Chapter 3

Analog Modulation System

3.1 Introduction

Analog modulation denotes the modulation of a carrier by utilizing a source baseband analog signal. Carrier is a deterministic periodic waveform. The carrier waveform discussed in this chapter is a cosine waveform, and its mathematical expression is

$$c(t) = A\cos(\omega_0 t + \varphi_0) \tag{3.1-1}$$

where A is amplitude; ω_0 is carrier angular frequency; and φ_0 is initial phase.

A carrier has three parameters: amplitude A, carrier angular frequency ω_0, and initial phase φ_0. Modulation will enable certain parameter of a carrier to vary with the signal, in other words, the value of signal from the source is represented by the value of a certain parameter of the carrier. Signal from the source is called modulating signal $m(t)$ thereafter, and the carrier after being modulated is called modulated signal $s(t)$. The device for modulation is called modulator (see Fig. 3.1.1). Modulation has the following two purposes:

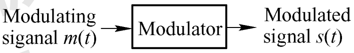


Figure 3.1.1 Modulator

(1) frequency spectrum of a baseband signal can be moved near the carrier frequency by modulation. Thus, baseband signal is transformed to bandpass signal. The frequency spectrum of a signal may be moved to the desired frequency band by selecting the required carrier frequency. Such a frequency spectrum movement is either for accommodating the requirement of channel transmission or for combining several signals for multichannel transmission.

(2) The anti-jamming ability of the signal transferred through channels can be improved by modulation. At the same time, modulation not only affects anti-jamming ability, but is also related to transmission efficiency. Specifically speaking, bandwidths of modulated signals produced by different modulation modes are different, hence affecting the utilization of transmission bandwidth.

Analog modulation can be classified into two kinds: linear modulation and nonlinear modulation. Frequency spectrum structure of a linear modulated signal is the same as that of modulating signal. In other words, frequency spectrum of modulated signal is the result of displacement of modulating signal along frequency axis. Linear modulated signals include amplitude modulated signal, single-sideband signal, double-sideband (suppressed carrier) signal, and vestigial side band signal. Nonlinear modulation is also called angle modulation. Frequency spectrum structure of a nonlinear modulated signal is much different from that of a modulating signal. In addition to the displacement of frequency spectrum, there are many other new frequency components. The bandwidth occupied may be much increased. Nonlinear modulated signals include frequency modulated signal and phase modulated signal.

Now the application of analog modulation in long distance transmission has gradually decreased thanks to the superiority of digital communication technology and its rapid development. However, analog modulation is still the basic modulation mode, and we need to have a basic understanding of it.

A brief introduction of the above amplitude and angle modulations is given in this chapter.

3.2 Linear Modulation

Let the carrier be

$$c(t) = A\cos\omega_0 t = A\cos 2\pi f_0 t \tag{3.2-1}$$

where A is amplitude (V); f_0 is frequency (Hz); $\omega_0 = 2\pi f_0$ is angular frequency (rad/s).

In the above definition equation of a carrier, it is already assumed that the initial phase is 0. Such assumption doesn't influence the generality of our discussion. In addition, it is assumed that modulating signal is $m(t)$, and modