, we have seen that relaxing the condition that A_c be a constant and making it a function of the message signal $m(t)$, gives rise to amplitude modulation. We shall now examine the case where A_c is a constant but $\theta(t)$, instead of being equal to $2\pi f_c t + \phi_0$, is a function of $m(t)$. This leads to what is known as the angle modulated signal. Two important cases of angle modulation are Frequency Modulation (FM) and Phase modulation (PM). Our objective in this chapter is to make a detailed study of FM and PM.

An important feature of FM and PM is that they can provide much better protection to the message against the channel noise as compared to the linear (amplitude) modulation schemes. Also, because of their constant amplitude nature, they can withstand nonlinear distortion and amplitude fading. The price paid to achieve these benefits is the increased bandwidth requirement; that is, the *transmission bandwidth* of the FM or PM signal with constant amplitude and which can provide noise immunity is much larger than $2W$, where W is the highest frequency component present in the message spectrum.

Now let us define PM and FM. Consider a signal $s(t)$ given by $s(t) = A_c \cos\left[\theta_i(t)\right]$ where $\theta_i(t)$, the instantaneous angle quantity, is a function of $m(t)$. We define the instantaneous frequency of the angle modulated wave $s(t)$, as

$$f_i(t) = \frac{1}{2\pi} \frac{d\,\theta_i(t)}{d\,t} \tag{5.1}$$

(The subscript i in $\theta_i(t)$ or $f_i(t)$ is indicative of our interest in the instantaneous behavior of these quantities). If $\theta_i(t) = 2\pi f_c t + \phi_0$, then $f_i(t)$ reduces to the constant f_c, which is in perfect agreement with our established notion of frequency of a sinusoid. This is illustrated in Fig. 5.1.

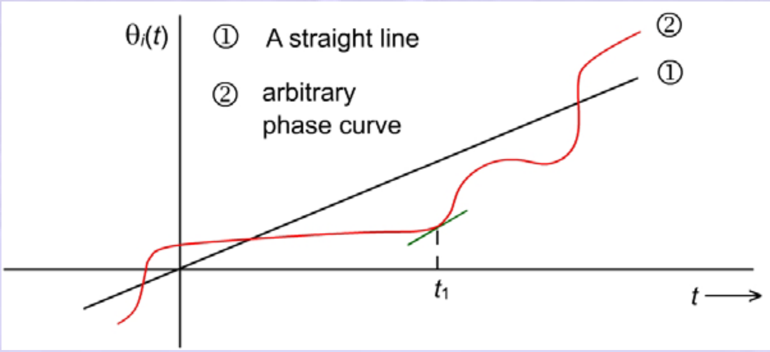


Fig. 5.1: Illustration of instantaneous phase and frequency

Curve ① in Fig. 5.1 depicts the phase behavior of a constant frequency sinusoid with $\phi_0 = 0$. Hence, its phase, as a function of time is a straight line; that is $\theta_i(t) = 2\pi f_c t$. Slope of this line is a constant and is equal to the frequency of the sinusoid. Curve ② depicts an arbitrary phase behavior; its slope changes with time. The instantaneous frequency (in radians per second) of this signal at $t = t_1$ is given by the slope of the tangent (green line) at that time.

a) Phase modulation

For PM, $\theta_i(t)$ is given by

$$\theta_i(t) = 2\pi f_c t + k_p m(t) \tag{5.2}$$

The term $2\pi f_c t$ is the angle of the *unmodulated* carrier and the constant k_p is the *phase sensitivity* of the modulator with the units, radians per volt. (For convenience, the initial phase angle of the unmodulated carrier is assumed to be zero). Using Eq. 5.2, the phase modulated wave $s(t)$ can be written as

$$\left[s(t)\right]_{PM} = A_c \cos\left[2\pi f_c t + k_p m(t)\right] \tag{5.3}$$

From Eq. 5.2 and 5.3, it is evident that for PM, the phase deviation of $s(t)$ from that of the unmodulated carrier phase is a linear function of the base-band message signal, $m(t)$. The instantaneous frequency of a phase modulated signal depends on $\dfrac{d\,m(t)}{d\,t} = m'(t)$ because $\dfrac{1}{2\pi}\dfrac{d\theta_i(t)}{d\,t} = f_c + \dfrac{k_p}{2\pi} m'(t)$.

b) Frequency Modulation

Let us now consider the case where $f_i(t)$ is a function of $m(t)$; that is,

$$f_i(t) = f_c + k_f m(t) \tag{5.4}$$

or

$$\theta_i(t) = 2\pi \int_{-\infty}^{t} f_i(\tau)\, d\tau \tag{5.5}$$

$$= 2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\tau)\, d\tau \tag{5.6}$$

k_f is a constant, which we will identify shortly. A frequency modulated signal $s(t)$ is described in the time domain by

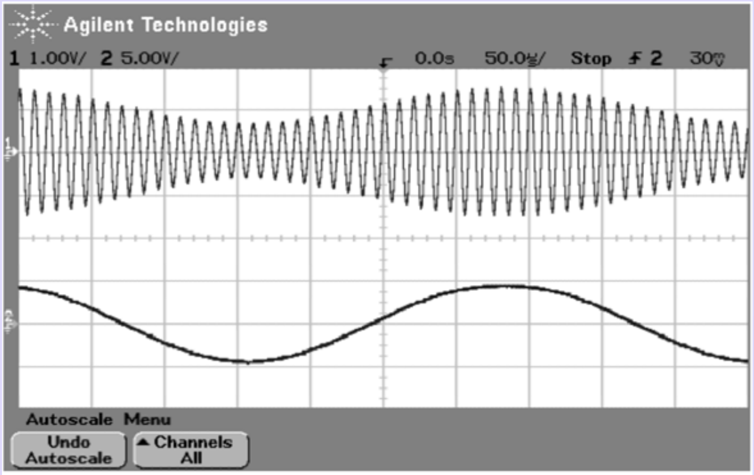
$$\left[s(t)\right]_{FM} = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\tau)\, d\tau\right] \tag{5.7}$$

k_f is termed as the *frequency sensitivity* of the modulator with the units Hz/volt. From Eq. 5.4 we infer that for an FM signal, the instantaneous frequency deviation of $s(t)$ from the (unmodulated) carrier frequency f_c is a linear function of $m(t)$. Fig. 5.2 to 5.5 illustrate the experimentally generated AM, FM and PM waveforms for three different base-band signals. From these illustrations, we observe the following:

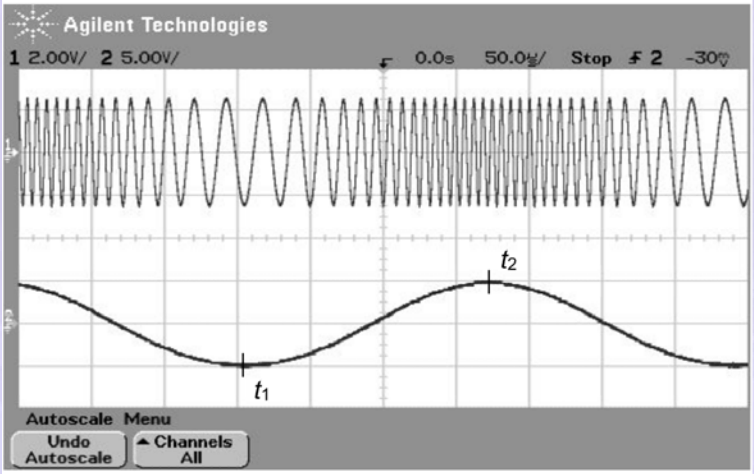
i) Unlike AM, the zero crossings of PM and FM waves are not uniform (zero crossings refer to the time instants at which a waveform changes from negative to positive and vice versa).

ii) Unlike AM, the envelope of PM or FM wave is a constant.

iii) From Fig. 5.2(b) and 5.3(b), we find that the minimum instantaneous frequency of the FM occurs (as expected) at those instants when $m(t)$ is most negative (such as $t = t_1$) and maximum instantaneous frequency occurs at those time instants when $m(t)$ attains its positive peak value, m_p (such as $t = t_2$). When $m(t)$ is a square wave (Fig. 5.4), it can assume only two possible values. Correspondingly, the instantaneous frequency has only two possibilities. This is quite evident in Fig. 5.4(b).

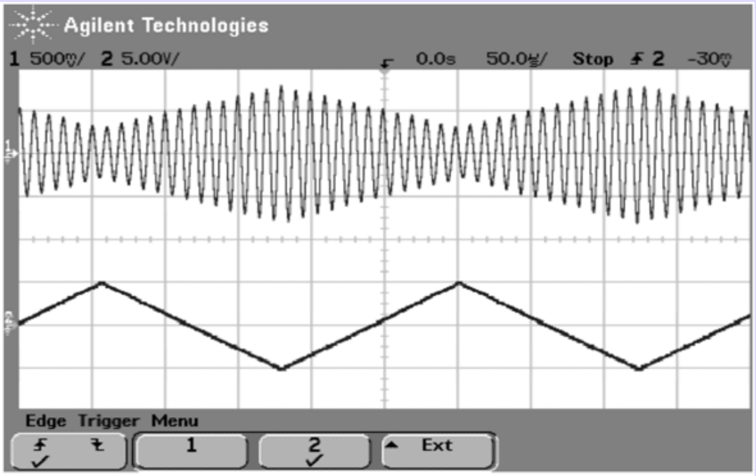


(a) AM

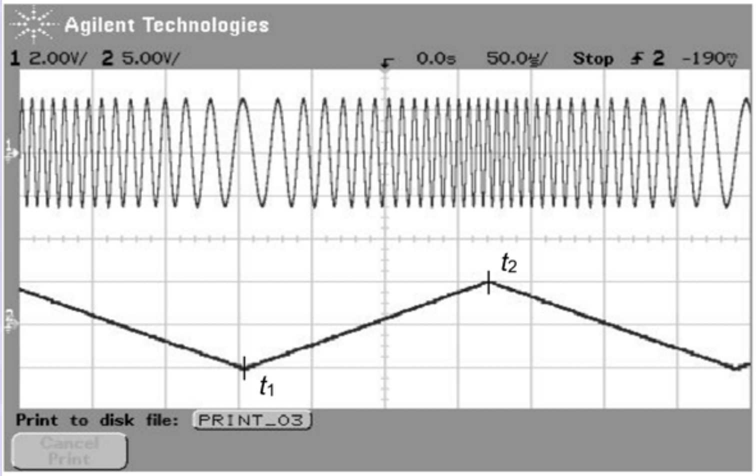


(b) FM

Fig 5.2: AM and FM with tone modulation

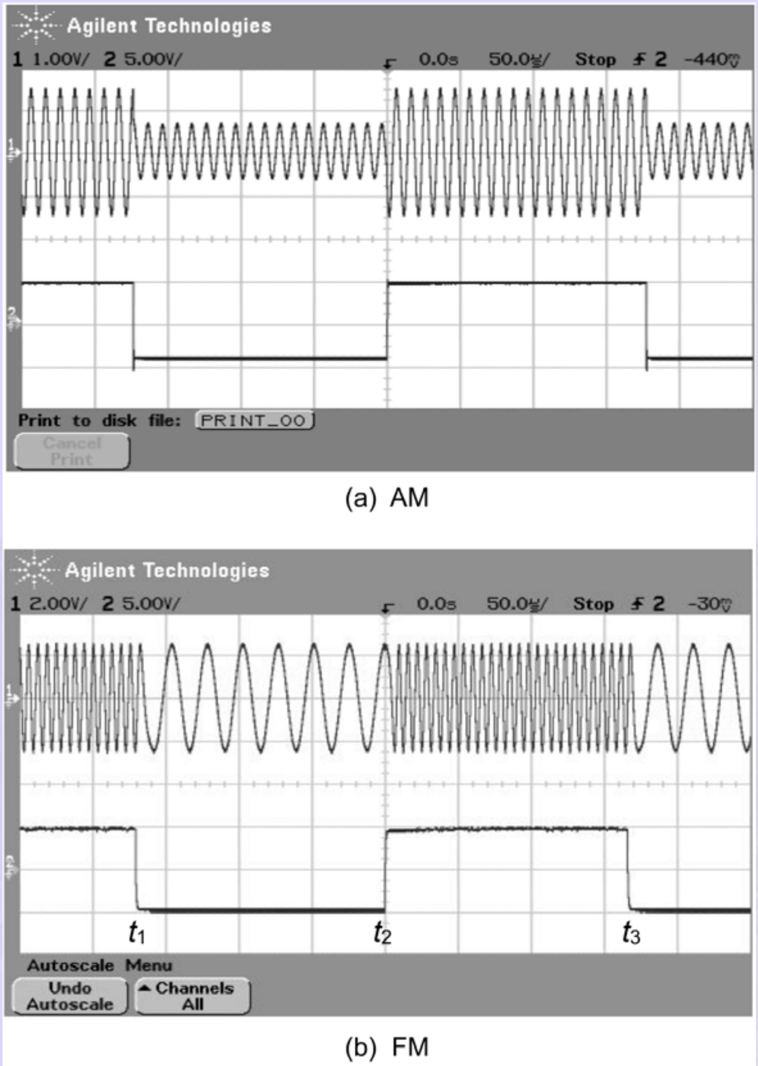


(a) AM



(b) FM

Fig. 5.3: AM and FM with the triangular wave shown as $m(t)$



(a) AM

(b) FM

Fig. 5.4: AM and FM with square wave shown as $m(t)$

iv) A triangular wave has only two possibilities for its slope. In Fig. 5.5(b), it has a constant positive slope between t_1 and t_2, and constant negative slope to the right of t_2 for the remaining part of the cycle shown. Correspondingly, the PM wave has only two values for its $f_i(t)$, which is evident from the figure.

v) The modulating waveform of Fig. 5.5(c) is a square wave. Except at time instants such as $t = t_1$, it has zero slope and $\left.\dfrac{d\,m(t)}{d\,t}\right|_{t = t_1}$ is an impulse.

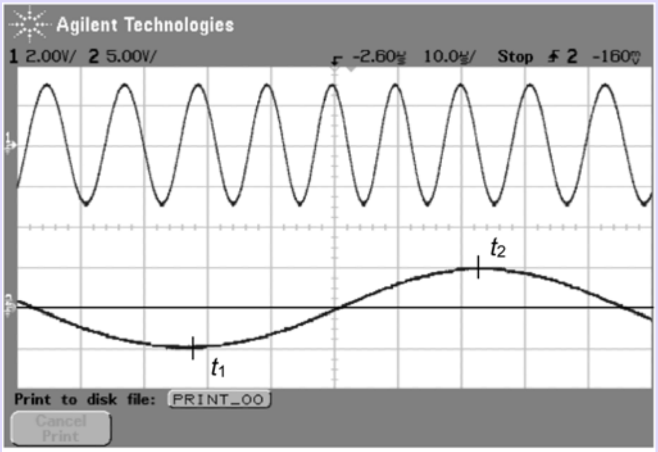
Therefore, the modulated carrier is simply a sinusoid of frequency f_c, except at the time instants when $m(t)$ changes its polarity. At $t = t_1$, $f_i(t)$ has to be infinity. This is justified by the fact that at $t = t_1$, $\theta_i(t)$ undergoes sudden phase change (as can be seen in the modulated waveform) which implies $\frac{d\,\theta_i(t)}{d\,t}$ tends to become an impulse.

Eq. 5.3 and 5.7 reveal a close relationship between PM and FM. Let $m_I(t) = \int_{-\infty}^{t} m(\tau)\,d\tau$. If $m_I(t)$ phase modulates a carrier with modulator sensitivity $k_p = 2\pi k_f$, then the resulting signal is actually an FM signal as given by Eq. 5.7. Similarly a PM signal can be obtained using frequency modulator by differentiating $m(t)$ before applying it to the frequency modulator. (Because of differentiation, $m(t)$ should not have any discontinuities.)

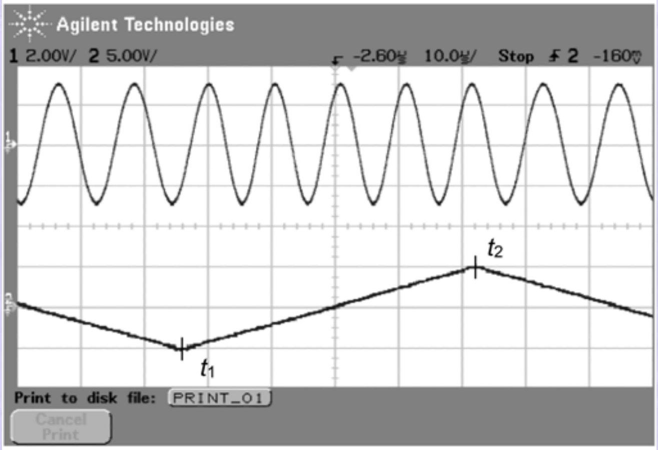
As both PM and FM have constant amplitude A_c, the average power of a PM or FM signal is,

$$P_{av} = \frac{A_c^2}{2},$$

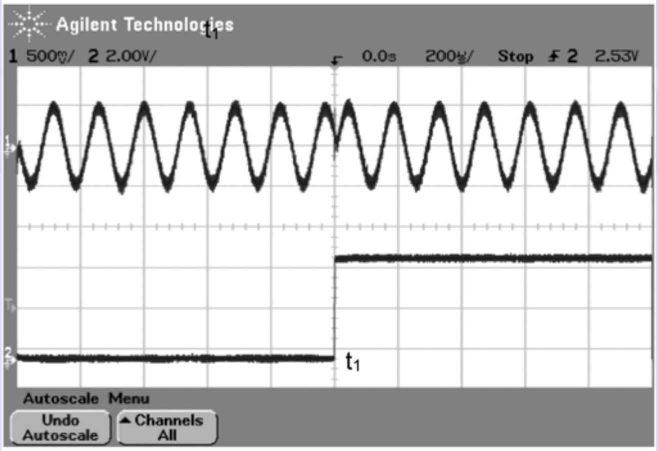
regardless of the value of k_p or k_f.



(a)



(b)



(c)

Fig 5.5: PM with $m(t)$ (a) a sine wave

(b) a triangular wave

(c) a square wave

Example 5.1

An angle modulated signal is given by

$$s(t) = \cos\left[2\pi\left(2 \times 10^6 t + 30\sin(150t) + 40\cos(150t)\right)\right]$$

Let us find the maximum phase and frequency derivations.

The terms $\left[30\sin(150t) + 40\cos(150t)\right]$ can be expressed in the form $\left[\cos\alpha\sin(150t) + \sin\alpha\cos(150t)\right]$.

As $\sqrt{30^2 + 40^2} = 50$, we have

$$30\sin(150t) + 40\cos(150t) = 50\left(\frac{3}{5}\sin(150t) + \frac{4}{5}\cos(150t)\right)$$

$$= 50\sin(150t + \varphi) \text{ where } \varphi = \tan^{-1}\frac{4}{3}$$

Hence $s(t) = \cos\left[4\pi \times 10^6 t + 100\pi\sin(150t + \varphi)\right]$

Evidently, the maximum phase deviation is (100π).

Let $\psi(t) = 100\pi\sin(150t + \varphi)$

$$\frac{1}{2\pi}\frac{d\psi(t)}{dt} = 50\cos(150t + \varphi)\cdot 150$$

$$= 7500\cos(150t + \varphi)$$

Hence maximum frequency deviation $= 7500$ Hz. ♦

Example 5.2

Let $m(t)$ be a periodic triangular wave with period 10^{-3} sec. with $m(t)_{max} = -m(t)_{min} = 1$ volt. We shall find the maximum and minimum values of the instantaneous frequency for

a) FM with $k_f = 10^4$ Hz/volt

b) PM with $k_p = \pi$ rad/volt

Assume the carrier frequency to be 100 kHz.

a) For FM, $f_i(t) = f_c + k_f\, m(t)$

$$\left(f_i(t)\right)_{\min} = 100 \times 10^3 - 10^4 = 90 \text{ kHz}$$

$$\left(f_i(t)\right)_{\max} = 100 \times 10^3 + 10^4 = 110 \text{ kHz}$$

b) For PM, $f_i(t) = f_c + \dfrac{k_p}{2\pi} m^{'}(t)$

Note that $m^{'}(t)$ is a square wave with maximum and minimum values as $\pm\, 4{,}000$.

Hence $\left(f_i(t)\right)_{\min} = 100 \times 10^3 - \dfrac{1}{2} \times 4{,}000$

$$= 98 \text{ kHz}$$

Similarly, $\left(f_i(t)\right)_{\max} = 102 \text{ kHz}$ ♦

Example 5.3

Let $s(t)$ be a general angle modulated signal given by

$$s(t) = A_c \cos\left[\theta_i(t)\right] = A_c \cos\left[\omega_c t + \varphi(t)\right]$$

It is given that when $m(t) = \cos \omega_m t$, $s(t)$ has the instantaneous frequency given by $f_i(t) = f_c + 2\pi k \left(f_m\right)^2 \cos \omega_m t$, where k is a suitable constant. Let us find the expression for $\theta_i(t)$. If $m(t)$ is different from $\cos \omega_m t$, what could be the expression for $\theta_i(t)$ and $s(t)$.

$$\frac{1}{2\pi} \frac{d\,\theta_i(t)}{d\,t} = f_i(t) = f_c + 2\pi k \left(f_m\right)^2 \cos \omega_m t$$

$$\frac{d\,\theta_i(t)}{d\,t} = 2\pi f_c + \omega_m^2\, k \cos\left(\omega_m t\right)$$

or $\qquad \theta_i(t) = 2\pi f_c t + \omega_m k \sin(\omega_m t)$

$$= 2\pi f_c t - k \frac{d}{dt}\left[\cos(\omega_m t)\right]$$

Generalizing,

$$\theta_i(t) = 2\pi f_c t - k \frac{d}{dt}\left[m(t)\right]$$

and $\qquad s(t) = A_c \cos\left[2\pi f_c t - k \frac{d\,m(t)}{dt}\right] \qquad \blacklozenge$

Exercise 5.1

A periodic signal $m(t)$ angle modulates a very high frequency carrier. Let the modulated signal be given by, $s(t) = A_c \cos\left[2\pi \times 10^8 t + k_p m(t)\right]$, where $m(t)$ is as shown in Fig. 5.6.

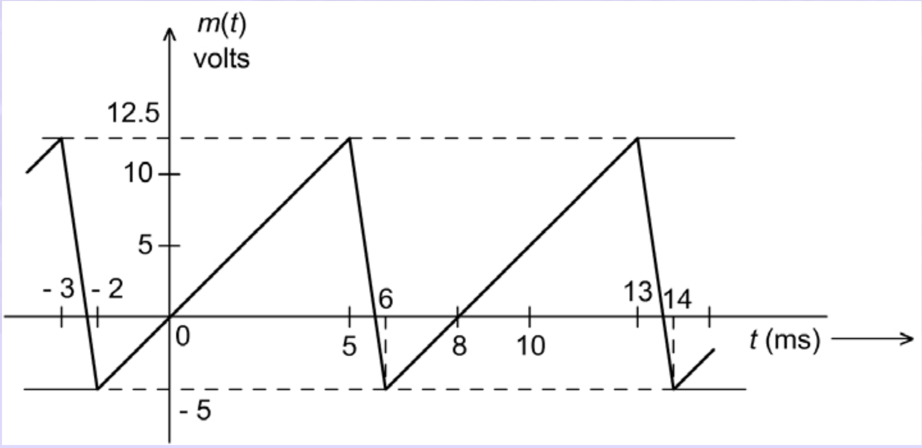


Fig. 5.6: Modulating signal for the exercise 5.1

If $(f_i)_{max} - (f_i)_{min}$ is to be 100 kHz show that $k_p = 10\pi$ rad/volt.

5.2 Bandwidth of FM

In this section, we shall make a detailed analysis of the bandwidth requirements of FM. (PM is considered in section 5.4.) Let $c_f = 2\pi k_f$. For convenience, let $s(t)$ denote the FM signal. With $m_I(t) = \int_{-\infty}^{t} m(\tau) d\tau$, the pre-envelope of $s(t)$ is

$$\left[s(t)\right]_{pe} = A_c \exp\left[j\left(\omega_c t + c_f m_I(t)\right)\right]$$

$$= A_c \left[1 + j c_f m_I(t) - \frac{c_f^2}{2!}\left(m_I(t)\right)^2 + \cdots + j^n \frac{c_f^n}{n!}\left(m_I(t)\right)^n + \cdots\right] e^{j \omega_c t} \tag{5.8}$$

As $s(t) = \text{Re}\left\{\left[s(t)\right]_{pe}\right\}$, we have

$$s(t) = A_c \left[\cos(\omega_c t) - c_f m_I(t) \sin(\omega_c t) - \frac{c_f^2}{2!}\left(m_I(t)\right)^2 \cos(\omega_c t) + \cdots\right] \tag{5.9}$$

If $m(t)$ is band-limited to W, then so is $m_I(t)$, as $M_I(f) = \dfrac{M(f)}{(j 2\pi f)}$. But $s(t)$ consists of terms $\left(m_I(t)\right)^2 \cos(\omega_c t)$, $\left(m_I(t)\right)^3 \sin(\omega_c t)$ etc. The spectrum of $\left(m_I(t)\right)^2$ is of width $2W$, and that of $\left(m_I(t)\right)^3$ is $3W$ etc. The spectrum of $\left(m_I(t)\right)^2 \cos(\omega_c t)$ occupies the interval $f_c \pm 2W$ and that of $\left(m_I(t)\right)^n \cos(\omega_c t)$ occupies the interval $f_c \pm nW$. Clearly, the spectrum of the FM signal is not band-limited. It appears that, at least theoretically, it has infinite bandwidth and $S(f)$ is not simply related to $M(f)$. A similar situation prevails even in the case of a PM signal. (Recall that the bandwidth of any linear modulated signal is less than or equal to $2W$, where W is the message band-width). Let us take a closer look at this issue of FM bandwidth.

5.2.1 NarrowBand FM (NBFM)

Assume that $\left|c_f\, m_I(t)\right|_{max} \ll 1$. Then $s(t)$ of Eq. 5.9 can be well approximated by the first two terms; that is,

$$s(t) \simeq A_c\left[\cos(\omega_c t) - c_f\, m_I(t)\sin(\omega_c t)\right] \qquad (5.10)$$

Eq. 5.10 is similar to that of an AM wave. Now $S(f)$ is band-limited to $2W$ as in the case of an AM signal[1]. For this reason, the case $\left|c_f\, m_I(t)\right|_{max} \ll 1$ is called *narrowband FM (NBFM)*. Similarly, the *narrowband PM (NBPM)* signal can be written as

$$s(t) \simeq A_c\left[\cos(\omega_c t) - k_p\, m(t)\sin(\omega_c t)\right] \qquad (5.11)$$

Though the expressions for NBFM or NBPM resemble fairly closely that of an AM signal, there is one important difference; namely, the sideband spectrum of PM or FM has a phase shift of $\frac{\pi}{2}$ with respect to the carrier. This makes the time domain signal of PM or FM very much different from that of an AM signal. A NBFM signal has phase variations with very little amplitude variations whereas the AM signal has amplitude variations with no phase variations.

Note that from Eq. 5.10, we have the envelope of the NBFM signal given by

$$A(t) = A_c\sqrt{1 + c_f^2\, m_I^2(t)}$$

If $\left|c_f\, m_I(t)\right| \ll 1$, then $A(t) \simeq A_c$.

Example 5.4

---

[1] In general, FM is nonlinear as

$$A_c \cos\left\{\omega_c t + c_f\left[m_{I1}(t) + m_{I2}(t)\right]\right\} \neq A_c \cos\left\{\omega_c t + c_f\, m_{I1}(t)\right\} + A_c \cos\left\{\omega_c t + c_f\, m_{I2}(t)\right\}$$

A NBFM signal, $s(t)$, is generated by using $m(t) = 10^3 \sin c^2 \left(10^3 t\right)$. We shall find $S(f)$.

From Eq. 5.10, we have

$s(t) = A_c \left[\cos(\omega_c t) - c_f \, m_I(t) \sin(\omega_c t)\right]$

But $\quad m_I(t) \longleftrightarrow \dfrac{M(f)}{j\,2\pi f} \qquad$ and $\quad c_f = 2\pi k_f$.

Hence

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

where $M(f) = tri\left(\dfrac{f}{10^3}\right)$ ♦

Exercise 5.2

Let $M(f) = \dfrac{1}{10^3}\, ga\left(\dfrac{f}{10^3}\right)$ and $f_c = 10^6$ Hz. It is given that $k_f = 250$ Hz/volt, and $A_c = 4$ V. Sketch $S(f)$ for a NBFM signal.

5.2.2 WideBand FM (WBFM)

If the condition $\left|c_f \, m_I(t)\right|_{max} << 1$ is not satisfied, then we have the wideband FM, which, as mentioned earlier has, at least theoretically, infinite bandwidth. However, as will be seen a little later, most of the power of the FM signal resides in a finite bandwidth, called the **transmission bandwidth**. In order to estimate this bandwidth, we observe that

$f_i(t) = f_c + k_f \, m(t)$

Let $m_p = m(t)_{max} = \left|m(t)_{min}\right|$. Then the instantaneous frequency varies in the range $(f_c - k_f\, m_p)$ to $(f_c + k_f\, m_p)$. As the total range of frequency deviation (with centre at f_c) is $2k_f\, m_p$, can we assume that, for all practical purposes, the transmission bandwidth of an FM signal, $(B_T)_{FM}$ is

$$(B_T)_{FM} \overset{?}{=} 2k_f\, m_p$$

Let Δf denote the *maximum deviation* of the carrier frequency f_c. That is, $\Delta f = k_f\, m_p$. Then, can we assume that, to a very good approximation, $(B_T)_{FM} = 2\Delta f$?

The above expression for the transmission bandwidth is valid only when $\Delta f \gg W$. In the case $\Delta f \ll W$, $(B_T)_{FM} \neq 2\Delta f$ but $2W$, as will be shown later. This is indeed the fallacy that gave birth to FM in the first place. In the 1920s radio engineers, thinking that the actual bandwidth of an FM signal is $2\Delta f$, felt that bandwidth requirement of FM can be made less than that of AM (that is, less than $2W$) by choosing Δf appropriately! The fallacy here lies in equating the instantaneous frequency to the spectral frequency. Although $f_i(t)$ is measured in Hz, it should not be equated with spectral frequency. Spectral frequency f is the independent variable of the frequency domain where as $f_i(t)$ is a time dependent quantity indicating the time behavior of a signal with angle modulation. When we refer to the spectrum $X(f)$ of a signal $x(t)$, we imply that $x(t)$ is composed of the complex exponentials with the magnitude and phase specified by $X(f)$, and each one of these exponentials exists for all t with the given frequency. $f_i(t)$, on the other hand represents the frequency of a cosine signal that can be treated to have a constant frequency for a very short duration, maybe only for a few cycles. We shall now give some justification that bandwidth of an FM signal is never less than $2W$.

Let the base-band signal $m(t)$ be approximated by a staircase signal, as shown in Fig. 5.7(a). From the sampling theorem (to be discussed in chapter 6), this staircase approximation is justified as long as the width of each rectangle (separation between two adjacent samples) is less than or equal to $\frac{1}{2W}$. For convenience, we shall treat the adjacent sample separation to be equal to $\frac{1}{2W}$.

The FM wave for the staircase-approximated signal will consist of a sequence of sinusoidal pulses, each of a constant frequency and duration of $\frac{1}{2W}$ sec. A typical pulse and its spectrum are shown in Fig. 5.7(b) and (c) respectively.

The frequency of the RF pulse in the interval $\left(t_k , t_k + \frac{1}{2W}\right)$ is $f_i(t_k) = f_c + k_f\, m(t_k)$. Hence, its spectral range (between the first nulls) as shown in Fig. 5.7(c), is from $\left[f_c + k_f\, m(t_k) - 2W\right]$ to $\left[f_c + k_f\, m(t_k) + 2W\right]$. Clearly the significant part of the spectrum of the FM signal will lie in the interval $\left(f_c - k_f\, m_p - 2W\right)$ to $\left(f_c + k_f\, m_p + 2W\right)$. (Note that about 92% of the energy of a rectangular pulse of duration T sec, lies in the frequency range $|f| \le \frac{1}{T}$). Hence, we can take as one possible value of the transmission bandwidth of an FM signal, as

$$(B_T)_{FM} = 2k_f\, m_p + 4W = 2\Delta f + 4W = 2(\Delta f + 2W) \qquad (5.12)$$

For the wideband FM case, where $\Delta f >> W$, $(B_T)_{FM}$ can be well approximated by $2\Delta f$.

In the literature, other rules of thumb for FM bandwidths are to be found. The most commonly used rule, called the *Carson's* rule, is

$$(B_T)_{FM} = 2(\Delta f + W) \qquad (5.13)$$

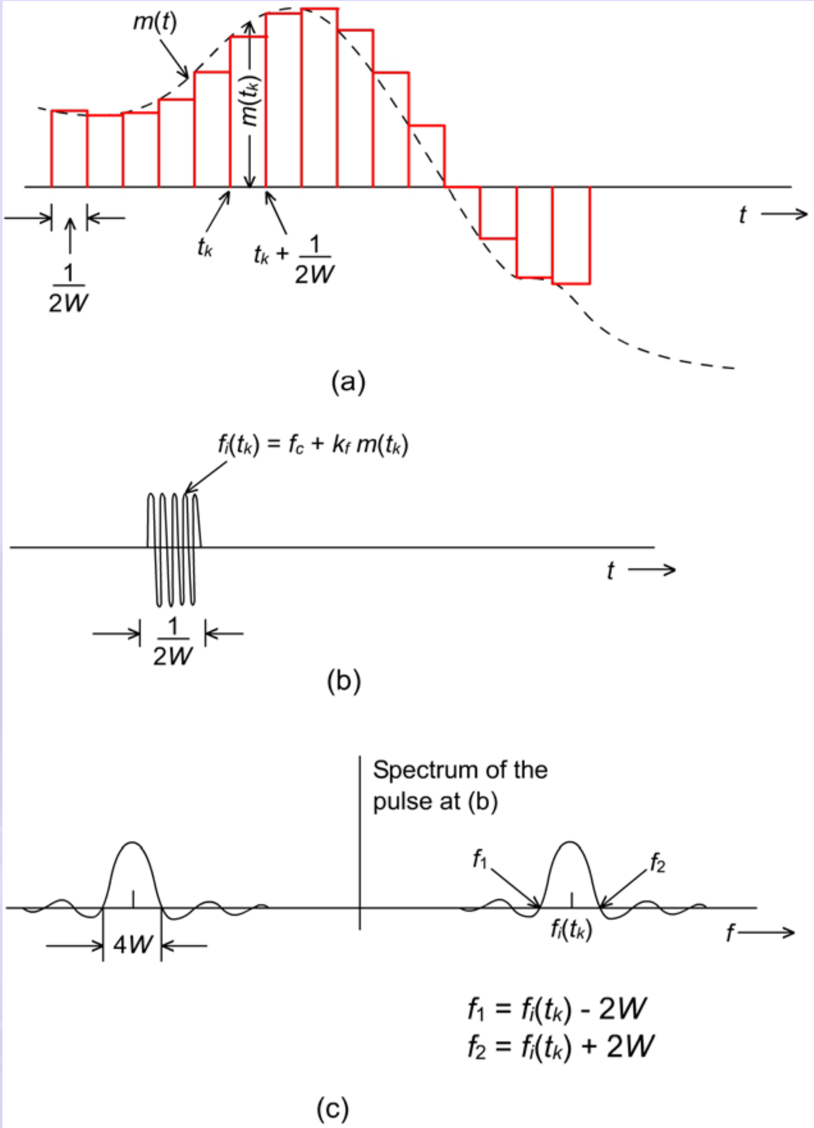


Fig. 5.7: Estimation of FM spectral width

Carson's rule gives a better bandwidth estimate than does Eq. 5.12 for the NBFM $(\Delta f \ll W)$ case. This is in agreement with our result on NBFM, namely, its bandwidth is approximately $2W$. In other cases, where $\Delta f \ll W$ is not satisfied (wideband and intermediate cases) Eq. 5.12 gives a better estimate than does Eq. 5.13.

We now define the *deviation ratio* D, as

$$D = \frac{\Delta f}{W} \tag{5.14}$$

Then Eq. 5.12 and 5.13 can be combined into

$$\left(B_T\right)_{FM} = 2W\left(D + k\right) \tag{5.15}$$

where k varies between 1 and 2. As can be seen, Eq. 5.12 and Eq. 5.13 are special cases of Eq. 5.15.

5.3 Tone Modulation

Let $m(t) = A_m \cos(2\pi f_m t)$

Then $m_I(t) = \frac{A_m}{\omega_m} \sin(\omega_m t)$, assuming $m_I(-\infty) = 0$

$$s_{pe}(t) = A_c \exp\left[j\left(\omega_c t + c_f \frac{A_m}{\omega_m} \sin(\omega_m t)\right)\right]$$

$$= A_c \exp\left[j\left(\omega_c t + \frac{k_f A_m}{f_m} \sin(\omega_m t)\right)\right]$$

For tone modulation, $\Delta f = k_f A_m$ and $W = f_m$. In this case, the deviation ratio is referred to as the **modulation index** and is usually denoted by the symbol β. That is, for tone modulation, $\beta = \frac{k_f A_m}{f_m}$ and

$$s_{pe}(t) = A_c e^{j\omega_c t}\left[e^{j\beta\sin(\omega_m t)}\right] \tag{5.16a}$$

$$s(t) = \text{Re}\left[s_{pe}(t)\right] = A_c \cos\left(\omega_c t + \beta\sin(\omega_m t)\right) \tag{5.16b}$$

5.3.1 NBFM

$$s(t) = A_c \cos\left[\omega_c t + \beta\sin(\omega_m t)\right]$$

$$= A_c \left\{\cos(\omega_c t)\cos\left(\beta\sin(\omega_m t)\right) - \sin(\omega_c t)\sin\left(\beta\sin(\omega_m t)\right)\right\}$$

NBFM: small β .

Let β be small enough so that we can use the approximations

$$\cos\left(\beta \sin\left(\omega_m t\right)\right) \simeq 1 \text{ , and}$$

$$\sin\left(\beta \sin\left(\omega_m t\right)\right) \simeq \beta \sin\left(\omega_m t\right).$$

Then, $$s(t) = A_c\left[\cos\left(\omega_c t\right) - \beta \sin\left(\omega_m t\right) \sin\left(\omega_c t\right)\right] \quad (5.17a)$$

$$= A_c\left\{\cos\left(\omega_c t\right) + \frac{\beta}{2}\left[\cos\left(\omega_c + \omega_m\right)t - \cos\left(\omega_c - \omega_m\right)t\right]\right\} \quad (5.17b)$$

Corresponding expression for AM (Eq 4.7) is

$$\left[s(t)\right]_{AM} = A_c\left\{\cos\left(\omega_c t\right) + \frac{\mu}{2}\left[\cos\left(\omega_c + \omega_m\right)t + \cos\left(\omega_c - \omega_m\right)t\right]\right\}$$

Eq. 5.17(b) can be written as

$$s(t) = \operatorname{Re}\left\{A_c e^{j \omega_c t} + \frac{A_c \beta}{2}\left[e^{j(\omega_c + \omega_m)t} - e^{-j(\omega_c - \omega_m)t}\right]\right\} \quad (5.17c)$$

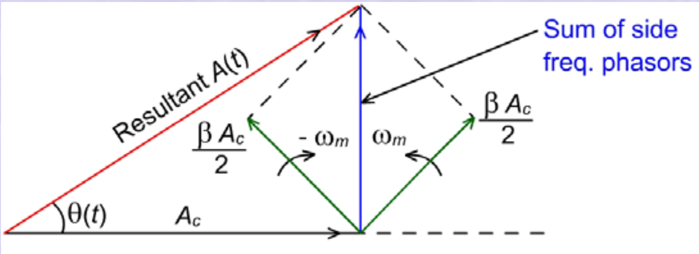


Fig. 5.8: Phasor diagram for NBFM with tone modulation

Using Eq. 5.17(c), we construct the phasor diagram for NBFM, shown in Fig. 5.8. (We have taken the carrier phasor as the reference.) Comparing the phasor diagram for the NBFM with that of the AM signal (Fig. 4.26), we make the following observation.

In the case of AM, the resultant of the side-band phasors is collinear with the carrier phasor whereas, it is perpendicular to the carrier phasor in NBFM.

It is this quadrature relationship between the two phasors of NBFM that produces angle variations resulting in corresponding changes in $f_i(t)$.

5.3.2 WBFM

The exponential term in the brackets in Eq. 5.16(a) is a periodic signal with period $\frac{1}{f_m}$. Expressing this term in Fourier series, we have

$$e^{j\beta\sin(\omega_m t)} = \sum_{n=-\infty}^{\infty} x_n \, e^{j n \omega_m t}$$

where $x_n = f_m \int_{-\frac{1}{2f_m}}^{\frac{1}{2f_m}} e^{j\beta\sin(\omega_m t)} \, e^{-jn\omega_m t} \, dt$

Let $\theta = \omega_m t$. Then,

$$x_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(\beta\sin\theta - n\theta)} \, d\theta \tag{5.18}$$

The integral on the RHS of Eq. 5.18 is recognized as the *nth order Bessel function of the first kind* and argument β and is commonly denoted by the symbol $J_n(\beta)$; that is,

$$x_n = J_n(\beta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(\beta\sin\theta - n\theta)} \, d\theta \tag{5.19a}$$

That is,

$$e^{j\beta\sin(\omega_m t)} = \sum_{n=-\infty}^{\infty} J_n(\beta) e^{jn\omega_m t} \tag{5.19b}$$

Hence,

$$e^{j\beta\sin(\omega_m t)} \longleftrightarrow \sum_{n=-\infty}^{\infty} J_n(\beta)\delta(f - nf_m) \tag{5.19c}$$

Using Eq. 5.19, we can express $s_{pe}(t)$ as

$$s_{pe}(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) e^{j\left[2\pi(f_c + n f_m)t\right]} \tag{5.20}$$

As $s(t) = \mathrm{Re}\left[s_{pe}(t)\right]$, we obtain

$$s(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \cos\left[2\pi(f_c + n f_m)t\right] \tag{5.21}$$

Taking the Fourier transform of $s(t)$, we have

$$S(f) = \frac{A_c}{2} \sum_{n=-\infty}^{\infty} J_n(\beta)\left[\delta(f - f_c - n f_m) + \delta(f + f_c + n f_m)\right] \tag{5.22}$$

Exercise 5.3

Eq. 5.19(c) gives the Fourier transform of $e^{j\beta\sin(\omega_m t)}$. Show that

a) $e^{j\beta\cos(\omega_m t)} \longleftrightarrow \sum_{n=-\infty}^{\infty} (j)^n J_n(\beta)\delta(f + n f_m)$

b) $A_c \cos\left[\omega_c t + \beta\cos(\omega_m t)\right] = A_c \sum_{n=-\infty}^{\infty} J_n(\beta)\cos\left[(\omega_c + n\omega_m)t + n\frac{\pi}{2}\right]$

Hint: $e^{j\beta\cos(\omega_m t)} = e^{j\beta\sin\left[\omega_m\left(\frac{\pi}{2\omega_m} - t\right)\right]}$

If $x(t) = e^{j\beta\sin\omega_m(-t)}$, then $e^{j\beta\cos(\omega_m t)} = x\left(t - \frac{\pi}{2\omega_m}\right)$.

Now use the Fourier transform properties for time reversal and time shift.

Properties of $J_n(\beta)$: The following properties of $J_n(\beta)$ can be established.

1) $J_n(\beta)$ is always *real* (For all n and β)

2) $J_n(\beta) = (-1)^n J_{-n}(\beta)$

3) For small values of β, $J_n(\beta) \simeq \frac{(\beta/2)^n}{n!}$

Hence, for small values of β,

$$J_0(\beta) \simeq 1$$

$$J_1(\beta) \simeq \frac{\beta}{2}$$

$$J_n(\beta) \simeq 0, n > 1$$

4) $$\sum_{n = -\infty}^{\infty} J_n^2(\beta) = 1$$

Fig 5.9 depicts the behavior of $J_n(\beta)$. We now make the following observations regarding the spectrum $(f > 0)$ of a tone modulated FM signal.

1) The spectrum of a tone modulated FM wave contains carrier component and an infinite set of side-frequencies located symmetrically on either side of the carrier at frequency separations of $n f_m, n = 1, 2, \cdots$

2) For small β, only $J_0(\beta)$ and $J_1(\beta)$ are significant. Then the FM spectrum has only three components: at $f_c, f_c \pm f_m$. This situation corresponds to the special case of NBFM.

3) The amplitude of the carrier component varies with β according to $J_0(\beta)$. Thus, in contrast to AM, this amplitude 'contains' part of the message information. Moreover, $J_0(\beta) = 0$ for some values $\beta(\beta = 2.4, 5.5, \text{etc.})$.

4) From Fig. 5.9(b), we find that $J_n(\beta)$ decays monotonically for $\frac{n}{\beta} > 1$ and that $|J_n(\beta)| << 1$ for $\left|\frac{n}{\beta}\right| >> 1$.

Fig, 5.10 gives the spectral plots obtained from a spectrum analyzer with $\beta \simeq 2.4$. As can be seen from the plot, the carrier component is very nearly zero (approximately 50 dB below the maximum value).

Fig 5.11(a) gives the theoretical plot of various spectral components for $\beta \simeq 4$. Fig. 5.11(b) is the spectral plot as observed on a spectrum analyzer. We see from this plot that the third spectral component has the largest magnitude which is an agreement with the theory. Let us normalize the dB plot with respect to the largest magnitude; that is, we treat the largest magnitude as 0 dB and compare the rest respect to with this value. For $n = 4$, Fig. 5.11(a), we have $J_4(\beta) = 0.281$. Hence $20 \log_{10} \dfrac{0.281}{0.43} = -3.7$ and value indicated by the spectrum analyzer agrees with this. Similarly, for $n = 5$, we expect $20 \log_{10} \dfrac{0.132}{0.430} = -10.25$; the value as observed on the plot of Fig. 5.11(b) is in close agreement with this. For $n = 1$, theoretical value is $20 \log \dfrac{0.066}{0.43} = -17.1$ dB. Spectrum analyzer display is in close agreement with this. The values of the remaining spectral components can be similarly be verified.

From Eq. 5.21, the average power P_{av} of the FM signal with tone modulation is $P_{av} = \dfrac{A_c^2}{2} \sum_{n = -\infty}^{\infty} J_n^2(\beta)$.

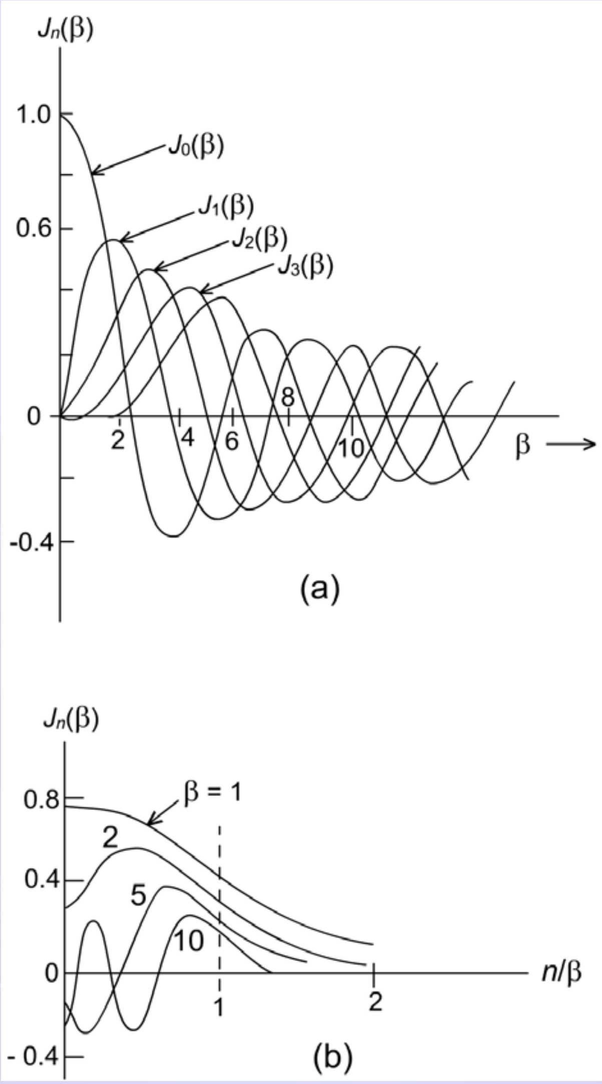


Fig. 5.9: Bessel Functions:

(a) $J_n(\beta)$ vs. β for fixed n

(b) $J_n(\beta)$ vs. $\dfrac{n}{\beta}$ for fixed β

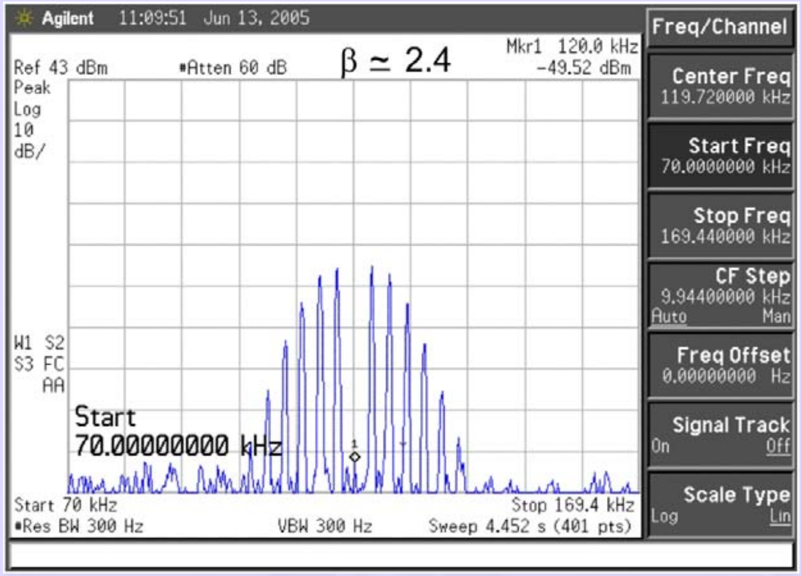


Fig. 5.10: Spectrum analyzer output with $\beta \simeq 2.4$

But from property 4 of $J_n(\beta)$, we have $\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1$, which implies that the P_{av} of the tone modulated FM signal is $\frac{A_c^2}{2}$. This result is true in general (whether $m(t)$ is a tone or not) because an FM signal is essentially a constant amplitude cosine signal with time varying frequency. As the RMS value of a sinusoid is independent of the frequency, we state that

$$(P_{av})_{FM} = \frac{A_c^2}{2} \tag{5.23}$$

For large β, if we assume that we can neglect $J_n(\beta)$ for $n > \beta + 2$, then

$$(B_T)_{FM} = 2 n f_m = 2(\beta + 2) f_m \tag{5.24a}$$

This in agreement with Eq. 5.15, with $k = 2$. If we neglect the values of $J_n(\beta)$ for $n > \beta + 1$, then,

$$(B_T)_{FM} = 2(\beta + 1) f_m \tag{5.24b}$$

which corresponds to Eq. 5.15 with $k = 1$.

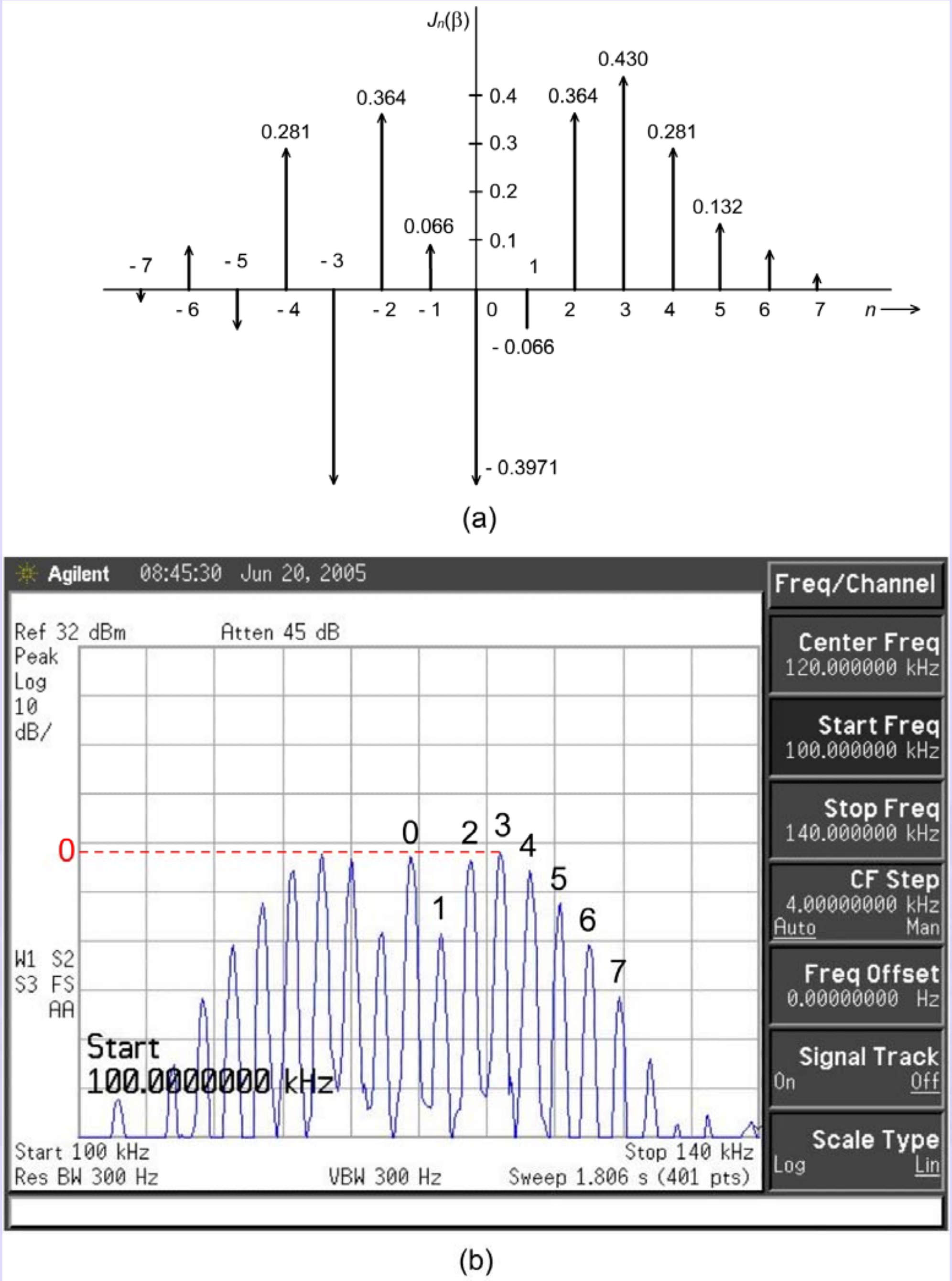


Fig. 5.11(a): $J_n(\beta)$ for $\beta = 4$

(b) Spectrum analyzer output for $\beta \simeq 4$

Consider $\beta = 5$. From Appendix A5.1, we have $J_7(5) = 0.053$ and $J_8(5) = 0.018$. P_1, the average power of these two is,

$$P_1 = \frac{A_c^2}{2}(0.0028 + 0.0003) = \frac{A_c^2}{2}(0.0031).$$

Hence, $\frac{P_1}{(P_{av})_{FM}} = 0.0031 = 0.31\%$. As $J_n(5)$ for $n \geq 9$ are much smaller than $J_8(5)$, we see that practically the entire power of the FM signal is confined to the frequency range $[f_c \pm 6 f_m]$. As such, we find that even Eq. 5.24(b) is a fairly good measure of $(B_T)_{FM}$ for tone modulation.

With tone modulation, it is possible to estimate the value of B_T to the accuracy desired. In this case, the magnitude of a spectral component depends only on $J_n(\beta)$ and these Bessel coefficients have been tabulated extensively. For a given β, let us define transmission bandwidth such that it includes all those spectral components whose magnitude is greater than *one percent of the unmodulated carrier amplitude*. If we take $A_c = 1$, then

$$B_T = 2 n_{sig} f_m \tag{5.24c}$$

where n_{sig} is such that $|J_n(\beta)| \geq 0.01$ for $n \leq n_{sig}$. For example, if $\beta = 20$,then n_{sig} would be 25. (For $\beta = 10, 5,$ etc., n_{sig} can be found from the table in Appendix A5.1.) Let $f_m = f_0$ when $\beta = 20$. Let us tabulate B_T for a few other values of β. We will keep $\Delta f = 20 f_0$, but reduce the value of β by increasing f_m. These are listed in Table 5.1. Also listed in the table are $B_{T,1}$ and $B_{T,2}$ where $B_{T,1} = 2(\beta + 1) f_m$ (Carson's rule), and

$$B_{T,2} = 2(\beta + 2) f_m$$

f_m	β	$2n_{sig}$	B_T	$B_{T,1}$	$B_{T,2}$
f_0	20	50	$50\ f_0$	$42\ f_0$	$44\ f_0$
$2\ f_0$	10	28	$56\ f_0$	$44\ f_0$	$48\ f_0$
$4\ f_0$	5	16	$64\ f_0$	$48\ f_0$	$56\ f_0$
$10\ f_0$	2	8	$80\ f_0$	$60\ f_0$	$80\ f_0$
$20\ f_0$	1	6	$120\ f_0$	$80\ f_0$	$120\ f_0$
$40\ f_0$	0.5	4	$160\ f_0$	$120\ f_0$	$200\ f_0$
$60\ f_0$	0.33	4	$240\ f_0$	$160\ f_0$	$280\ f_0$
$200\ f_0$	0.1	2	$400\ f_0$	$440\ f_0$	$840\ f_0$

Table 5.1: B_T of Eq. 5.24c for various values of β with $\Delta f = 20 f_0$

From the above table, we see that

i) For small values of β (less than or equal to 0.5), Δf becomes less and less significant in the calculation of B_T. For $\beta = 0.1$, we find that

$$\frac{2\Delta f}{B_T} = \frac{40 f_0}{400 f_0} = \frac{1}{10}$$

ii) For small values of β, bandwidth is essentially decided by the highest frequency component in the input spectrum. As such, as $\beta \to 0$, B_T does not go zero. In fact, in absolute terms, it increases. (From the table, with $\beta = 20$, we require a B_T of $50 f_0$ where as with $\beta = 0.1$, B_T required is $400 f_0$.)

iii) $B_{T,1}$, which is based on Carson's rule is in close agreement with B_T only for very small values of β. Otherwise, $B_{T,2}$ is better.

It is interesting to note that for $\beta =$ 5, 10 and 20, $B_{T,2}$ (which is generally considered to overestimate the bandwidth requirement) is less than B_T as given by Eq. 5.24(c). This is because B_T of Eq. 5.24(c) neglects only those frequency components whose magnitude is less than or equal to $0.01\,A_c$. In other words, the ratio

$$\frac{\text{Power of any spectral component not included in } B_T}{P_{av}}$$

is less than $\dfrac{(0.01\,A_c)^2 / 2}{A_c^2 / 2} = 10^{-4}$

With $\beta = 5$, $B_{T,2}$ takes into account only upto $J_7(5)$ and $\dfrac{(J_8(5)A_c)^2 / 2}{\left(A_c^2 / 2\right)} = 3.38 \times 10^{-4}$. That is, B_T as given by Eq. 5.24(c) includes few more spectral components in the bandwidth than that taken into account by $B_{T,2}$ resulting in a larger value for the bandwidth.

It is instructive to analyze the FM signal when $m(t)$ is a sum of sinusoids. This has been done in the Appendix A5.3. Another measure of bandwidth that is useful in the study of frequency modulation is the **rms bandwidth**. This has been dealt with in Appendix A5.4.

5.4 Phase Modulation

All the results derived for FM can be directly applied to PM. We know that, for phase modulation,

$$f_i(t) = \frac{1}{2\pi}\frac{d\,\theta_i(t)}{d\,t} = f_c + \frac{k_p}{2\pi}\,m'(t) \tag{5.25a}$$

where $m'(t)$ is the derivative of $m(t)$. Therefore,

$$\Delta f = \frac{k_p}{2\pi} m_p', \text{ where } m_p' = \left| m'(t) \right|_{\max} \tag{5.25b}$$

$$(B_T)_{PM} = 2[\Delta f + kW], \; 1 < k < 2$$

$$= 2\left[\frac{k_p m_p^{'}}{2\pi} + kW\right] \quad (5.26)$$

With respect to Δf, there is one important difference between PM and FM. For FM, as $\Delta f = k_f m_p$, it depends only on the peak value of the modulating signal. Hence the bandwidth of WBFM is essentially independent of the spectrum of $m(t)$ (It has a weak dependence on W). On the other hand, for PM, Δf is a function of $m_p^{'}$ and this value is strongly dependent on $M(f)$. Strong high frequency components in $M(f)$ would result in large values of $m_p^{'}$. Similarly, predominance of lower-frequency components will result in a lower value of $m_p^{'}$. As a result, for signals with strong high frequency components, $(B_T)_{PM}$ will be much larger than for the case of the signals with strong spectral components at lower frequencies.

Example 5.5

To determine the frequency sensitivity of an FM source, the following method has been used. The FM generator is fed a tone modulating signal $A_m\left[\cos\left(4\pi\times 10^3\right)t\right]$. Starting near about zero, A_m is gradually increased and when $A_m = 2$ V, it has been found that the carrier component goes to zero for the first time. What is the frequency sensitivity of the source? Keeping A_m at 2 V, frequency f_m is decreased until the carrier component goes zero for the second time. What is the value of f_m for this to happen?

With tone modulation, we know that carrier goes to zero for the first time when $\beta = 2.4$. (That is, the smallest value of β for which $J_0(\beta) = 0$ is $\beta = 2.4$.)

But $\beta = \dfrac{\Delta f}{f_m} = 2.4$. That is,

$$\Delta f = 2.4\, f_m = 4.8 \text{ kHz} = k_f\, m_p$$

As m_p is 2 V, we have $k_f = \dfrac{4.8}{2} = 2.4$ kHz/V

The carrier component is zero for the second time when $\beta = 5.5$. Hence,

$$5.5 = \frac{4.8 \times 10^3}{f_m}$$

or $$f_m = \frac{4.8 \times 10^3}{5.5} = 872 \text{ Hz}$$ ♦

Example 5.6

A 1.0 kHz tone is used to generate both an AM and an FM signal. Unmodulated carrier amplitude is the same for both AM and FM. The modulation index β of FM is 8. If the frequency components at $(f_c \pm 1000)$ Hz have the same magnitude in AM as well as FM, find the modulation index of AM.

For AM, the magnitude of the spectral components at $(f_c \pm 1000)$ Hz is $\dfrac{A_c\,\mu}{2}$. For FM, the magnitude of the spectral components at $(f_c \pm 1000)$ Hz is $A_c\, J_1(8)$.

$$\frac{A_c\,\mu}{2} = A_c\, J_1(8) \Rightarrow \mu = 2\, J_1(8)$$

$$= 0.47$$ ♦

Example 5.7

In this example we will show that the bandwidth of the PM signal is strongly dependent on $M(f)$ whereas this dependence is weak in the case of an FM signal.

Let $m(t) = \cos(2\pi f_1 t) + 0.85\cos(2\pi f_2 t)$ be used to generate an FM as well as a PM signal. Modulator constants are $k_f = 5$ kHz/V and $k_p = 10$ rad/V. Let us compute (a) $(B_T)_{FM}$ and (b) $(B_T)_{PM}$ for the following combinations of f_1 and f_2.

i) $f_1 = 500$ Hz and $f_2 = 700$ Hz

ii) $f_1 = 1000$ Hz and $f_2 = 1400$ Hz

We will assume that $J_n(\beta)$ can be neglected if $n > \beta + 2$.

a) $(B_T)_{FM}$

Case i) $f_1 = 500$ Hz and $f_2 = 700$ Hz

From Eq. A5.3.1 (Appendix 5.3), we have

$$s(t) = A_c \sum_m \sum_n J_m(\beta_1) J_n(\beta_2) \cos\left[(\omega_c + m\omega_1 + n\omega_2)t\right]$$

$$\beta_1 = \frac{5\times 10^3}{500} = 10$$

$$\beta_2 = \frac{5\times 10^3 \times 0.85}{700} \simeq 6$$

We shall take into account $J_m(\beta_1)$ upto $m = 12\ (= \beta_1 + 2)$ and $J_n(\beta_2)$ upto $n = 8\ (= \beta_2 + 2)$. Then,

$$(B_T)_{FM} = 2(12\times 500 + 8\times 700)$$

$$= 23.2\times 10^3 \text{ Hz}$$

In the equation for $s(t)$ above, we see that the magnitude of any spectral component depends on the product. $J_m(\beta_1) J_n(\beta_2)$. Let us calculate the ratio $\dfrac{J_{13}(10) J_8(6)}{\left[J_m(10) J_n(6)\right]_{\max}}$ so as to get an idea of the magnitude of the spectral components that have been neglected in calculating $(B_T)_{FM}$.

Let $A = J_{13}(10) J_9(6) = 0.03 \times 0.02$

$= 0.0006$

From the tables (Appendix A5.1), we find that the maximum value of the product $J_m(10) J_n(6)$ occurs for $m = 8$ and $n = 5$. Let,

$B = J_8(10) J_5(6) = 0.318 \times 0.362$

$= 0.115$,

Then, $\dfrac{A}{B} = 0.0052$.

Case ii) $f_1 = 1000$ Hz and $f_2 = 1400$ Hz

Now, we have $\beta_1 = 5$ and $\beta_2 \simeq 3$.

If we account for $J_m(\beta_1)$ and $J_n(\beta_2)$ upto $m = 7$ and $n = 5$, then

$$(B_T)_{FM} = 2\left(7 \times 10^3 + 5 \times 1400\right)$$

$$= 28 \times 10^3 \text{ Hz}$$

Of course the maximum value of $J_m(\beta_1)$ and $J_n(\beta_2)$ occurs for $m = 4$ and $n = 2$, and this product is

$C = J_4(5) J_2(3) = 0.391 \times 0.486 = 0.19$

However, $D = J_8(5) J_6(3) = 0.018 \times 0.0114$

$= 0.0002$

Hence $\frac{D}{C} = \frac{0.0002}{0.19} \simeq 0.001$

This is much less than (A/B) which implies that we have taken into account more number of spectral components than case (i) above. If we restrict ourselves to the number of spectral components upto $m = \beta + 1$, then,

$$\frac{J_7(5)\,J_6(3)}{J_4(5)\,J_2(3)} = 0.003 \text{ which is closer to } (A/B) \text{ than } 0.001.$$

Then,

$$(B_T)_{FM} = 2(6000 + 5 \times 1400) = 26 \times 10^3 \text{ Hz}$$

which is fairly close to the previous value of 23.2×10^3 Hz.

b) $(B_T)_{PM}$

$$m^{'}(t) = -\left[2\pi f_1 \sin(2\pi f_1 t) + 0.85\,(2\pi f_2)\sin(2\pi f_2 t)\right]$$

Now $m^{'}(t)$ frequency modulates the carrier

$$\beta_1 = \frac{k_p}{2\pi} \cdot \frac{2\pi f_1}{f_1} = k_p = 10$$

Similarly,

$$\beta_2 = 0.85 k_p = 8.5$$

Case i) $f_1 = 500$ Hz and $f_2 = 700$ Hz

$$(B_T)_{PM} = 2(12 \times 500 + 10 \times 700) = 26 \times 10^3 \text{ Hz}$$

Case ii) $f_1 = 1000$ Hz and $f_2 = 1400$ Hz

As β_1 and β_2 remains unchanged, we have

$$(B_T)_{PM} = 2\left(12 \times 10^3 + 10 \times 1400\right)$$

$$= 52 \times 10^3 \text{ Hz}$$

This is twice $(B_T)_{PM}$ of case (i). This example shows that the bandwidth of the PM signals is strongly dependent on $M(f)$. ♦

Exercise 5.4

Consider the scheme shown in Fig. 5.12.

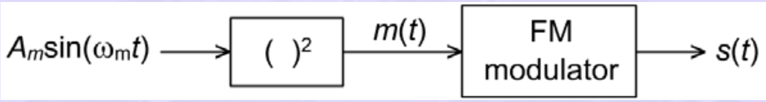


Fig. 5.12: Scheme for the Exercise 5.4

a) Let f_c be the frequency of the modulator when $m(t) = 0$. If $s(t)$ can be expressed in the form

$$s(t) = A_c \cos\left[\omega'_c\, t - \beta \sin(2\omega_m t)\right],$$

what are the values of $f_c^{'}$ and β?

b) What are the frequency components in the spectrum of $s(t)$?

Ans: $f_c^{'} = \left(f_c + \dfrac{k_f A_m^2}{2}\right)$

$$\beta = \frac{k_f A_m^2}{2 f_m}$$

5.5 Generation of FM

We had earlier identified two different categories of FM, namely, NBFM and WBFM. We shall now present the schemes for their generation.

5.5.1 Narrowband FM

One of the principal applications of NBFM is in the (indirect) generation of WBFM as explained later on in this section. From the approximation (5.10) we have,

$$s(t) \simeq A_c \left[\cos(\omega_c t) - c_f\, m_I(t) \sin(\omega_c t)\right]$$

The system shown in Fig. 5.13 can be used to generate the NBFM signal.

Applying $m(t)$ directly to the balanced modulator, results in NBPM with a suitable value for k_p.

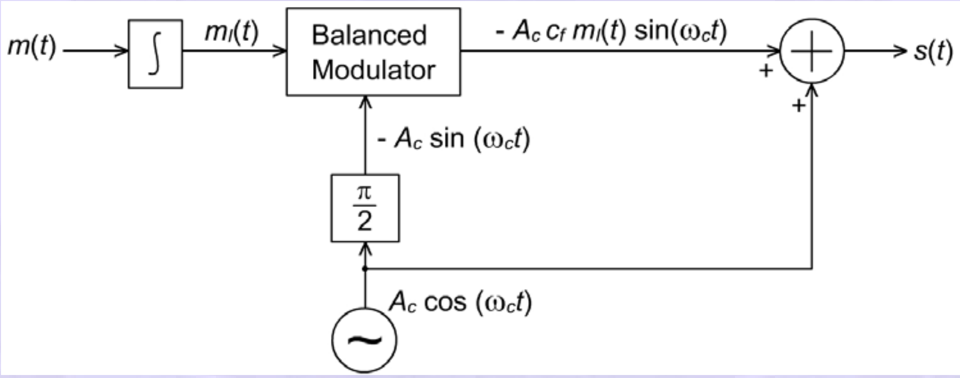


Fig. 5.13: Generation of NBFM signal

5.5.2 WBFM: Indirect and direct methods

There are two distinct methods of generating WBFM signals: a) Direct FM b) Indirect FM. Details on their generation are as follows.

a) Indirect FM (Armstrong's method)

In this method - attributed to Armstrong - first a narrowband FM signal is generated. This is then converted to WBFM by using frequency multiplication. This is shown schematically in Fig. 5.14.

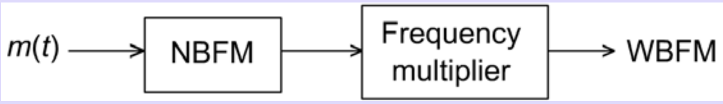


Fig. 5.14: Generation of WBFM (Armstrong method)

The generation of NBFM has already been described. A frequency multiplier is a nonlinear device followed by a BPF. A nonlinearity of order n can give rise to frequency multiplication by a factor of n. For simplicity, consider a square law device with output $y(t) = x^2(t)$ where $x(t)$ is the input. Let $x(t)$ be the FM signal given by,

$$x(t) = \cos\left[\theta(t)\right], \text{ where } \theta(t) = \omega_c t + 2\pi k_f \int_{-\infty}^{t} m(\alpha)\, d\alpha$$

Note that we have dropped the subscript i from $\theta_i(t)$. Then,

$$\begin{aligned} y(t) &= \cos^2\left[\theta(t)\right] \\ &= \frac{1}{2}\left\{1 + \cos\left[2\theta(t)\right]\right\} \\ &= \frac{1}{2} + \frac{1}{2}\cos\left[2\omega_c t + 4\pi k_f \int_{-\infty}^{t} m(\alpha)\, d\alpha\right] \end{aligned} \quad (5.27)$$

The DC term in Eq. 5.27 can be filtered out to give an FM output with the carrier frequency $2f_c$ and frequency deviation twice that of the input FM signal. An input-output relation of the type $y(t) = a_1 x(t) + a_2 x^2(t) + \ldots . a_n x^n(t)$ will give rise to FM output components at the frequencies $f_c, 2f_c, \ldots ., n f_c$ with the corresponding frequency deviations $\Delta f, 2\Delta f, \ldots ., n\Delta f$, where Δf is the frequency deviation of the input NBFM signal. The required WBFM signal can be obtained by a suitable BPF. If necessary, frequency multiplication can be resorted to in more than one stage. The multiplier scheme used in a commercial FM transmitter is indicated in Fig. 5.15.

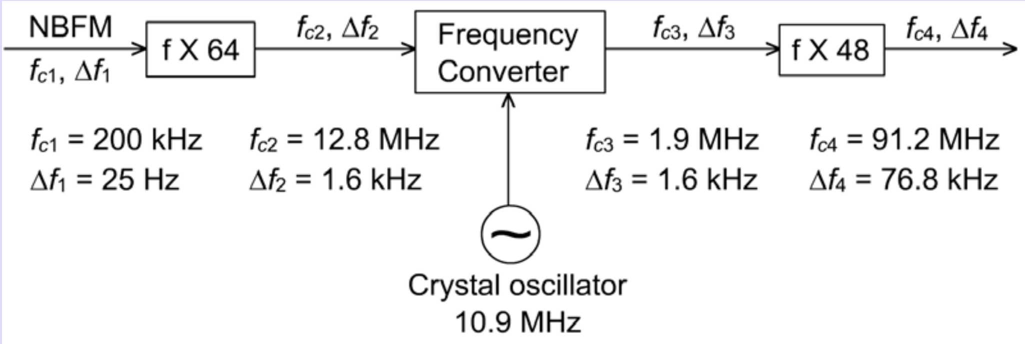


Fig. 5.15: Multiplier chain used in typical commercial FM transmitter

The carrier frequency of the NBFM signal f_{c1}, is 200 kHz with the corresponding $\Delta f_1 = 25$ Hz. Desired FM output is to have the frequency deviation $\Delta f_4 \simeq 75$ kHz and a carrier $\left(f_{c4}\right)$ of 91.2 MHz.

To obtain $\Delta f_4 = 75$ kHz starting from $\Delta f_1 = 25$ Hz, we require a total frequency multiplication of $\frac{\left(75 \times 10^3\right)}{25} = 3000$. In the scheme of Fig. 5.11, this has been accomplished in two stages, namely, multiplication by 64 followed by multiplication by 48, giving a total multiplication by the factor $64 \times 48 = 3072$. (Actually each stage of multiplication is implemented by a cascade of frequency doublers or triplers. Thus multiplication by 64 is obtained by 6 doublers in cascade where as multiplication by 48 is implemented by a cascade of a frequency tripler and 4 doublers.) Multiplication of $f_{c1} = 200$ kHz by 3072 gives a carrier frequency $f_{c4} = 614.4$ MHz. As the final required carrier frequency is 91.2 MHz, a frequency conversion stage is used to down convert f_{c2} (12.8 MHz) to f_{c3} (1.9 MHz). In this process of down conversion, frequency deviation is unaffected $\left(\Delta f_2 = \Delta f_3 = 1.6 \text{ kHz}\right)$. The possible drawbacks of this scheme are the introduction of noise in the process of multiplication and distortion in the generation of NBFM signal especially for low modulating frequencies as $\frac{\Delta f}{f_m}$ could become excessive.

Example 5.8

Armstrong's method is to be used to generate a WBFM signal. The NBFM signal has the carrier frequency $f_{c1} = 20$ kHz. The WBFM signal that is required must have the parameters $f_c = 6$ MHz and $\Delta f = 10$ kHz. Only frequency triplers are available. However, they have a limitation: they cannot produce frequency components beyond 8 MHz at their output. Is frequency conversion stage required? If so, when does it become essential? Draw the schematic block diagram of this example.

Total frequency multiplication required $= \frac{6 \times 10^6}{20 \times 10^3} = 300$. Using only frequency triplers, we have $3^5 = 243$ and $3^6 = 729$. Hence a set of six multipliers is required. But these six cannot be used as a single cascade because, that would result in a carrier frequency equal to $20 \times 10^3 \times 3^6 = 14.58$ MHz and the last multiplier cannot produce this output. However, cascade of 5 triplers can be used. After this, a frequency conversion stage is needed.

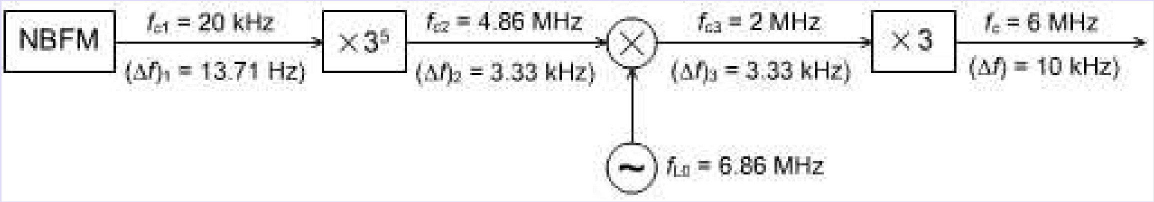


Fig 5.16: Generation of WBFM from NBFM of example 5.8

Block diagram of this generation scheme is shown in Fig. 5.16. As the final frequency deviation required is 10 kHz, the NBFM must have a

$(\Delta f)_1 = \frac{10 \times 10^3}{729} = 13.71$ Hz. After the frequency conversion stage, we have one more stage multiplication by 3. If f_{c3} is the carrier frequency at the mixer output, then $f_{c3} \times 3 = 6$ MHz $\Rightarrow f_{c3} = 2$ MHz. Assuming that f_{Lo} is greater than incoming carrier frequency of 4.86 MHz, we require $f_{Lo} = 6.86$ MHz so that the difference frequency component is 2 MHz. ♦

Exercise 5.5

In the indirect FM scheme shown in Fig. 5.17, find the values of $f_{c,i}$ and Δf_i for $i = 1, 2$ and 3. What should be the centre frequency, f_0, of the BPF. Assume that $f_{LO} > f_{c,2}$.

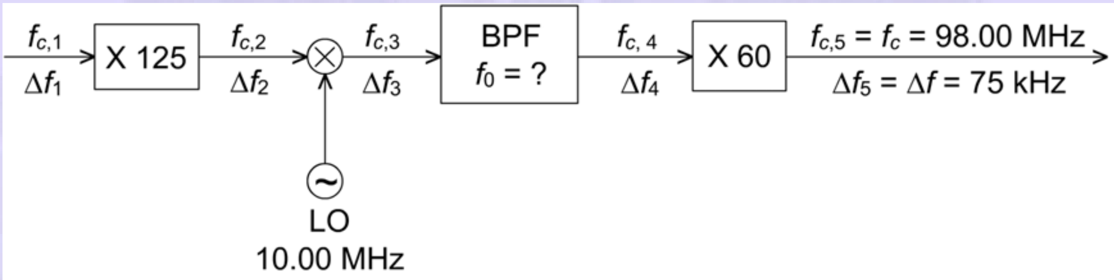


Fig. 5.17: Scheme for the Exercise 5.5

Exercise 5.6

In the indirect FM scheme shown in Fig. 5.18, find the values of the quantities with a question mark. Assume that only frequency doublers are available. It is required that $f_{LO} < f_{c,2}$.

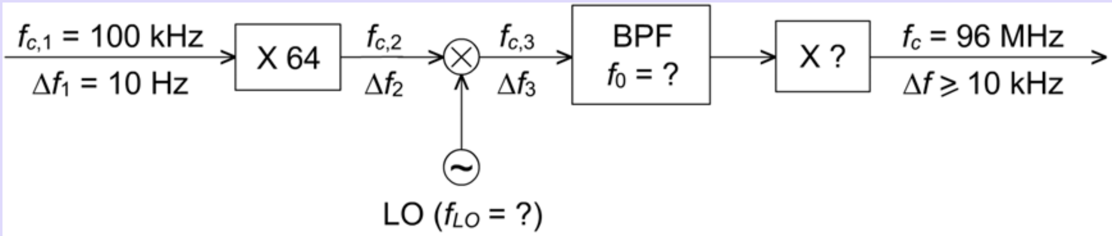


Fig. 5.18: Scheme for the Exercise 5.6

b) Direct FM

Conceptually, the direct generation of FM is quite simple and any such system can be called as a *Voltage Controlled Oscillator* (VCO). In a VCO, the oscillation frequency varies linearly with the control voltage. We can generate the required FM signal from a VCO by using the message signal $m(t)$ in the control voltage. We shall explain below two methods of constructing a VCO.

One can construct a VCO by using an integrator and hysteric comparator (such as Schmitt trigger circuit). Method 1 below gives the details of this scheme. Another way of realizing a VCO is to vary the reactive component of a tuned circuit oscillator such as Colpitt's oscillator or Hartley oscillator. Method 2 below gives the details of this scheme.

Method 1: Consider the scheme shown in Fig. 5.19(a). The Schmitt trigger output $y(t)$ is $+V_0$ when the integrator output $x(t)$ is increasing, and $-V_0$ when $x(t)$ is decreasing. Further, $y(t)$ changes from $+V_0$ to $-V_0$ when $x(t)$ reaches $+E_0$ and from $-V_0$ to $+V_0$ when $x(t)$ reaches $-E_0$. The relationship between $x(t)$ and $y(t)$ is shown in Fig. 5.19(b). The electronic switch is designed such that it is in position 1 when $y(t) = V_0$ and goes to position 2 when $y(t) = -V_0$. Let us first examine the case when the input $v_m(t)$ is a positive constant v_0. Consider the situation at $t = 0$ when $y(t)$ has just switched to V_0 and the switch going from position 2 to position 1. At this point, $x(t)$ has attained the value $-E_0$. Then, for $0 < t \le t_1$, we have

$$x(t) = -E_0 + \frac{1}{RC}\int_0^t v_0 \, d\tau$$

When $t = t_1$, let $x(t)$ become E_0. Then $y(t)$ goes to $-V_0$ and the electronic switch assumes position 2. The value of t_1 can be obtained from

$$E_0 = -E_0 + \frac{1}{RC} \int_0^{t_1} v_0 \, d\tau$$

or $$t_1 = \frac{2RCE_0}{v_0}$$

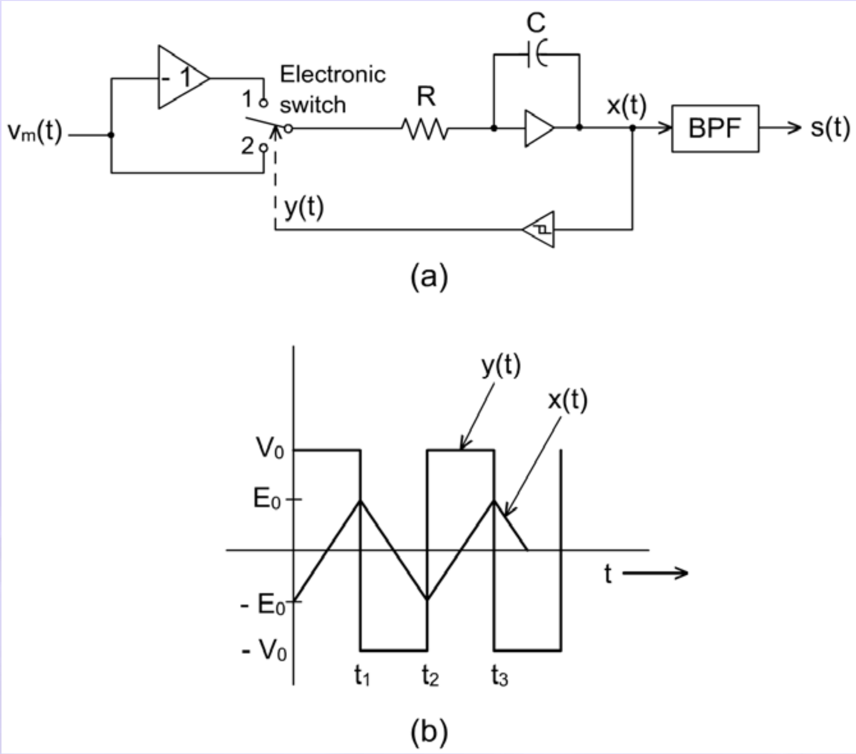


Fig. 5.19: Direct FM generation (method 1)

The output $x(t)$ now keeps decreasing until $t = t_2$ when $x(t) = -E_0$. It is easy to see that $(t_2 - t_1) = \frac{2RCE_0}{v_0}$. That is, $x(t)$ and $y(t)$ are periodic with period $\frac{4RCE_0}{v_0}$ or the fundamental frequency of these waveforms is $f_0 = \left[\frac{4RCE_0}{v_0}\right]^{-1}$.

Note that f_0 depends on the input signal v_0. If $v_0 = E_0$ then $f_0 = \frac{1}{4RC}$. By properly choosing the values of R and C, we can have $f_0 = f_c = \frac{1}{4RC}$, where

f_c is the required carrier frequency. If the triangular wave $x(t)$ is the input to a narrowband BPF tuned to f_c, the output $s(t)$ is $A_c \cos(2\pi f_c t)$. Then, by making $v_m(t) = v_0 + k\,m(t)$ where k is a constant such that $k|m(t)| < v_0$, we can have the instantaneous frequency of $s(t)$ as $f_c + k_f\,m(t)$, which is the desired result.

Method 2: The oscillation frequency f_0 of a parallel tuned circuit with inductance L and capacitance C is given by

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \text{ or } \omega_0 = \frac{1}{\sqrt{LC}}$$

Let C be varied by the modulating signal $m(t)$, as given by

$$C(t) = C_0 - k\,m(t)$$

where k is an appropriate constant.

Then $$\omega_i(t) = \frac{1}{\sqrt{LC_0\left[1 - \frac{k\,m(t)}{C_0}\right]}}$$

$$\simeq \frac{1}{\sqrt{LC_0}}\left[1 + \frac{k\,m(t)}{2C_0}\right] \quad \text{when } \frac{k|m(t)|}{C_0} << 1$$

Let $\omega_c = \dfrac{1}{\sqrt{LC_0}}$. Then,

$$\omega_i(t) = \omega_c + c_f\,m(t), \text{ where } c_f = \frac{k\,\omega_c}{2C_0}.$$

One of the more recent devices for obtaining electronically variable capacitance is the *varactor* (also called *varicap*, or *voltacap*). In very simple terms, the varactor is a junction diode. Though all junction diodes have inherent junction capacitance, varactor diodes are designed and fabricated such that the value of the junction capacitance is significant (varactors are available with nominal ratings from 0.1 to 2000 pF). Varactor diodes, when used as voltage-variable capacitors are reverse biased and the capacitance of the junction varies inversely with the applied (reverse) voltage.

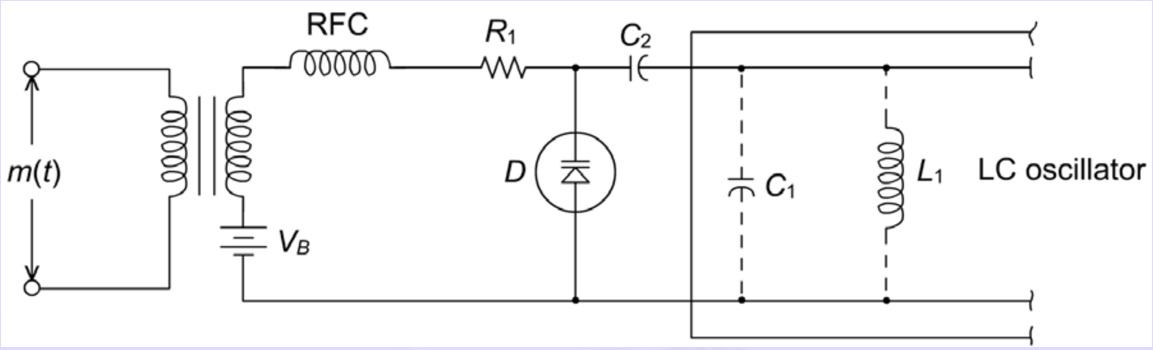


Fig. 5.20: Direct FM generation (method 2)

Consider the scheme shown in Fig. 5.20. The frequency of the oscillator output depends on L_1 and C_{eq} where $C_{eq} = C_1 \,||\, C_d$, C_d being the capacitance of the varactor D. V_B reverse biases D such that when $m(t) = 0$, C_{eq} is of the correct value to result in output frequency f_c. When the message signal $m(t)$ is on, C_d can be taken as, $C_d = C_0^{'} - k\,m(t)$ where $C_0^{'}$ is the value of C_d, with $m(t) = 0$. Hence $C(t)$ of the oscillator circuit is,

$$C(t) = \left(C_1 + C_0^{'}\right) - k\,m(t) = C_0 - k\,m(t), \text{ where } C_0 = C_1 + C_0^{'}$$

The other components in Fig. 5.14 have the following functions. C_2 is a blocking capacitor to isolate the DC level of the varactor from the rest of the modulator circuit. Note that $C_2 >> C_d$ or C_1 and is essentially a short at the required FM frequencies. RFC is an RF choke that prevents the RF energy of the oscillator circuit from feeding into the audio transformer. R_1 limits the current in the circuit in the event that the peaks of the audio signal voltage exceed the voltage of the DC source and momentarily forward bias the diode D.

Direct FM generation can produce sufficient frequency deviation and requires little frequency multiplication. However, it suffers from the carrier frequency instability as the carrier frequency is not obtained from a highly stable oscillator. Some additional circuitry is required to achieve frequency stability.

Voltage controlled oscillators are available in the form of IC chips. For example XR-2206 of the EXAR Corporation, USA has an operating frequency range of 0.01 Hz to 1 MHz. Another chip with an operating frequency range of about 1 MHz is LM 566. Texas Instruments CD 4046 is an inexpensive VCO chip with a normal operating frequency range upto 1.4 MHz. (Actually CD 4046 is a PLL and VCO is a part of it. PLLs are discussed in sec. 5.5.4.) To make the VCO chip functional, what is required is an external capacitor and one or two external resistors. CD 74HC7046 is another PLL chip with VCO. Free running frequency of the VCO is 18 MHz. Koster, Waldow and Ingo Wolf describe a VCO operating in the frequency range 100 MHz to 4 GHz [1]. Donald Tillman describes a new VCO design (called a quadrature trapezoid VCO) especially suited for electronic music applications. Details are available at:
http://www.till.com/articles/QuadTrapVCO/discussion.html.

Exercise 5.7

Consider the circuit of Fig. 5.20 for the direct generation of FM. The diode capacitance C_d, is related to the reverse bias as,

$$C_d = \frac{100}{\sqrt{1 + 2 v_d}} \text{ pF}$$

where v_d is the voltage across the varactor. Let $V_B = 4$ V and $m(t) = 0.054 \sin\left[\left(10\pi \times 10^3\right) t\right]$. It is given that $C_1 = 250$ pF and the circuit resonates at 2 MHz when $m(t) = 0$.

a) Show that, by using the binomial approximation, C_d can be put in the form

$$C_d = \frac{10^{-10}}{3} - 0.2 \times 10^{-12} \sin\left[\left(10\pi \times 10^3\right) t\right]$$

b) Show that

$$f_i(t) = \left(2 \times 10^6\right) + 705.86 \sin\left[\left(10\pi \times 10^3\right) t\right]$$

(Note that C_2 is a short circuit at the frequencies being generated.)

c) Let A_c be the amplitude of oscillations of the VCO. Write the expression for the generated FM signal.

Exercise 5.8

A simple variation of the circuit of Fig. 5.20 is shown in Fig. 5.21.

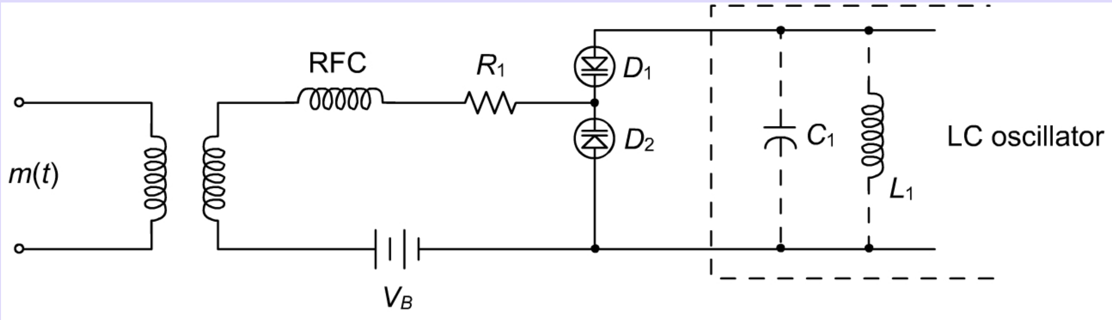


Fig. 5.21: A variation of the VCO circuit of Fig. 5.14

The varactor diodes D_1 and D_2 are connected back-to-back. This arrangement helps to mitigate the effect of the RF signal of the tuned circuit (also called tank circuit) driving a single diode into conduction on its peaks which will change the bias voltage (and thereby the frequency that is generated). Of course, as far as the tuned circuit is concerned, the two diodes are in series; this means the capacitance of the combination is one-half that of a single diode.

Let $V_B = 5$ V, $m(t) = 0.5\cos\left(2\pi \times 10^3 t\right)$. Calculate the carrier frequency f_c, $(f_i)_{max}$ and $(f_i)_{min}$ from the above scheme, given that $C_1 = 50$ pF, $L_1 = 50\ \mu$H and the capacitance of the varactor diode C_d, follows the relation

$$C_d = \frac{50}{\sqrt{v_d}} \text{ pF}$$

where v_d is the voltage across the diode. Do not use any approximations. What are the values of $(f_i)_{max}$ and $(f_i)_{min}$, if you use binomial approximation.

Ans: $f_c \simeq 91$ kHz, $(f_i)_{max} \simeq 91.4$ kHz and $(f_i)_{min} \simeq 90.6$ kHz, without any approximation.

5.6 Demodulation of FM

A variety of techniques and circuits have been developed for demodulating FM signals. We shall consider a few of these techniques falling under the following categories:

1) FM-to-AM conversion
2) Phase shift discrimination
3) Zero crossing detection
4) Phase Locked Loop (PLL)

5.6.1 FM-to-AM conversion

The instantaneous frequency of an FM signal is given by $f_i = f_c + k_f\, m(t)$. Hence a frequency selective network with a transfer function of the from $|H(f)| = \alpha f + \beta$, ($f > 0$, and α and β are constants) over the FM band would yield an output proportional to the instantaneous frequency. That is, the circuit converts the frequency deviation into a corresponding amplitude change, which in this case is proportional to $m(t)$, the message signal. It is assumed that the time constant of the network is small enough in comparison with the variations in the instantaneous frequency of the FM signal. We shall now indicate three ways of implementing this demodulation scheme.

Consider the scheme shown in Fig. 5.22 where $\frac{d}{dt}$ represents a band-pass differentiator with the magnitude characteristic $|H(f)| = \alpha f + \beta$, (for $f > 0$), over the required bandwidth. BPL[1] is a Band-Pass Limiter which eliminates amplitude fluctuations from the received FM signal.

[1] Band-pass limiters have been analyzed in section 5.7.

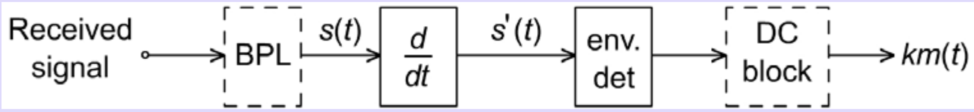


Fig. 5.22: Schematic of an FM demodulator based on FM - to - AM conversion

$s(t)$ is the constant amplitude FM signal to be demodulated. $s^{'}(t)$, the output of the differentiator, is given by

$$s^{'}(t) = -A_c\left[\omega_c + 2\pi k_f m(t)\right] \sin\left[\omega_c t + 2\pi k_f \int_{-\infty}^{t} m(\alpha)\, d\alpha\right] \tag{5.28}$$

Eq. 5.28 represents a signal that is both amplitude and frequency modulated. The envelope of $s^{'}(t) = A_c\left[\omega_c + 2\pi k_f m(t)\right]$ (we assume that $\Delta f = k_f m_p \le f_c$; hence, $\left[\omega_c + 2\pi k_f m(t)\right] \ge 0$). As $A_c \omega_c$ represents a DC term, signal $m(t)$ can be obtained from $s^{'}(t)$, after the DC-block.

The need for a BPL is as follows. Assume that the received FM signal (with amplitude fluctuations) is applied directly as the input to the differentiator. Let $A_c(t)$ denote the envelope of the FM signal. Then, there would be an additional term, $\dfrac{d A_c(t)}{d t}$ on the RHS of Eq. 5.28. Even if this term were to be neglected, the envelope of $s^{'}(t)$ would be $A_c(t)\left[\omega_c + 2\pi k_f m(t)\right]$, which implies that the envelope of $s^{'}(t)$ does not contain a term proportional to $m(t)$. Therefore, it is essential to maintain the FM envelope at a constant level. (Several factors such as channel noise, fading etc. cause variations in A_c). Band-pass limiter eliminates the amplitude fluctuations, giving rise to an FM signal with constant envelope.

We shall now indicate some schemes to implement this *method of demodulation.*

Scheme 1: Circuit implementation of the scheme of Fig. 5.22, can be carried out by constructing an op-amp differentiator circuit followed by an envelope detector.

Scheme 2 (Slope detection): Another way of implementing the FM-to-AM conversion scheme is through a simple tuned circuit followed by an envelope detector. The transfer characteristic of a tuned circuit in a small region off resonance is approximately linear.

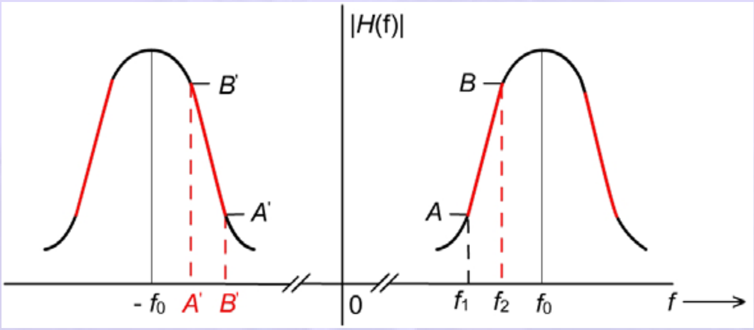


Fig. 5.23: Magnitude characteristics of a tuned circuit

In Fig. 5.23, we have shown the resonance characteristic of a tuned circuit. The parts of the characteristic shown in red have been drawn as straight lines. (This is a good approximation to the actual characteristic.) Assuming that $m(t) > 0$ produces an increase in the instantaneous frequency, we can use the straight line segment between A and B (for $f > 0$; for $f < 0$, we have the corresponding segment, $A^{'}$ to $B^{'}$) for demodulation purposes. Let us assume that the FM signal to be demodulated has the carrier frequency $f_c = \frac{f_1 + f_2}{2}$ and $f_1 \leq f_c - \frac{B_T}{2}$ where as $f_2 \geq f_c + \frac{B_T}{2}$. This is shown in Fig. 5.24.

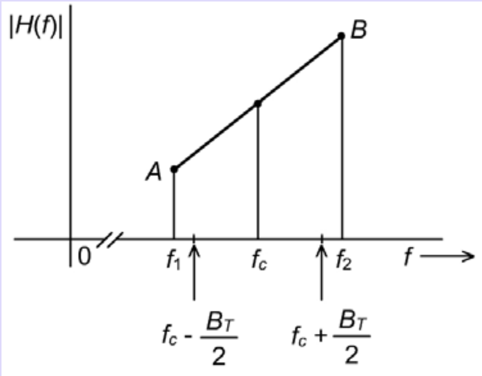


Fig. 5.24: Segment A to B of Fig. 5.23

As can be seen from the figure, changes in the instantaneous frequency will give rise to corresponding changes in the output amplitude. Envelope detection of this output will produce the required message signal.

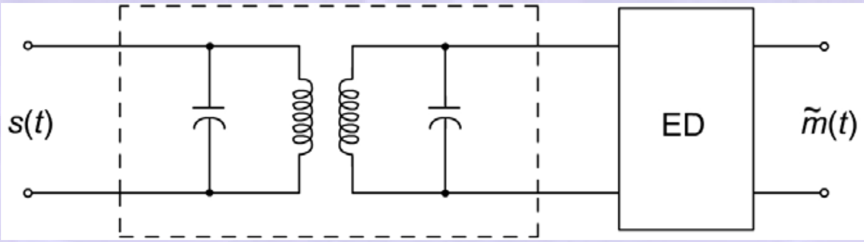


Fig. 5.25: Tuned circuit demodulator

Consider the demodulator circuit shown in Fig. 5.25. The primary of the coupled circuit is tuned to f_c [1] whereas the secondary is tuned to f_0, where $f_0 > f_c$. If, over the frequency range $|f \pm f_c| \leq \frac{B_T}{2}$, the output of the primary is fairly constant, then, we can expect the $\tilde{m}(t)$ to resemble $m(t)$ fairly closely. This method of demodulating an FM signal is also known as *slope detection.*

Though the method is fairly simple, it suffers from the following disadvantages: the resonant circuits on the primary and secondary side are

[1] In a superheterodyne receiver, the detector stage follows the IF stage. As such, f_c gets converted to f_{IF}.

tuned to two different frequencies; the frequency range of linear amplitude response of a tuned circuit is somewhat limited, making the demodulation of a WBFM unsatisfactory.

Scheme 3 (Balanced slope detection): This latter problem is partially overcome by using a balanced configuration (*balanced slope detection*). The scheme, shown in Fig. 5.26(a) has three tuned circuits: two on the secondary side of the input transformer and one on the primary. The resonant circuit on the primary is tuned to f_c whereas the two resonant circuits on the secondary side are tuned to two different frequencies, one above f_c and the other, below f_c. The outputs of the tuned circuits on the secondary are envelope detected separately; the difference of the two envelope detected outputs would be proportional to $m(t)$.

Though the balanced configuration has linearity over a wider range (as can be seen from Fig. 5.26(b), the width of linear frequency response is about $3B$, where $2B$ is the width of the 3-dB bandwidth of the individual tuned circuits) and does not require any DC bock (The two resonant frequencies of the secondary are appropriately selected so that output of the discriminator is zero for $f = f_c$), it suffers from the disadvantage that the three tuned circuits are to be maintained at three different frequencies.

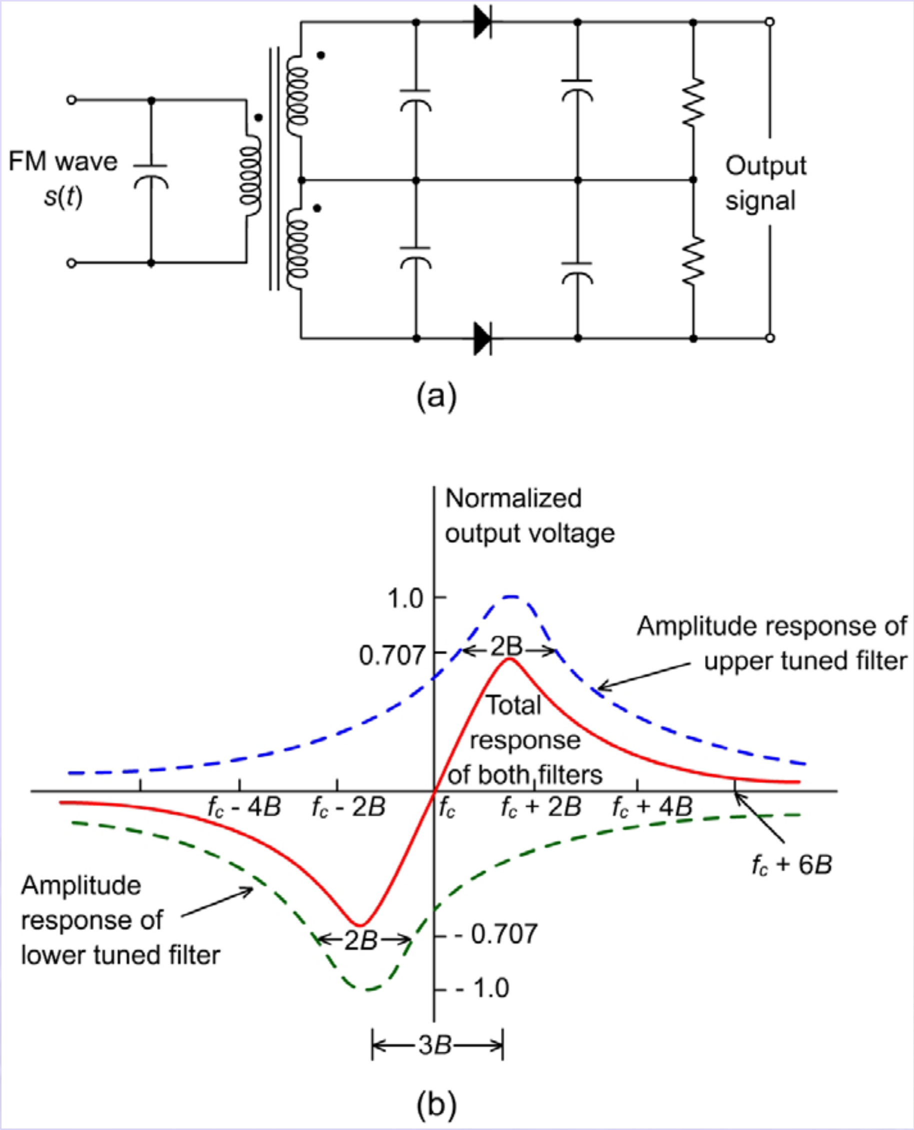


Fig. 5.26: Balanced slope detection (a) circuit schematic (b) response curve

Example 5.9

Let the FM waveform of Fig. 5.4(b) be the input to a differentiator followed by an envelope detector. Let us find an expression for the output of the differentiator and sketch the output of the envelope detector. We shall assume that the differentiator will produce the appropriate step change in the output for sudden changes in the input frequency.

Let the FM waveform between the time instants (t_1, t_2) be taken as $\cos(2\pi f_1 t)$ and that between time instants (t_2, t_3) be taken as $\cos(2\pi f_2 t)$, where $f_2 > f_1$.

Then the output of the differentiator is,

$t_1 < t < t_2 : 2\pi f_1 \sin(2\pi f_1 t) = A_1 \sin(2\pi f_1 t)$

$t_2 < t < t_3 : 2\pi f_2 \sin(2\pi f_2 t) = A_2 \sin(2\pi f_2 t)$

with $A_2 > A_1$ and $A_1, A_2 > 0$.

(Imagine Figure 5.4(a), with two different frequencies.) The output of the ED is proportional to A_1 during $t_1 < t < t_2$ and proportional to A_2 during $t_2 < t < t_3$. Taking the constant of proportionality as unity, we have the output of ED as shown in Fig 5.27

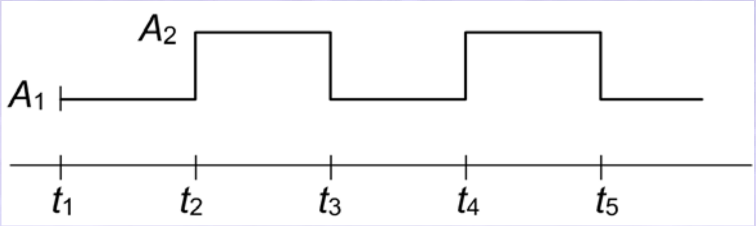


Fig. 5.27: Output of the ED

After DC block, we obtain the modulating square wave signal. Hence, $\dfrac{d}{dt}$ followed by ED with a DC block, would act as a demodulator for the FM. ♦

Example 5.10

a) Consider the RC network shown in Fig. 5.28. For the values of R and C given, we will show that for frequencies around 1.0 MHz, this can act as a differentiator.

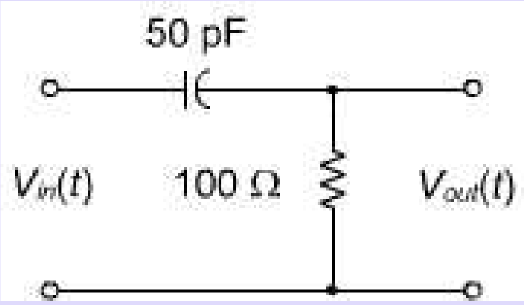


Fig. 5.28: The RC network of Example 5.10

b) Let us find the condition on R and C such that this network can act as a differentiator for frequencies around some frequency f_c.

c) For the above network, let $v_{in}(t) = s(t) = A_c \cos\left[2\pi \times 10^6 t + c_f m_I(t)\right]$ where $|k_f m(t)| << 10^6$ Hz. If $v_{out}(t)$ is envelope detected, we will show that, we can recover $m(t)$ from the ED output.

a) $H(f) = \dfrac{j 2\pi f RC}{1 + j 2\pi f RC}$

$RC = 100 \times 50 \times 10^{-12} = 5 \times 10^{-9}$ and for frequencies around 1.00 MHz,

$2\pi f RC = 2\pi \times 10^6 \times 5 \times 10^{-9} = 10\pi \times 10^{-3} = \dfrac{\pi}{100}$

As $\dfrac{\pi}{100} << 1$, we can take $H(f)$ as

$H(f) \simeq j 2\pi f RC$, which is a differentiator.

$= \left(5 \times 10^{-9}\right) j 2\pi f$

b) for frequencies around some f_c, we require, $2\pi f RC << 1$.

c) With $s(t)$ as the input, $v_{out}(t)$ is

$$v_{out}(t) = 5 \times 10^{-9} \frac{d}{dt}\left[s(t)\right]$$

$$= 5 \times 10^{-9}\, A_c \sin\left[2\pi \times 10^6 t + c_f\, m_I(t)\right]\left\{2\pi \times 10^6 + 2\pi k_f\, m(t)\right\}$$

$$\text{Output of the ED} = 5 \times 10^{-9}\, A_c \left[2\pi \times 10^6 + 2\pi k_f\, m(t)\right]$$

$$= 2\pi \times 10^6 \times 5 \times 10^{-9}\, A_c \left[1 + \frac{k_f}{10^6}\, m(t)\right].$$

$$= \frac{\pi A_c}{100}\left[1 + \frac{k_f}{10^6}\, m(t)\right]$$

♦

5.6.2 Phase shift discriminator

This method of FM demodulation involves converting frequency variations into phase variations and detecting the phase changes. In other words, this method makes use of linear phase networks instead of the linear amplitude characteristic of the circuits used in the previous method. Under this category, we have the **Foster-Seely discriminator** (and its variant **the ratio detector**) and the **quadrature detector.** Foster-Seely discriminator and the ratio detector have been discussed in appendix A5.2. We shall now explain the operation of the quadrature detector.

Quadrature Detector (QD): Consider the FM signal

$$s(t) = A_c \cos\left[\omega_c t + \phi(t)\right] \quad \text{where}$$

$$\phi(t) = 2\pi k_f \int_{-\infty}^{t} m(\lambda)\, d\lambda$$

Then $\phi'(t) = \dfrac{d\phi(t)}{dt} = 2\pi k_f\, m(t)$

$$\simeq \frac{1}{\Delta t}\left[\phi(t) - \phi(t - \Delta t)\right], \text{ provided } \Delta t \text{ is small.}$$

$\phi(t - \Delta t)$ can be obtained from $\phi(t)$ with a delay line or a network with linear phase.

Consider the scheme shown in Fig. 5.29. $H(f)$ has the transfer characteristic given by

$$H(f) = e^{j\,A(f)}$$

where $A(f)$, the phase function, can be well approximated by a linear phase function, namely,

$$A(f) \simeq \begin{cases} -\dfrac{\pi}{2} - 2\pi(f - f_c)\,\Delta t\,, & f > 0\,,\ |f - f_c| < \dfrac{B_T}{2} \\ \dfrac{\pi}{2} - 2\pi(f + f_c)\,\Delta t\,, & f < 0\,,\ |f + f_c| < \dfrac{B_T}{2} \end{cases} \tag{5.29}$$

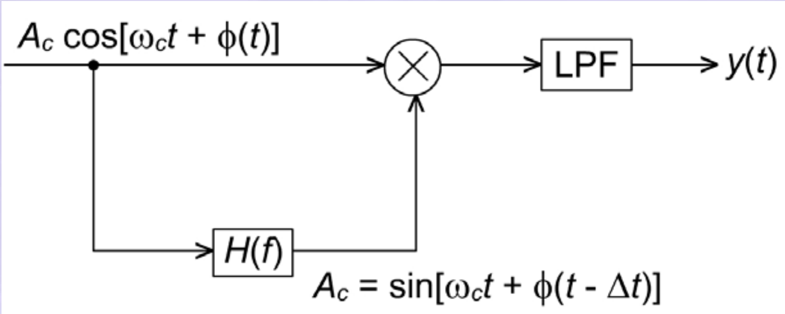


Fig. 5.29: Block diagram of a quadrature detector

With $H(f)$ specified as above, let us calculate the output of the network with the input $A_c\cos[\omega_c t + \phi(t)] = A_c\,\dfrac{e^{j[\omega_c t + \varphi(t)]} + e^{-j[\omega_c t + \varphi(t)]}}{2}$. We can take $\dfrac{1}{2}A_c\,e^{j[\omega_c t + \varphi(t)]}$ to represent the positive part of the spectrum. As $H(f)$ has the linear phase term $e^{-j\,2\pi f\,\Delta t}$, (which will contribute a delay of Δt), the term at the output of the filter is

$$\frac{1}{2}A_c\,e^{j\left[\omega_c(t - \Delta t) + \varphi(t - \Delta t) - \frac{\pi}{2} + \omega_c\,\Delta t\right]} = \frac{1}{2}A_c\,e^{j\left[\omega_c t + \varphi(t - \Delta t) - \frac{\pi}{2}\right]}.$$

Similarly, the output corresponding to the negative part of the input spectrum would be $\frac{1}{2} A_c \, e^{-j\left[\omega_c t + \varphi(t - \Delta t) - \frac{\pi}{2}\right]}$. Combining these two terms, we have at the output of the filter, the quantity

$$A_c \cos\left[\omega_c t + \varphi(t - \Delta t) - \frac{\pi}{2}\right] = A_c \sin\left[\omega_c t + \varphi(t - \Delta t)\right].$$

The $\frac{\pi}{2}$ phase shift provided by $H(f)$ at $f = \pm f_c$ gives rise to the term *quadrature detector*. Multiplication of this output by $A_c \cos\left[\omega_c t + \varphi(t)\right]$ followed by low pass filtering yields the output $y(t)$ proportional to $\sin\left[\phi(t) - \phi(t - \Delta t)\right]$. Assuming Δt to be very small, $y(t)$ can be approximated as,

$$\begin{aligned} y(t) &\simeq k_1 \{\phi(t) - \phi(t - \Delta t)\} \\ &= k_1 \Delta t \, \varphi'(t) \\ &= k_2 \, m(t) \end{aligned}$$

where k_1 and k_2 are constants with $k_2 = c_f k_1 \Delta t$.

Several tuned circuits, when properly designed, can provide the band-pass response with the phase characteristic given by Eq. 5.29. Consider the series RLC circuit shown in Fig. 5.30.

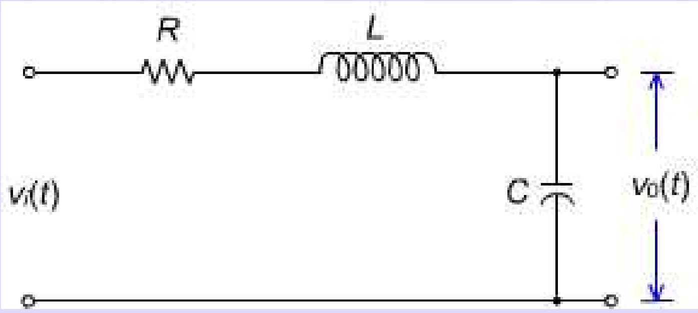


Fig. 5.30: A network to provide the phase of Eq. 5.29

$$H(f) = \frac{V_0(f)}{V_i(f)} = \frac{\dfrac{1}{j\,2\pi f C}}{R + j\,2\pi f L + \dfrac{1}{j\,2\pi f C}}$$

$$= \frac{1}{j\,2\pi f RC - (2\pi f)^2 LC + 1} \qquad (5.30)$$

Let $f_0 = \dfrac{1}{2\pi\sqrt{LC}}$ and $f_b = \dfrac{R}{2\pi L}$. Then, Eq. 5.30 reduces to

$$H(f) = \frac{f_0^2}{\left(f_0^2 - f^2\right) + j\,(f f_b)}$$

Consider $f > 0$

$$\arg\left[H(f)\right] = -\tan^{-1}\left(\frac{f f_b}{f_0^2 - f^2}\right)$$

$$= -\frac{\pi}{2} - \tan^{-1}\left(\frac{f^2 - f_0^2}{f f_b}\right)$$

$$= -\frac{\pi}{2} - \tan^{-1}\left[\frac{(f + f_0)(f - f_0)}{f f_b}\right]$$

Let the circuit be operated in a small frequency interval, around f_0 so that $f \simeq f_0$.

Then, $f + f_0 \simeq 2 f_0$, $\dfrac{(f + f_0)}{f f_b} \simeq \dfrac{2}{f_b}$ and

$$\arg\left[H(f)\right] \simeq -\frac{\pi}{2} - \tan^{-1}\frac{2(\delta f)}{f_b}$$

$$\simeq -\frac{\pi}{2} - \tan^{-1}\frac{2(\delta f)Q}{f_0}$$

where $Q = \dfrac{f_0}{f_b}$ and $\delta f = (f - f_0)$

If $\dfrac{2(\delta f)Q}{f_0} << 1$, then,

$$\arg\left[H(f)\right] \simeq -\frac{\pi}{2} - \frac{2Q(\delta f)}{f_0} \qquad (5.31)$$

As $H(f)$ is the frequency response of a network with real impulse response, we will have

$$\arg\left[H(-f)\right] = -\arg\left[H(f)\right]$$

By choosing $f_0 = f_c$, we have $\delta f = (f - f_c)$ and $\Delta t = \frac{Q}{\pi} \cdot \frac{1}{f_c}$

There are other circuit configurations (other than the one given in Fig. 5.30) that can provide the required phase shift for the quadrature detector. The circuit given in Fig. 5.31 is another possibility.

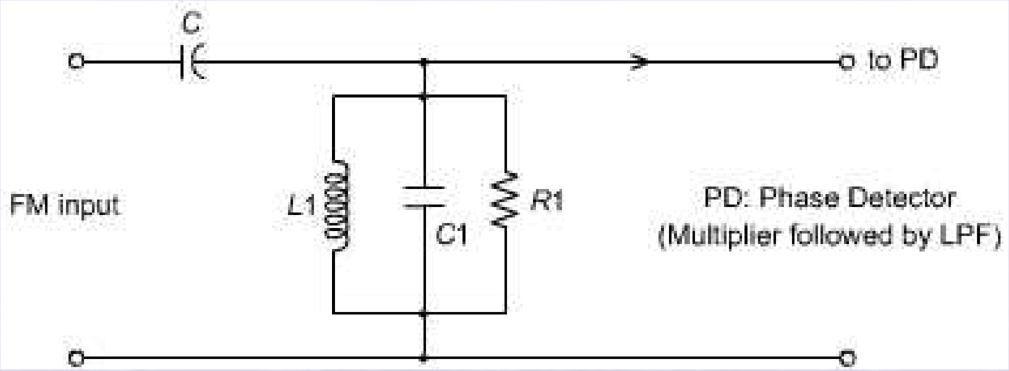


Fig. 5.31: Another phase shift circuit for the QD

Here C provides a very high reactance at the carrier frequency and the parallel tuned circuit resonates at $f = f_c$.

Quadrature detector is well suited to IC construction. Companies such as Signetics, have built high quality FM receivers using QD as the FM demodulator. Some details on these FM receivers can be found in Roddy and Coolen [2].

5.6.3 Zero-crossing detection

Consider the scheme shown in Fig. 5.32(a). When the input to the hard limiter is a sine wave of period T_0, it produces at its output a square wave of the same period, with the transitions in the square wave occurring at the zero crossings of the input sine wave. When the input to the hard limiter is an FM signal, the hard limiter output appears as a square wave of varying frequency. The hard limiter output $v_H(t)$, triggers a monostable pulse generator, which produces a short pulse of amplitude A and duration τ at each upward (or downward) transition of $v_H(t)$.

Consider a time interval T such that $\frac{1}{W} >> T >> \frac{1}{f_c}$, where W is the highest frequency present in the input signal. We shall assume that during the T sec. interval, the message signal $m(t)$ is essentially constant which implies that instantaneous frequency $f_i(t)$ is also a near constant (Fig. 5.32(b)). Then the monostable output, $v_p(t)$, looks like a pulse train of nearly constant period. The number of such pulses in this interval $n_T \simeq T f_i(t)$ with an average value,

$$v_I(t) = \frac{1}{T} \int_{t-T}^{t} v_p(\lambda)\, d\lambda$$

$$= \frac{1}{T} n_T A \tau \simeq A \tau f_i(t)$$

After the DC block, we will have $y(t)$ being proportional to $m(t)$, which is the desired result.

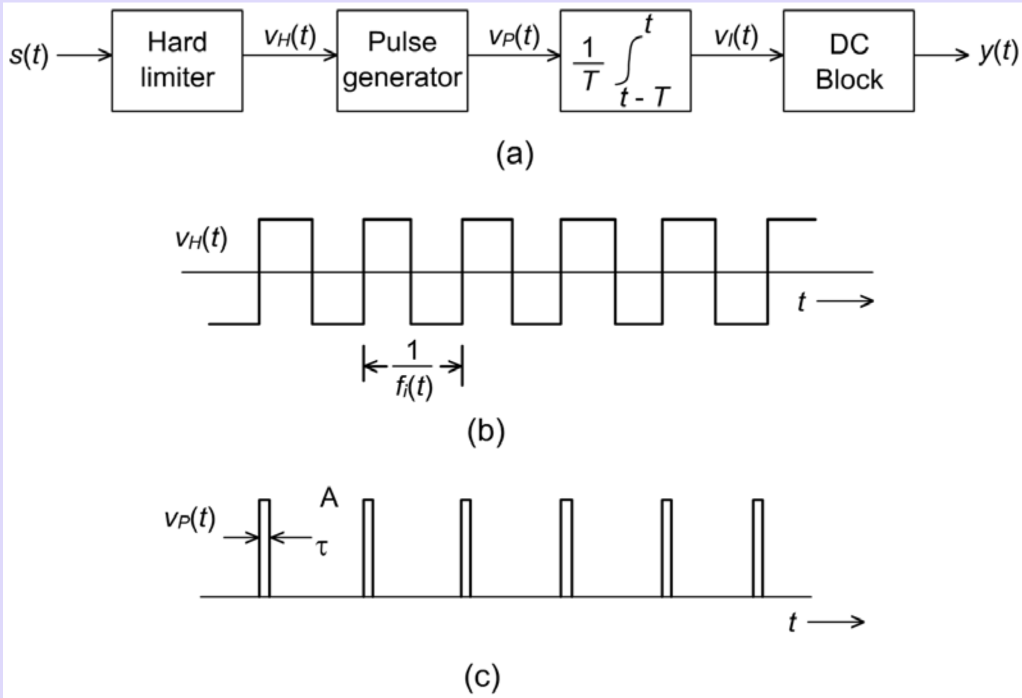


Fig 5.32: Zero-crossing detector

5.6.4 FM demodulation using PLL

PLL is a versatile building block of the present day communication systems. Besides FM demodulation, it has a large number of other applications such as carrier tracking (in schemes with a pilot carrier and even suppressed carrier; recall the squaring loop, Section 4.2.3), timing recovery, frequency synthesis etc.

The basic aim of a PLL is to *lock* (or synchronize) the instantaneous angle of a VCO output to the instantaneous angle of a signal that is given as input to the PLL. In the case of demodulation of FM, the input signal to PLL is the received FM signal.

In its simplest form, PLL consists of a phase detector and a VCO connected as shown in Fig. 5.33(a).

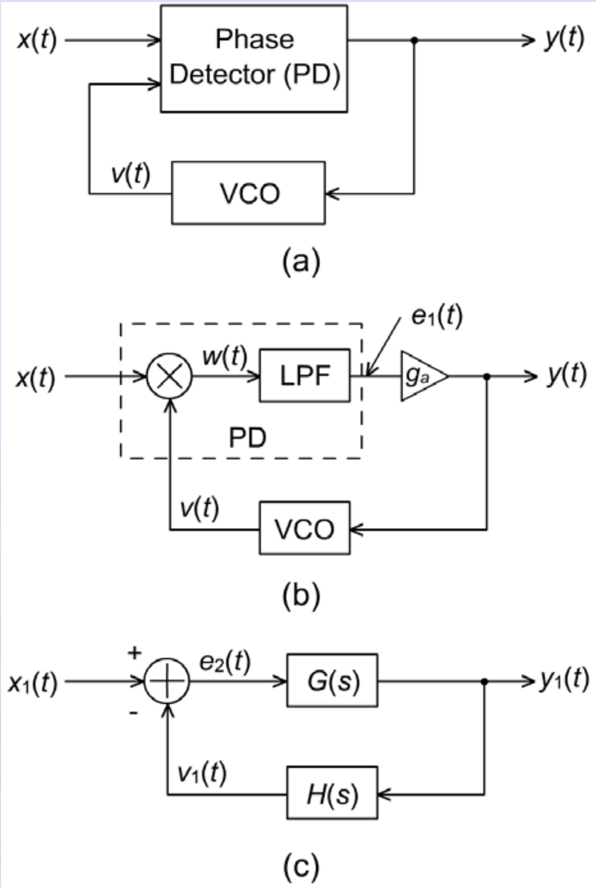


Fig. 5.33: Phase lock loop (a) Basic configuration
(b) PD of (a) in functional form
(c) PLL as a negative feedback loop

PD makes the comparison of the instantaneous phase of $x(t)$ and $v(t)$, and is designed such that $\theta_v(t)$, the instantaneous phase of $v(t)$ locks on to $\theta_x(t)$, the instantaneous phase of $x(t)$, if necessary with some fixed phase difference. (This will become evident later.)

A number of circuits are available which have been used as phase detectors. In the context of FM demodulation, the most common PD is that of an analog multiplier followed by a LPF (Fig. 5.33(b)). The scheme resembles closely that of a negative feedback amplifier configuration, shown in Fig. 5.33(c). In this figure, s is the variable of the Laplace transform, $G(s)$ is system function in the forward path whereas $H(s)$ is the network in the feedback path. A properly designed negative feedback system ensures that the error quantity $e_2(t)$ is fairly close to zero so that $v_1(t) \simeq x_1(t)$. This is ensured by providing sufficiently high loop gain. Similarly, by making the amplifier gain g_a sufficiently large, it is possible to make $\theta_v(t)$ follow the changes in $\theta_x(t)$.

Let $x(t) = \cos\left[\omega_c t + \varphi(t)\right]$ and let the VCO output be, $v(t) = \cos\left[\omega_c t + \psi(t)\right]$. Then, from Fig. 5.33(b),

$$w(t) = x(t)v(t) = \cos\left[\omega_c t + \varphi(t)\right]\cos\left[\omega_c t + \psi(t)\right]$$

$$= \frac{1}{2}\left\{\cos\left[2\omega_c t + \varphi(t) + \psi(t)\right] + \cos\left[\varphi(t) - \psi(t)\right]\right\}$$

Only the term $\frac{1}{2}\cos\left[\varphi(t) - \psi(t)\right]$ will appear at the output of the LPF; that is

$$e_1(t) = \frac{1}{2}\cos\left[\varphi(t) - \psi(t)\right] \tag{5.32}$$

(We are assuming that the LPF has unit gain)

As the phase detector, we want $e(t)$ to be zero where $\varphi(t) = \psi(t)$; but from Eq. 5.32, $e_1(t)$ is maximum when $\varphi(t) = \psi(t)$. This is not the characteristic of a proper phase detector. This anomaly can be corrected, if the loop provides a $\frac{\pi}{2}$ phase shift so that the output of the VCO is $\sin\left[\omega_c t + \psi(t)\right]$. That is, the loop locks in *phase quadrature*. Here after, we shall assume this $\frac{\pi}{2}$ phase shift in the VCO output.

Now let us look at the demodulation of FM. Let

$$x(t) = s(t) = A_c\cos\left[\omega_c t + \varphi(t)\right]$$

where $\varphi(t) = 2\pi k_f \int_{-\infty}^{t} m(\tau)\, d\tau$. The VCO is designed such that when $y(t)$, the control voltage is zero, its frequency is f_c. (As mentioned earlier, in a superhet, the demodulator follows the IF stage and f_c is actually f_{IF}, which is a known quantity). This is called the **free running frequency** of the VCO. Hence, the VCO output can be written as

$$v(t) = A_v \sin\left[\omega_c t + \psi(t)\right] \tag{5.33a}$$

where $\psi(t) = 2\pi k_v \int^{t} y(\tau)\, d\tau$ (5.33b)

and k_v is the voltage sensitivity of the VCO, in units of Hz/volt.

Or $\psi'(t) = 2\pi k_v\, y(t)$ (5.33c)

$= K_1 y(t)$, where $K_1 = 2\pi k_v$

$e_1(t)$ of Fig. 5.33(b) is,

$$e_1(t) = \frac{A_c A_v}{2} \sin\left[\varphi(t) - \psi(t)\right] * h(t) \tag{5.34a}$$

where $h(t)$ is the impulse response of the LPF,

and $y(t) = g_a\, e_1(t)$ (5.34b)

Let $\theta_e(t) = \varphi(t) - \psi(t)$ (5.35)

Then $y(t) = \dfrac{A_c A_v g_a}{2} \sin\left[\theta_e(t)\right] * h(t)$

$= K_2 \sin\left[\theta_e(t)\right] * h(t)$, where $K_2 = \dfrac{A_c A_v g_a}{2}$ (5.36)

Using Eq. 5.35, 5.36 and 5.33(c), we can draw following block diagram (Fig. 5.34) for the PLL.

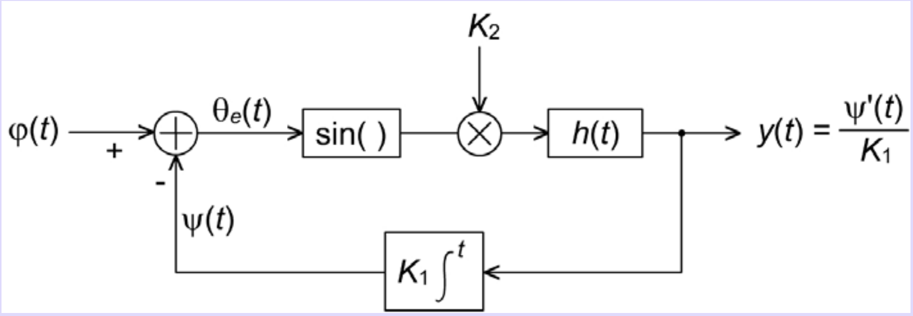


Fig. 5.34: Equivalent circuit of a PLL

Fig. 5.34 brings out much more clearly the negative feedback nature of the PLL, where in the quantities involved are instantaneous phase deviations of the input and the VCO output.

Let the loop be in lock so that $|\theta_e(t)| << 1$ for all t. Then $\sin \theta_e(t) \simeq \theta_e(t)$, $\psi(t) \simeq \varphi(t)$ and $\psi'(t) \simeq \varphi'(t)$. As $\varphi(t) = 2\pi k_f \int_{-\infty}^{t} m(\tau)\, d\tau$, we have

$$\varphi'(t) = 2\pi k_f\, m(t)$$

Hence $y(t) = \frac{\psi'(t)}{K_1} \simeq \frac{2\pi k_f}{K_1} m(t)$. That is, $y(t) \propto m(t)$.

The above has been a very elementary analysis of the operation of PLL. Literature on PLL is very widespread. For a more detailed and rigorous analysis, refer Taub and Shilling [3].

Summary of the Detectors:

Slope detection using a single tuned circuit has been presented to show how simple an FM demodulator could be; it is not used in practice because of its limited range of linearity. Though balanced slope detection offers linearity over a wider frequency range, it suffers from the problem of tuning. The Foster-Seely discriminator and the ratio detector have been the work horses of the FM industry until recently; they are now becoming less and less important as better circuit configurations have been developed especially from the point of view of IC design. The quadrature detector offers very high linearity and is commonly used in high quality receivers. Except the phase-shifting network, the rest of the detector, including the IF amplifier is available in a single chip (e.g. RCA CA 3089E). Commercial zero crossing detectors have better than 0.1 % linearity and can operate from 1Hz to 10 MHz. A divide by ten counter inserted after the hard limiter extends the range up to 100 MHz. This type of detector is best suited when exceptional linearity over a very large frequency deviation is required. It is less useful when the frequency deviation is a small fraction of the carrier frequency. PLL performs better than other demodulators when the signal-to-noise ratio at the input to the detector is somewhat low.

5.7 BandPass Limiter (BPL)

We know that information in an FM wave resides in the instantaneous frequency (or in the zero-crossings) of the signal. As such, the amplitude changes of the carrier are irrelevant. In fact, as was pointed out in section 5.6.1, if the envelope $A(t)$ is not a constant, it will give rise to distortion in the demodulated output. In other words, from the point of view of proper demodulation, we want $A(t) = A$, a constant.

Though the FM signal that is generated at the transmitter has a constant envelope, the received signal may not possess this property. This is due to various impairments during propagation on the channel, namely, channel noise, distortion introduced by the channel, etc. BPL helps us to recover the constant envelope FM signal from the one that has envelope fluctuations.

A band-pass limiter consists of a hard limiter followed by a band-pass filter. The input-output relationship of a hard-limiter is given by

$$y(t) = \begin{cases} A, & x(t) > 0 \\ -A, & x(t) < 0 \end{cases} \tag{5.37}$$

where $x(t)$ is the input, $y(t)$ is the output and A is a constant. Let $x(t)$ be as shown in Fig. 5.35(a).

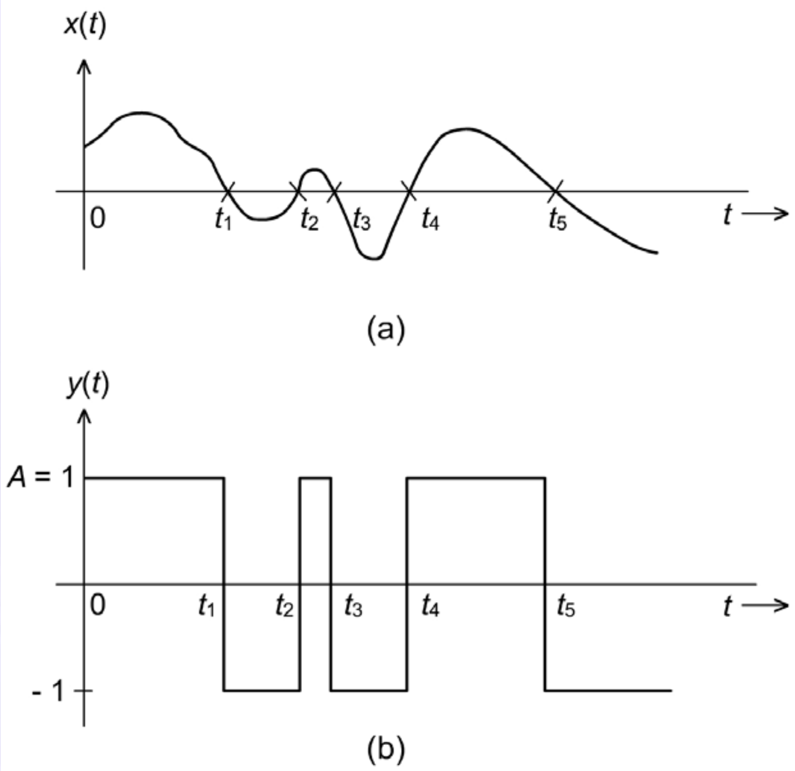


Fig. 5.35: (a) Input to the hard limiter

(b) Output of the hard limiter

Then the output with A taken as 1, would be as shown in Fig. 5.35(b). If the input to the hard limiter is the FM signal (with or without envelope fluctuations), then the output would be a sequence of alternate positive and negative rectangular pulses with durations that are not uniform. Such a wave is difficult to analyze. However, $\cos\theta$ as a function of θ is always periodic with period 2π. Hence the hard limiter output, when considered as a function of θ, will be a periodic square wave (with period 2π) when the input to the limiter is a cosine signal. Hence, if $x(t) = \cos \omega t = \cos\theta$, then

$$y(\theta) = \begin{cases} 1, & \cos\theta > 0 \\ -1, & \cos\theta < 0 \end{cases}$$

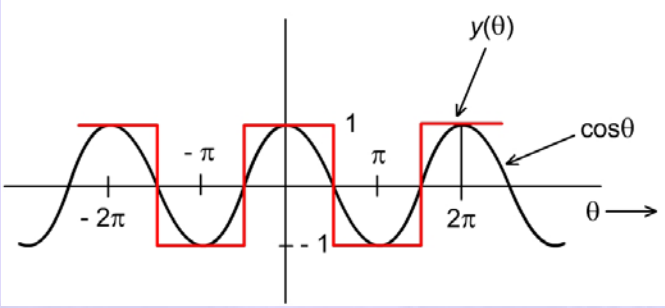


Fig. 5.36: Hard limiter output with $\cos\theta$ as input

$y(\theta)$ being a periodic signal (Fig. 5.36), we can expand it in terms of Fourier series, namely

$$y(\theta) = \frac{4}{\pi}\left[\cos\theta - \frac{1}{3}\cos 3\theta + \frac{1}{5}\cos 5\theta + \cdots\right] \tag{5.38}$$

(Note that $y(\theta)$ is real, and has half-wave symmetry. Hence Fourier series consists of only cosine terms with the even harmonics missing. See Sec. 1.2.2.)

Let θ be the instantaneous angle of the FM signal; that

$$\theta = \theta_i(t) = \omega_c t + c_f \int^{t} m(\tau)\, d\tau = \omega_c t + \varphi(t)$$

Then, from Eq. 5.38, we have

$$y(\theta) = \frac{4}{\pi}\cos[\omega_c t + \varphi(t)] - \frac{4}{3\pi}\cos\left[3\left(\omega_c t + \varphi(t)\right)\right] + \cdots$$

At the output of BPL, we have the constant envelope FM waves with carrier frequencies $n f_c$ and frequency $n\Delta f$ respectively, where $n = 1, 3, 5$ etc. With an appropriate BPF, it is possible for us to obtain constant envelope FM signal with carrier frequency f_c and deviation Δf. We will assume that BPF will pass the required FM signal, suppressing the components with spectra centered at the harmonics of f_c. Incidentally, BPL can be used as a frequency multiplier, multiplication factors being 3, 5, 7, etc.

A hard limiter, as defined by Eq. 5.37 can be easily be realized in practice. One such circuit is shown in Fig. 5.37.

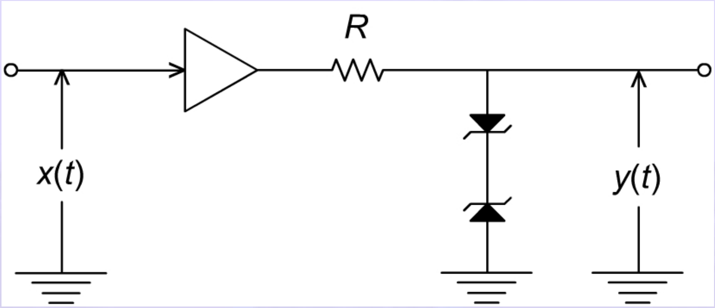


Fig. 5.37: Circuit realization of a hard limiter

As shown in the figure, the circuit consists of a high gain amplifier, a current limiting resistor R and two zener diodes arranged in a back-to-back configuration. Zeners can be chosen to have the appropriate break down voltages.

5.8 Broadcast FM

5.8.1 Monophonic FM Reception

FM stations operate in the frequency range 88.1 to 107.9 MHz with stations being separated by 200 kHz; that is, the transmission bandwidth allocation for each station is about 200 kHz. The receiver for the broadcast FM is of the superheterodyne variety with the intermediate frequency of 10.7 MHz. As the audio bandwidth is 15 kHz, these stations can broadcast high quality music.

Let us now look at the receiver for the single channel (or monophonic) broadcast FM. Like the superhet for AM, the FM receiver (Fig. 5.38) also has the front end tuning, RF stage, mixer stage and IF stage, operating at the allocated frequencies. This will be followed by the limiter-discriminator combination, which is different from the AM radio.

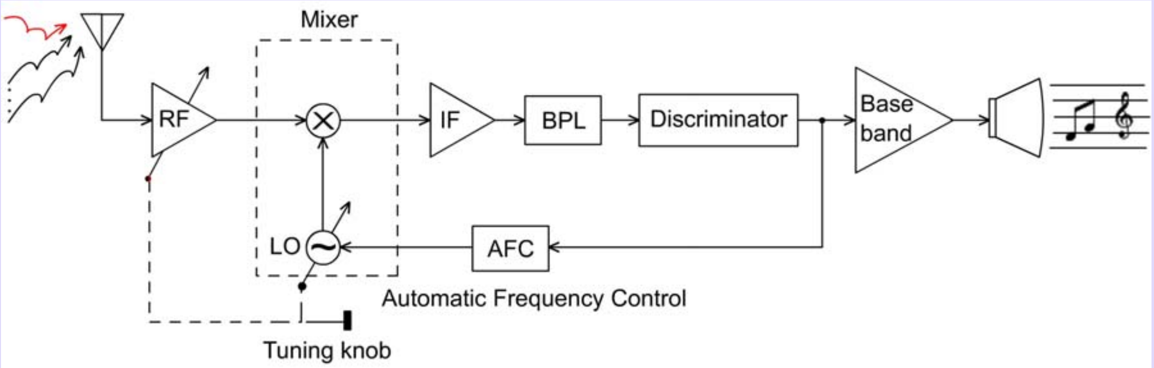


Fig. 5.38: FM broadcast superhet receiver

The need for a band-pass limiter has already been explained. As far as the frequency discriminator is concerned, we have a lot of choices: Foster-Seely, ratio detector, quadrature detector, PLL etc. This is unlike AM, where the envelope detector is the invariable choice. Also, the output of the discriminator is used in a feedback mode to control the frequency stability of the local oscillator. This is very important as a change in the frequency of the LO can result in improper demodulation. (Recall that, in the case AM, the envelope detector output is used for AVC.) However, most of the present day receivers might be using frequency synthesizers in place of a LO. These being fairly stable, AFC is not a requirement.

Basic operation of the AFC block is as follows. Assume that the receiver is properly tuned so that $f_{LO} - f_c = f_{IF}$. Then the discriminator input will have equal frequency variations with respect to f_{IF}; hence the discriminator output will vary symmetrically with respect to zero output. As such, the net DC voltage is zero and LO frequency will not be changed. Assume, however, that LO frequency is not correct. Let $f_{LO} - f_c = f_{IF}^{'}$, where $f_{IF}^{'} < f_{IF}$. Then the input to the S-curve of the discriminator will be operating mostly in the negative output region. This means that the discriminator output will have a net negative voltage. If this voltage is applied to the varicap in the LO circuit, f_{LO} will increase which implies $f_{IF}^{'}$ will be increased. In other words, $f_{IF}^{'}$ tends towards f_{IF}. Similarly, if $f_{LO} - f_c = f_{IF}^{'}$ where $f_{IF}^{'} > f_{IF}$, then the discriminator output is positive most of the time which implies a net positive DC value. This voltage will increase the capacitance of the varicap which implies f_{LO} decreases and this makes $f_{IF}^{'}$ tend to f_{IF}. Discriminator output goes through a base-band amplifier[1] (with a bandwidth of 15 kHz), whose output drives the speaker.

5.8.2 Two-channel (stereo) FM

Two-channel (stereo) FM is fairly common these days and stereo transmission has been made compatible with mono-aural reception. As the name indicates, in the two-channel case, audio signal is derived as the output of two separate microphones. These are generally called as left microphone and the right microphone. Let the corresponding output signals be denoted by $m_L(t)$ and $m_R(t)$. The two- channel transmitter is shown in Fig. 5.39.

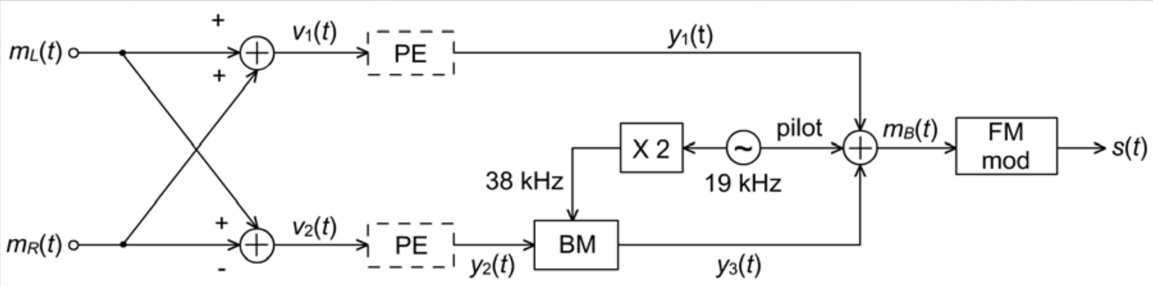
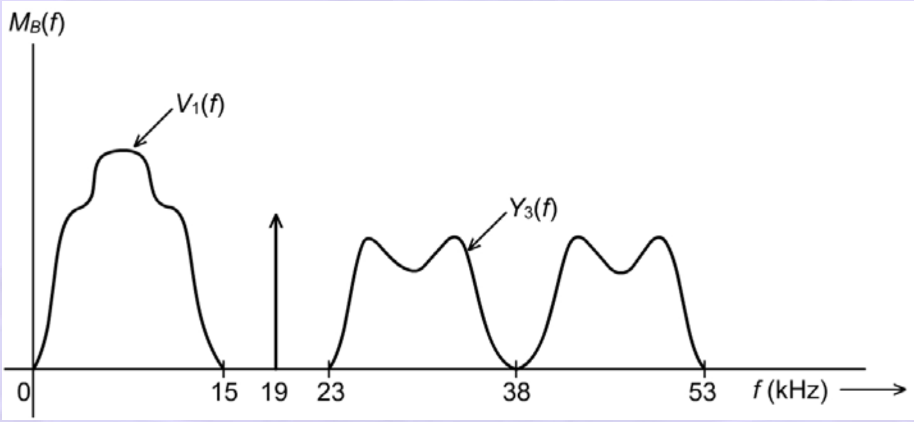


Fig. 5.39: FM stereo transmission scheme

[1] Output of the discriminator goes through a de-emphasis network before being applied to a base-band amplifier. Pre-emphasis at the transmitter and de-emphasis at the receiver are used to improve the signal-to-noise ratio performance. Pre-emphasis and de-emphasis will be discussed in Chapter 7.

As shown in the figure, $v_1(t) = m_L(t) + m_R(t)$ and $v_2(t) = m_L(t) - m_R(t)$. The signals $v_1(t)$ and $v_2(t)$ go through a pre-emphasis stage. The signal $y_2(t)$ is applied to a balanced modulator, the other input to modulator being the carrier with a frequency of 38 kHz. Hence $y_3(t)$ is a DSB-SC signal. The carrier 38 kHz is derived from the primary source at 19 kHz and a frequency doubler. The final base-band signal $m_B(t)$ consists of the sum of $y_1(t)$, $y_3(t)$ and 19 kHz primary carrier. Typical spectrum (for $f > 0$) of the base-band signal as shown in Fig. 5.40.



$$V_1(f) = m_L(f) + m_R(f)$$

$$Y_3(f) = K\left[Y_2(f - f_0) + Y_2(f + f_0)\right]$$

$f_0 = 38$ kHz and K is a constant

Fig. 5.40: Spectrum of the final base-band signal

The signal $m_B(t)$ is used to frequency modulate the carrier allotted to the station. The resulting signal $s(t)$ is transmitted on to the channel.

The block schematic of the stereo receiver upto the discriminator is the same as shown in Fig. 5.38 (that is, monophonic case). Hence, let us look at the operations performed by the stereo receiver after recovering $m_B(t)$ from $s(t)$.

These operations are indicated in Fig. 5.41. The DSB-SC signal is coherently demodulated by generating the 38 kHz carrier from the pilot carrier of 19 kHz. (Note that if a pilot carrier of 38 kHz had been sent, it would have been difficult to extract it at the receiver.) $r_1(t)$ and $r_2(t)$, after de-emphasis will yield $v_1(t)$ and $v_2(t)$. (Note that constants of proportionality are ignored and are taken as 1.)

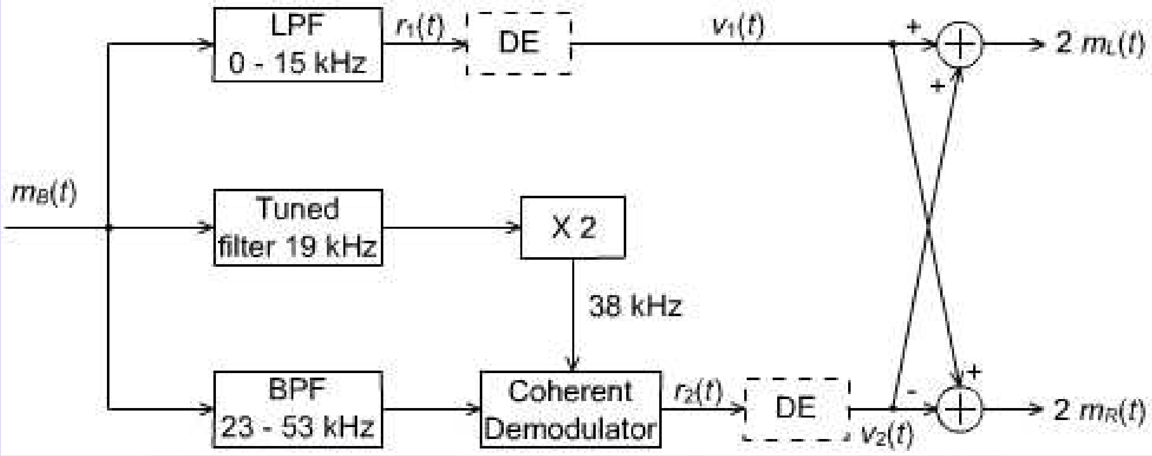


Fig. 5.41: Scheme to recover $m_L(t)$ and $m_R(t)$

From $v_1(t)$ and $v_2(t)$, the individual channels signals, namely $m_L(t)$ and $m_R(t)$ are obtained. These signals, after suitable power amplification, will drive the two speakers, arranged such that one is on the left and the other is on the right. If the receiver is not stereophonic, it would respond only to $v_1(t)$ thereby making it stereo transmission and monophonic reception compatible.

Appendix A5.1

Table of Bessel Functions

$J_n(\beta)$

n \ β	0.5	1	2	3	4	5	6	8	10
0	0.9385	0.7652	0.2239	- 0.2601	- 0.3971	- 0.1776	0.1506	0.1717	- 0.2459
1	0.2423	0.4401	0.5767	0.3391	- 0.0660	- 0.3276	- 0.2767	0.2346	0.0435
2	0.0306	0.1149	0.3528	0.4861	0.3641	0.0465	-0.2429	- 0.1130	0.2546
3	0.0026	0.0196	0.1289	0.3091	0.4302	0.3648	0.1148	- 0.2911	0.0584
4	0.0002	0.0025	0.0340	0.1320	0.2811	0.3912	0.3576	- 0.1054	- 0.2196
5	-	0.0002	0.0070	0.0430	0.1321	0.2611	0.3621	0.1858	- 0.2341
6		-	0.0012	0.0114	0.0491	0.1310	0.2458	0.3376	- 0.0145
7			0.0002	0.0025	0.0152	0.0533	0.1296	0.3206	0.2167
8			-	0.0005	0.0040	0.0184	0.0565	0.2235	0.3179
9				0.0001	0.0009	0.0055	0.0212	0.1263	0.2919
10				-	0.0002	0.0014	0.0070	0.0608	0.2075
11					-	-	0.0020	0.0256	0.1231
12							0.0005	0.0096	0.0634
13							0.0001	0.0033	0.0290
14							-	0.0010	0.0120

Appendix A5.2

Phase Shift Discriminator

(i) Foster-Seely Discriminator

(ii) The Ratio Detector

A5.2.1 Foster-Seely discriminator

Fig. A5.2.1 illustrates the circuit diagram of this discriminator where all the resonant circuits involved are tuned to the same frequency. Note the similarity between this circuit and the circuit of Fig. 5.26. Major differences are a by-pass capacitor C between the primary and secondary, an additional inductance L and only a single tuned circuit on the secondary $\left(L_2 \,||\, C_2\right)$.

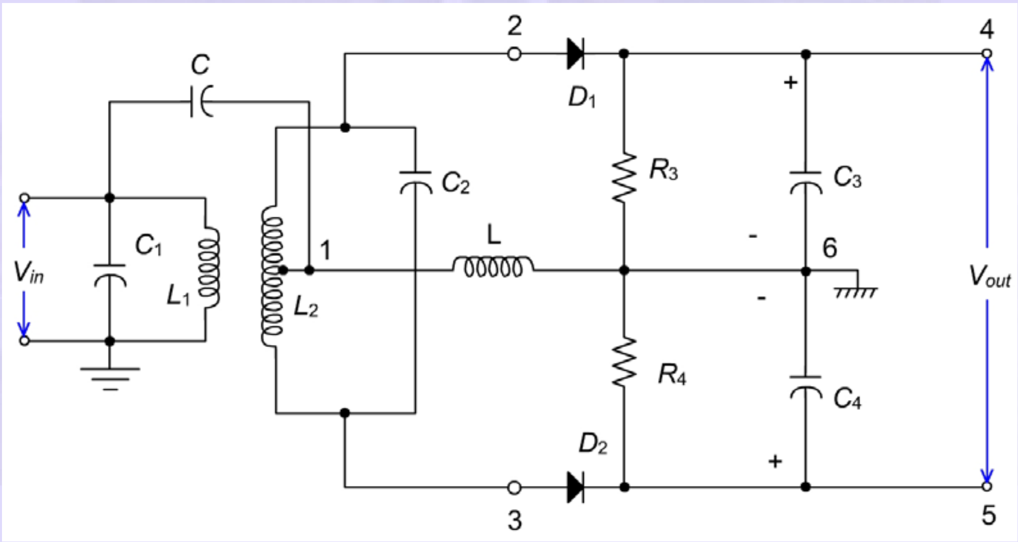


Fig. A5.2.1: Circuit schematic of Foster-Seely discriminator

In the frequency band of operation, C, C_3 and C_4 are essentially short circuits, which implies that the entire input voltage V_{in} would appear across L. The mutually coupled double tuned circuit has high primary and secondary Q and low mutual inductance. When evaluating the primary current, we may neglect the primary resistance and any impedance coupled from the secondary. Hence, the primary current

$$I_p = \frac{V_{in}}{j\,\omega L_1} \qquad \text{(A5.2.1)}$$

(Note that all the voltages and currents are phasor quantities)

The voltage induced in series with the secondary as a result of the primary current is given by

$$V_s = \pm\, j\,\omega M\, I_p \qquad \text{(A5.2.2)}$$

where the sign depends on the direction of the winding. Taking the negative sign and using Eq. A5.2.1 in Eq. A5.2.2, we have

$$V_s = -\left(\frac{M}{L_1}\right) V_{in}$$

Assuming the diode circuit will draw very little current, we calculate the current in the secondary because of V_s as,

$$I_s = \frac{V_s}{R_2 + j\left(X_{L_2} - X_{C_2}\right)}$$

where R_2 is the resistance associated with the inductance L_2, $X_{L_2} = \omega L_2$ and $X_{C_2} = \frac{1}{\omega C_2}$. Hence, the voltage across the terminals 2, 3 is given by

$$V_{23} = I_s\left(-\, j\, X_{C_2}\right)$$

$$= \frac{V_s\left(-\, j\, X_{C_2}\right)}{R_2 + j\left(X_{L_2} - X_{C_2}\right)}$$

$$= \frac{j\,M}{L_1}\,\frac{V_{in}\, X_{C_2}}{R_2 + j\, X_2} \quad \text{where } X_2 = \left(X_{L_2} - X_{C_2}\right)$$

The voltage applied to diode D_1, V_{62}, is

$$V_{62} = V_L + \frac{1}{2} V_{23}$$

$$= V_{in} + \frac{1}{2} V_{23} \qquad \text{(A5.2.3)}$$

Similarly, the voltage applied to diode D_2, V_{63}, is

$$V_{63} = V_{in} - \frac{1}{2} V_{23} \quad \text{(A5.2.4)}$$

The final output voltage V_{54} is, $V_{54} = V_{64} - V_{65}$ which is proportional to $\{|V_{62}| - |V_{63}|\}$. We will consider three different cases: $f = f_c$ [1], $f > f_c$ and $f < f_c$.

i) When the input frequency $f = f_c$ we have

$$V_{23} = \frac{j\,M}{L_1} \frac{V_{in}\,X_{C_2}}{R_2} = j\left(\frac{M\,X_{c_2}}{L_1 R_2}\right) V_{in} \quad \text{(A5.2.5a)}$$

That is, the secondary voltage V_{23} leads the primary voltage by 90°. Thus $\frac{1}{2} V_{23}$ will lead V_{in} by 90° and $-\frac{1}{2} V_{23}$ will lag V_{in} by 90°. Let us construct a phasor diagram by taking V_{in} as reference. This is shown in Fig. A5.2.2(a). As the magnitude of the voltage vectors applied to the diodes D_1 and D_2, V_{62} and V_{63} respectively are equal, the voltages V_{64} and V_{65} are equal and hence the final output V_{54} is zero.

---

[1] Note that in a superheterodyne receiver, the demodulator follows the IF stage. Hence, f_c is actually f_{IF}, and the discriminator circuit is always tuned to f_{IF}, irrespective of the incoming carrier frequency.

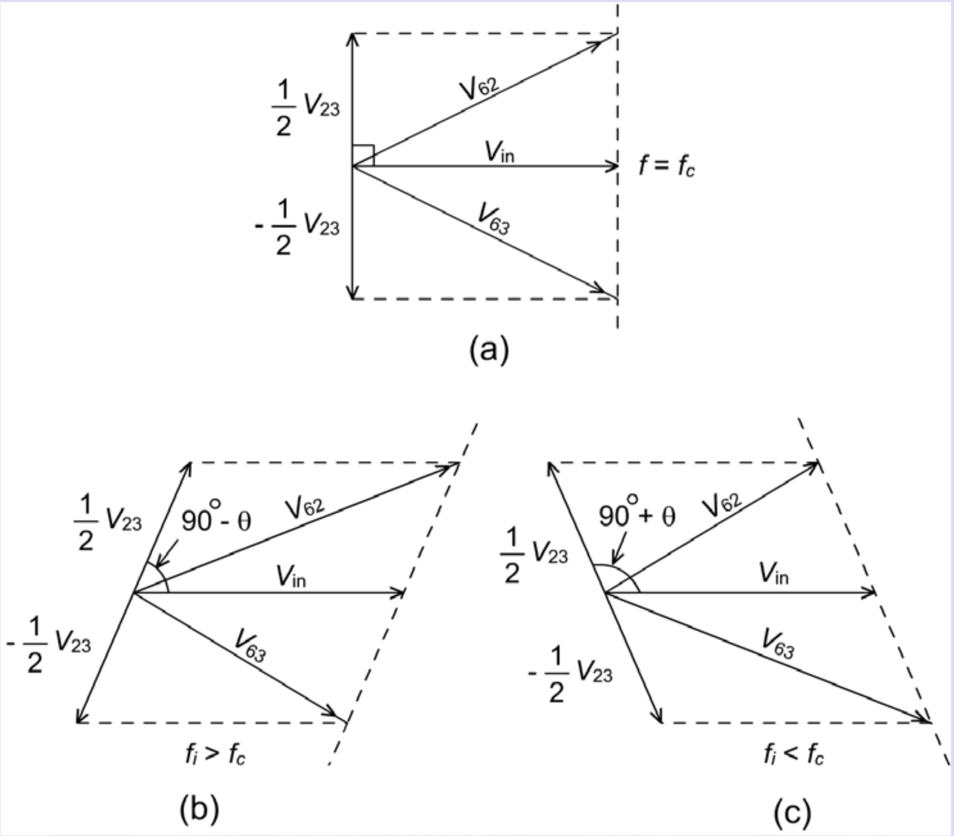


Fig. A5.2.2: Phasor diagram illustrative of the operation of Foster-Seely discriminator

ii) when the input frequency exceeds f_c, $X_2 = X_{L_2} - X_{C_2}$ is positive. Let $R_2 + j X_2 = |Z_2| e^{j\theta}$. Then,

$$V_{23} = \frac{j M}{L_1} \frac{V_{in} X_{C_2}}{R_2 + j X_2} = \frac{V_{in} X_{C_2} M}{L_1 |Z_2|} e^{j(90 - \theta)} \quad \text{(A5.2.5b)}$$

That is, V_{23} leads in V_{in} by less than 90° and $-V_{23}$ lags in V_{in} by more than 90°. This is shown in Fig. A5.2.2(b). As the magnitude of the vector V_{62} is greater than that of V_{63}, $V_{64} > V_{65}$ which implies the final output $V_{54} = V_{64} - V_{65}$ is positive.

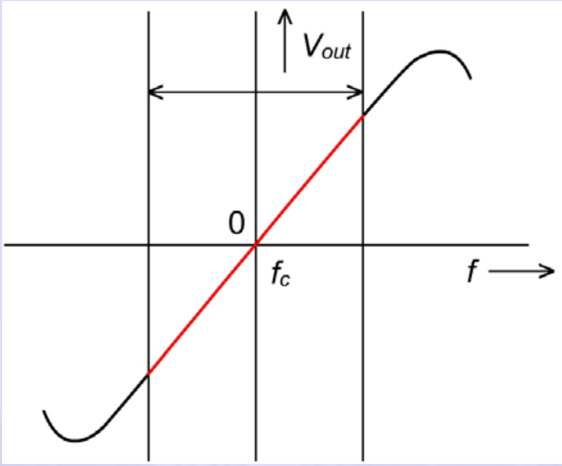


Fig. A5.2.3: Response curve of the Foster-Seely discriminator

iii) Similarly, based on the phasor diagram of Fig. A5.2.2(c), we can easily argue that the final output would be negative when $f < f_c$. The actual value of the final output depends on how far away the input frequency is from f_c. Fig. A5.2.3 gives the plot of the frequency response of the Foster-Seely discriminator, which is usually termed as the S-curve of the discriminator. Useful range of the discriminator (frequency range of linear response, shown in red in Fig. A5.2.3) normally lies between the 3 dB points of the tuned circuit which forms part of the discriminator circuit.

Foster-Seely discriminator responds also to input amplitude variations. Let the input to the discriminator be $f_i(t) = f_c$. Then, the voltages across R_3 and R_4 are equal and let this value be 3 V. Now, let $f_i(t)$ be such that voltage across R_3 increases while that across R_4 decreases. Let the voltage increase on R_3 be 2 Volts. We can take the voltage decrease on R_4 also as 2 V. In other words, for the given frequency deviation, say Δf_1, we have the voltage at point 4 equal to 5 volts where as the voltage at point 5 equal to 1 V. This implies

$V_{out} = (5 - 1) = 4$. Let V_{64} denote voltage across R_3 and V_{65}, the voltage across R_4. Then $\frac{V_{64}}{V_{65}} = 5$.

Let the input signal strength be increased such that, when $f_i(t) = f_c$, $V_{64} = V_{65} = 6$. Now let $f_i(t)$ change such that we have the deviation Δf_1 as in the previous case. Then V_{64} will become 10 Volts whereas V_{65} becomes 2 V, with their difference being equal to 8 V. Though the ratio $\frac{V_{64}}{V_{65}}$ remains at 5, V_{out} changes from the previous value. That is, the circuit responds not only to frequency changes but also to changes in the incoming carrier strength. Hence, Foster-Seely discriminator has to be the preceded by a BPL.

A5.2.2 Ratio Detector

By making a few changes in the Foster-Seely discriminator, it is possible to have a demodulator circuit which has built in capability to handle the amplitude changes of the input FM signal, thereby obviating the need for an amplitude limiter. The resulting circuit is called the *ratio detector* which has been shown in Fig. A5.2.4.

Comparing the ratio detector circuit with that of the Foster-Seely discriminator, we find the following differences: direction of D_2 is reversed, a parallel RC combination consisting of $(R_5 + R_6)$ and C_5 has been added and the output V_{out} is taken across a different pair of points. We shall now briefly explain the operation of the circuit.

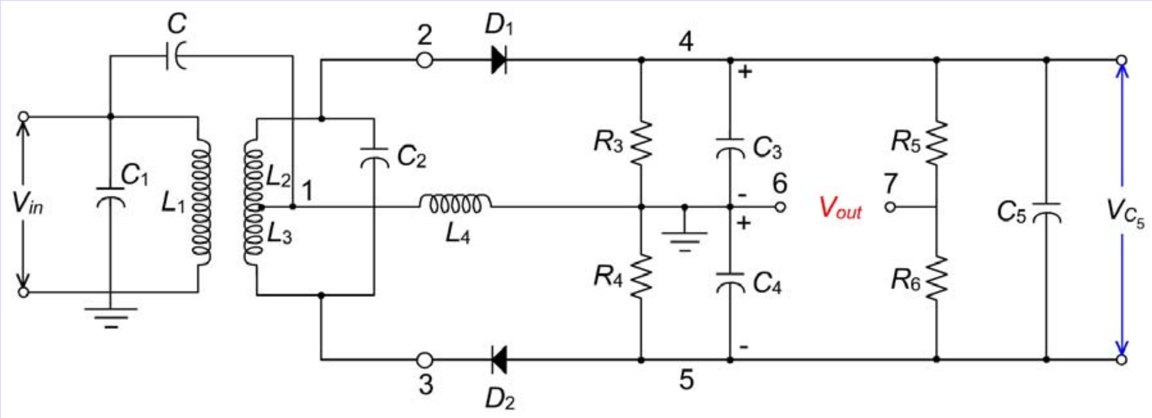


Fig. A5.2.4: Circuit schematic of a ratio detector

Reexamining Fig.A5.2.1 and the corresponding phasor diagrams, we find that by and large, the sum $V_{62} + V_{63}$ remains constant. Hence, any variation in the magnitude of this sum voltage can be considered to be spurious. Suppression of these spurious variations would result in a detector that is unaffected by input voltage fluctuations which implies that circuit does not require a separate limiter stage. How the sum voltage is kept constant would be explained a little later.

With the diode D_2 being reversed, we find that the voltages V_{64} and V_{65} are series aiding rather than series opposing and as such, the voltage V_{54} represents the sum voltage. Taking $R_5 = R_6$, we find

$$
\begin{aligned}
V_{out} &= V_{64} + V_{47} = V_{64} - V_{74} \\
&= V_{64} - \frac{1}{2} V_{54} \\
&= V_{64} - \frac{V_{56} + V_{64}}{2} \\
&= \frac{1}{2}\left[V_{64} - V_{56}\right] \\
&= k\left[\left|V_{62}\right| - \left|V_{63}\right|\right]
\end{aligned}
$$

Usually, $C_3 = C_4$ and $R_3 = R_4$. Hence at resonance, $V_{64} = V_{56}$ which implies that V_{out} is zero. Above resonance, as $V_{64} > V_{56}$, the output is positive whereas below resonance $V_{56} > V_{64}$, and the output is negative.

In the circuit Fig. A5.2.4, C_5 is a capacitor of a rather large value. For example, C_5 is of the order of 5 μF whereas C_3 and C_4 are of the order 300 pF. If V_{in} is constant, C_5 charges to the full potential existing between the points 5 and 4, which, as indicated earlier is essentially a constant. If V_{in} tries to increase, C_5 will tend to oppose any rise in V_{out}. This is because as the input voltage tries to rise, extra diode current flows trying to charge C_5. But V_{54} remains constant at first because C_5 is a fairly large capacitance and it is not possible for the voltage across it to change instantaneously. The situation now is that the current in the diodes' load has risen but the voltage across the load has not changed. This being so, the secondary of the ratio detector transformer is more heavily damped, Q falls and so does the gain of the amplifier driving the detector. This nearly counteracts the rise in the input voltage. Similarly, when V_{in} increases, the damping is reduced. The gain of the driving amplifier increases thereby counteracting the fall in the input voltage. Thus the ratio detector provides *variable damping.*

For a large number of years, the Foster-Seely discriminator and the ratio detector have been the work horses of the FM industry. As these circuit configurations are not very convenient from the point of view of IC fabrication, of late, their utility has come down. Companies such as Motorola have built high quality FM receivers using the Foster-Seely discriminator and the ratio detector. Some details can be found in Roddy and Coolen [2].

Appendix A5.3

Multi-tone FM

Let $m(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t$ where f_1 and f_2 are arbitrary. Then,

$$s_{pe}(t) = A_c \left[e^{j \beta_1 \sin(\omega_1 t)} \right] \left[e^{j \beta_2 \sin(\omega_2 t)} \right] e^{j \omega_c t}$$

where $\beta_1 = \dfrac{A_1 k_f}{f_1}$ and $\beta_2 = \dfrac{A_2 k_f}{f_2}$

$$s_{pe}(t) = A_c \left\{ \left[\sum_m J_m(\beta_1) e^{j m \omega_1 t} \right] \left[\sum_n J_n(\beta_2) e^{j n \omega_2 t} \right] \right\} e^{j \omega_c t}$$

Hence, $s(t) = A_c \sum_m \sum_n J_m(\beta_1) J_n(\beta_2) \cos\left[(\omega_c + m \omega_1 + n \omega_2) t \right]$ (A5.3.1)

The (discrete) spectrum of $s(t)$ can be divided into 4 categories:

1) Carrier component: amplitude $= J_0(\beta_1) J_0(\beta_2)$ when $f = f_c$

2) A set of side frequency components due to f_1: These components have amplitudes $J_m(\beta_1) J_0(\beta_2)$ at frequencies $(f_c \pm m f_1)$, $m = 1, 2, 3, \cdots$

3) A set of side frequency components due to f_2: These components have amplitudes $J_0(\beta_1) J_n(\beta_2)$ at frequencies $(f_c \pm n f_2)$, $n = 1, 2, 3, \cdots$

4) A set of cross modulation (or beat frequency) terms with amplitudes $J_m(\beta_1) J_n(\beta_2)$ at frequencies $(f_c \pm m f_1 \mp n f_2)$ where $m = 1, 2, 3, \cdots$; $n = 1, 2, 3, \cdots$

Cross spectral terms of the type given at (4) above clearly indicate the non-linear nature of FM. These terms are not present in linear modulation. Even with respect to the terms of the type (2) and (3), linear modulation generates only those components with $m = 1$ and $n = 1$.

Appendix A5.4

RMS Bandwidth of WBFM

We have already used a number of measures for the bandwidths of signals and systems, such as 3-dB bandwidth, null-to-null bandwidth, noise equivalent bandwidth, etc. In the context of WBFM, another meaningful and useful bandwidth quantity is the **rms bandwidth**. The basic idea behind the rms bandwidth is as follows.

Let $S_X(f)$ denote the power spectral density of a random process $X(t)$. Then, the normalized PSD, $S_{X,N}(f) = \dfrac{S_X(f)}{\int_{-\infty}^{\infty} S_X(f)\,df}$, has the properties of a PDF. Let σ_f denote the standard deviation of $S_{X,N}(f)$. Then, $2\,\sigma_f$ can be used as a measure of the spectral width of the process.

The WBFM process, with $f_c >> \Delta f$, can be treated as a band-pass random process. We shall define the *rms bandwidth* of a band-pass process as

$$B_{rms} = 2\left(\overline{(f - f_0)^2}\right)^{\frac{1}{2}} \tag{A5.4.1a}$$

or
$$B_{rms}^2 = 4\,\overline{(f - f_0)^2} \tag{A5.4.1b}$$

where
$$\overline{(f - f_0)^2} = \frac{\int_0^{\infty} (f - f_0)^2 S_X(f)\,df}{\int_0^{\infty} S_X(f)\,df}$$

$$= \frac{\int_0^{\infty} (f - f_0)^2 S_X(f)\,df}{P_T/2} \tag{A5.4.2a}$$

where P_T is the total power of the process and

$$f_0 = \frac{2}{P_T} \int_0^{\infty} f \, S_X(f) \, df \tag{A5.4.2b}$$

Let $M(t)$ be a strict sense stationary message process with $m(t)$ as a sample function. Let $p_M(m)$ denote the PDF of the process. (Note that we are using the symbol $p(\ )$ to denote the PDF, instead of the earlier symbol $f(\ )$; this has been done to avoid confusion. In this derivation, f denotes only the frequency variable.) For the FM case, we have

$$f_i = f_c + k_f m$$

where m is a specific value of $m(t)$ for some t. Then

$$m = \frac{(f_i - f_c)}{k_f} \tag{A5.4.3}$$

Using *quasi-static* approximation, it has been shown by Peebles [4] that,

$$\left[S_X(f)\right]_{WBFM} = \frac{P_T}{2k_f}\left[p_M\left(\frac{f - f_c}{k_f}\right) + p_M\left(\frac{f + f_c}{k_f}\right)\right] \tag{A5.4.4}$$

(Note that in quasi-static approximation, it is assumed that $f_i(t)$ remains constant for a sufficiently long period; as such, FM wave appears to be a regular sinusoid and f_i can be replaced by f.) For the WBFM process, $P_T = S_T = \frac{A_c^2}{2}$.

In Eq. A5.4.4, $p_M\left(\frac{f - f_c}{k_f}\right)$ is the positive part of the spectrum. We will now show that $(B_{rms})_{WBFM} = 2k_f\left[R_M(0)\right]^{\frac{1}{2}}$ where $R_M(0) = \overline{M^2(t)}$, the mean square value of the process. (Note that $R_M(\tau)$ is the ACF of the process.)

For $f > 0$, $S_X(f)$ is symmetrical with respect to $f = f_c$ (we assume that the input PDF is symmetric about zero) and hence $f_0 = f_c$. Using Eq. A5.4.4 in A5.4.2(a), we have

$$\left(B_{rms}^2\right)_{WBFM} = \frac{8}{P_T} \int_0^{\infty} (f - f_c)^2 \cdot \frac{P_T}{2 k_f} \, p_M\left(\frac{f - f_c}{k_f}\right) df .$$

Let $\frac{f - f_c}{k_f} = \lambda$. Then,

$$\left(B_{rms}^2\right)_{WBFM} = 4 \int_{-\frac{f_c}{k_f}}^{\infty} (\lambda k_f)^2 \, p_M(\lambda) \, d\lambda$$

$$\simeq 4 k_f^2 \int_{-\infty}^{\infty} \lambda^2 \, p_M(\lambda) \, d\lambda = 4 k_f^2 \, \overline{M^2(t)}$$

$$= 4 k_f^2 \, R_M(0)$$

That is,

$$\left(B_{rms}\right)_{WBFM} = 2 k_f \sqrt{R_M(0)} \qquad \text{(A5.4.5)}$$

Example A5.4.1

Let $m(t)$ be a sample function of a strict sense stationary process $M(t)$. A WBFM signal is generated using $m(t)$ as the message signal. It is given that

$$p_M(m) = \begin{cases} \frac{1}{2}, & |m| < 1 \\ 0, & otherwise \end{cases}$$

a) Find $\left[S_X(f)\right]_{WBFM}$.

b) What is the value of B_{rms}?

As $p_M(m)$ is uniform, we expect the PSD of the resulting WBFM also to be uniform over the appropriate frequency range. When $m = 0$, we have $f = f_c$, whereas $m = \pm 1$ results in the instantaneous frequency to be $f_c \pm k_f$ respectively. In other words, the frequency range over which $S_M(f)$ exists is $f_c - k_f < |f| < f_c + k_f$. Let $p_M\left(\frac{f - f_c}{k_f}\right) + p_M\left(\frac{f + f_c}{k_f}\right)$ be as shown in Fig. A5.4.1.

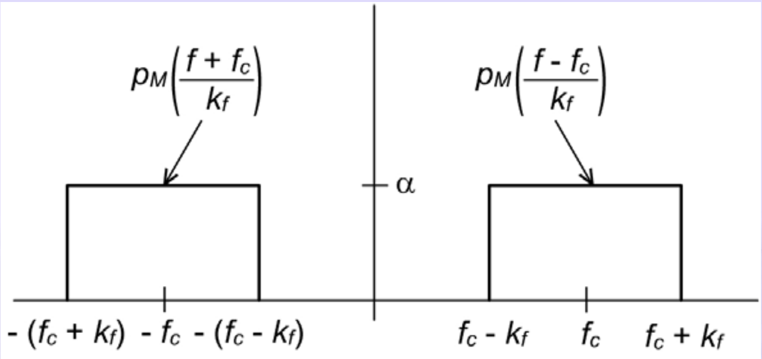


Fig. A5.4.1: Shape of the PSD of Example A5.4.1

To find the value of α, we require that the RHS of Eq. A5.4.4, when integrated over the entire range should be equal to $\frac{A_c^2}{2}$. That is,

$$2\left[\frac{A_c^2}{4k_f}\int_{f_c - k_f}^{f_c + k_f} \alpha \, df\right] = \frac{A_c^2}{2}$$

or $$\frac{A_c^2}{2k_f}\cdot \alpha(2k_f) = \frac{A_c^2}{2}$$

That is, $\alpha = \frac{1}{2}$

Hence, $$\left[S_X(f)\right]_{WBFM} = \begin{cases} \frac{A_c^2}{8k_f}, & f_c - k_f < |f| < f_c + k_f \\ 0 & , otherwise \end{cases}$$

In a few situations (B_{rms}) might be quite meaningful. For example, if $m(t)$ is the sample function of a Guassian process, there is a small but finite probability that $m(t)$ will assume very large values. This implies that Δf becomes excessive and as such B_T given by Carson's rule or its variants would be extremely high. As (B_{rms}) weights the large frequency derivations with small probabilities, (B_{rms}) will not be unduly excessive. For example if $m(t)$ belongs to a Guassian process and say $p_M(m)$ is $N(0, 9)$, then

$$(B_{rms})_{WBFM} = 2 \times 3 \times k_f = 6 k_f$$

(Note that $\sqrt{R_M(0)} = \sqrt{9} = 3$)

Exercise A5.4.1

We define the RMS bandwidth of any lowpass process $M(t)$ as,

$$\left[(B_{rms})_M\right]^2 = \frac{\int_{-\infty}^{\infty} f^2 S_M(f)\, df}{R_M(0)} \quad \text{(A5.4.6)}$$

Let $(B_{rms})_{PM}$ denote the RMS bandwidth of the PM signal. Show that

$$(B_{rms})_{PM} = 2 k_p \sqrt{R_M(0)} \left[(B_{rms})_M\right] \quad \text{(A5.4.7)}$$

Appendix A5.5

Modulation Techniques in TV

Color television is an engineering marvel. At the flick of a button on a remote, we have access to so much information and entertainment in such a fine detail (audio as well as video with all its hues and colors) that the angels could envy the humans on this count. Distances are no longer a barrier; we have real time reception almost at any point on this globe. Maybe, not too far into the future, we may be able to watch any program of our choice in the *language we prefer* to listen either in real time or near real time. We shall now take a closer look at the modulation techniques used in (commercial) TV, based on the NTSC standard used in North America and Japan. Our purpose here is to illustrate how various modulation techniques have been used in a practical scheme. Hence PAL and SECAM systems have not been discussed.

As the TV transmission and reception started with monochrome (black and white) signals we shall begin our discussion with this scheme. (Note that color transmission can be viewed on a monochrome receiver. Similarly, black-and-white transmission can be viewed on a color receiver.)

The bandwidth of a monochrome video signal is about 4.2 MHz. Let $m_v(t)$ denote this signal. The bandwidth allocated to each TV station by the regulatory body is about 6 MHz. Hence the use of DSB is ruled out. It is very difficult to generate the SSB of the video signal, $m_v(t)$. (Filtering method is ruled out because of appreciable low frequency content in $M_v(f)$ and because of the fairly wide bandwidth, designing the HT with required specifications in extremely difficult.) As such, VSB becomes the automatic choice. With a suitable carrier component, we have seen that VSB can be envelope detected. Actually at the transmitter, it is only partially VSB; because of the high power levels at the transmitter, it is difficult to design a filter with an exact vestigial sideband. It is at the IF stage of the receiver that perfect VSB shaping is achieved; subsequently it is demodulated.

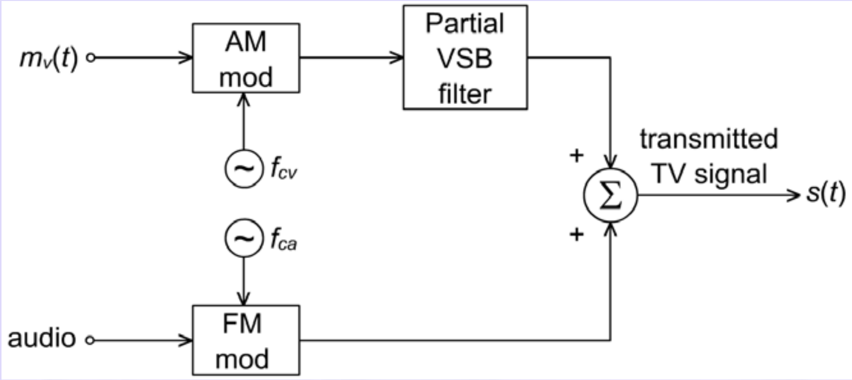


Fig. A5.5.1: Modulation stages in a monochrome TV transmitter

Fig. A5.5.1 shows the block diagram of the modulation scheme of a monochrome TV transmitter. Details of the magnitude spectrum of the transmitted signal is given in Fig. A5.5.2.

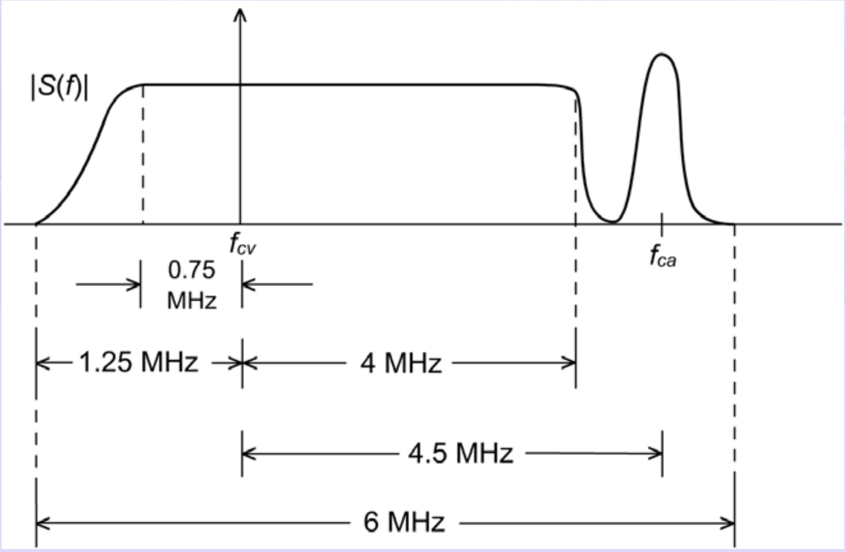


Fig. A5.5.2: Audio and video spectrum of a monochrome TV signal

In this figure, f_{cv} is the video carrier and f_{ca} is the carrier for audio, which is higher than f_{cv} by 4.5 MHz. (The video carrier f_{cv} is 1.25 MHz above the lower frequency limit allotted to the station)

The USB of the video occupies the full bandwidth of 4.2 MHz. That part of LSB spectrum between $(f_{cv} - 0.75)$ to f_{cv} is transmitted without any attenuation. Below the frequency $(f_{cv} - 0.75)$ MHz, LSB is gradually attenuated reaching almost zero level at $(f_{cv} - 1.25 \text{ MHz})$. As this is not exactly a VSB characteristic, it is termed as partial VSB.

The audio signal, band-limited to 10 kHz frequency modulates the audio carrier. The maximum frequency deviation is 25 kHz. Hence, the audio bandwidth can be taken as $2(\Delta f + W) = 2(25 + 10) = 70$ kHz.

TV receiver is of the superheterodyne variety. A part of the reciever structure is shown in Fig. A5.5.3.

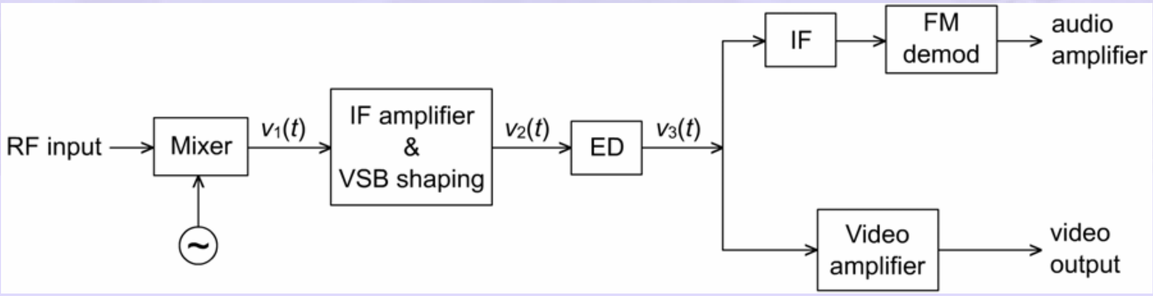


Fig. A5.5.3: (Partial) Block diagram of monochrome TV receiver

The mixer output, $v_1(t)$, is applied to an IF amplifier and VSB shaping network. The IF amplifier has a pass-band of 41 to 47 MHz. The characteristic of the VSB shaping filter is shown in Fig. A5.5.4.

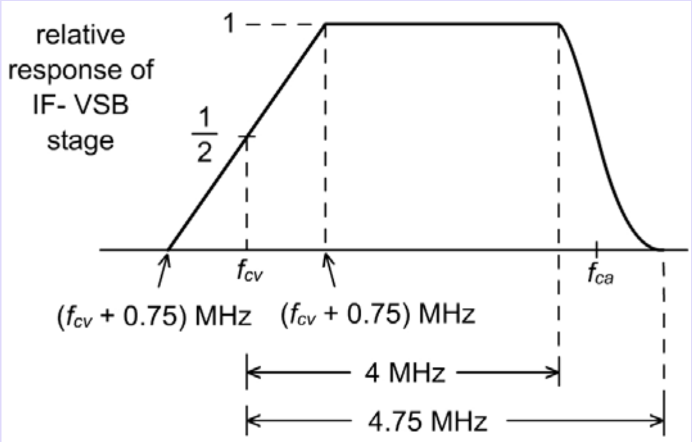


Fig. A5.5.4: VSB shaping in TV receiver

The frequency modulated audio signal is also passed by the IF stage but with much less gain than the video part. That is, $v_2(t)$ consists of the video signal of the (VSB+C) type and the FM audio signal, with a carrier at 4.5 MHz and with $A_{ca} << A_{cv}$ where A_{ca} is the audio carrier amplitude and A_{cv} is the amplitude of the video carrier. Under these conditions, it can be shown that $v_3(t)$, the output of the envelope detector, has the video as well as required audio. The video amplifier removes the audio from $v_3(t)$. The output of the video amplifier is processed further and is displayed on the picture tube. $v_3(t)$ is also applied as input to an IF stage, with the IF of 4.5 MHz. The audio part of the $v_3(t)$ is passed by this IF stage; the FM demodulator that follows produces the audio signal.

Color TV: The three primary colors, whose linear combination can give rise to other colors are: Red, Blue and Green. These color signals, pertaining to the scene that is being shot, are available at the outputs of three color cameras. Let us denote these signals as $m_R(t)$, $m_G(t)$ and $m_B(t)$ respectively. These basic color components are linearly combined to produce (i) the video signal of the monochrome variety (this is called the luminance signal and is denoted by $m_L(t)$)) and (ii) two other independent color signals (called the in-phase component of the color signal and the quadrature component of the color signal).

Let us denote these components as $m_I(t)$ and $m_Q(t)$ respectively. The equations of the linear transformation are given below in the form of a matrix equation.

$$\begin{bmatrix} m_L(t) \\ m_I(t) \\ m_Q(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0.3 & 0.59 & 0.11 \\ 0.6 & -0.28 & -0.32 \\ 0.21 & -0.52 & 0.31 \end{bmatrix}}_{M} \begin{bmatrix} m_R(t) \\ m_G(t) \\ m_B(t) \end{bmatrix} \tag{A5.5.1}$$

From Eq. A5.5.1, we obtain

$$m_L(t) = 0.3\, m_R(t) + 0.59\, m_G(t) + 0.11 m_B(t) \tag{A5.5.2}$$

It has been found that $m_L(t)$ as given by Eq. A5.5.2 closely resembles $m_v(t)$ of the monochrome system. $m_L(t)$ has the same bandwidth as $m_v(t)$ namely, 4.2 MHz. This is required to preserve the sharp transitions in the intensity of light at the edges in a scene. However, the eye is not as sensitive to color transitions in a scene and is possible to reduce the bandwidth occupancy of $m_I(t)$ and $m_Q(t)$. In the NTSC system, bandwidth allocation for $m_I(t)$ is 1.5 MHz and that of $m_Q(t)$ is 0.5 MHz. These chrominance signals are quadrature multiplexed (QAM) on the color subcarrier as shown in Fig. A5.5.5.

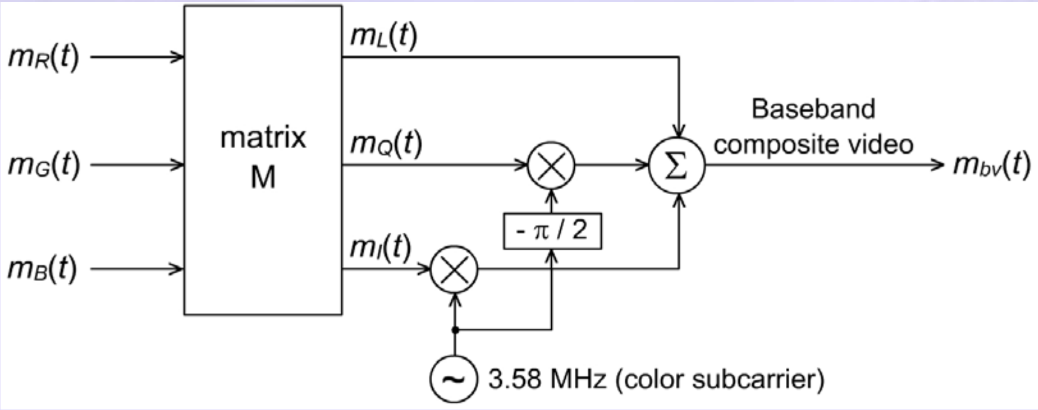


Fig. A5.5.5: Generation of composite baseband video

The composite video signal $m_{bv}(t)$ is modulated on to the carrier f_{cv}. This implies that color subcarrier, f_{cc}, is 3.58 MHz above f_{cv} and f_{ca} is 4.5 MHz above f_{cv}. As such, the color information gets interleaved in between the spectral lines of the luminance signal and no additional bandwidth is required for the color TV system. In order to facilitate coherent demodulation of the QAM signal, a few cycles of the color subcarrier (called color burst) is sent along with the transmitted signal. This reference carrier is tracked by a PLL in the receiver. The VCO output of the PLL is used in the demodulation of the chrominance signal. For more details on TV transmission and reception, refer to Carlson, Crilly and Rutledge [5] or Leon Couch [6].

TV modulator circuit is available in an IC chip. For example, Motorola MC 1374 includes an FM audio modulator, sound carrier oscillator, RF dual input modulator. It is designed to generate a TV signal from audio and video inputs. It is also suited for applications such as video tape recorders, video disc players, TV games etc. The FM system can also be used in the base station of a cordless telephone. Circuit details and other parameters can be obtained from the manual.

Exercise A5.5.1

Indirect method can used to generate the FM signal for the audio in TV. Required carrier is 4.5 MHz and $\Delta f = 25$ kHz. Using $f_{c1} = 200$ kHz and $\Delta f_1 < 20$ Hz, design the modulator such that frequency at any point in the modulator does not exceed 100 MHz. Use the shortest possible chain of frequency doublers and triplers.

References

1) Bettina Koster, Peter Waldow and Ingo Wolff, A unique low voltage, source-coupled J-FET VCO, RF signal processing, April 2001, PP58-66 (website: http://www.rfdesign.com)
2) Dennis Roddy and John Coolen, Electronic communications, (4th ed.), PHI, 1995
3) Herbert Taub and Donald L. Schilling, Principles of communication systems, (2nd ed.), McGraw Hill International ed., 1986
4) Peebles P. Z., Communication system principles, Addison-Wesley, 1976
5) A. Bruce Carlson, Paul B. Crilly and Janet C. Rutledge, Communication systems (4th ed.), McGraw Hill International ed., 2002
6) Leon W. Couch II, Digital and Analog Communication systems (6th ed.), Pearson Education-Asia, 2001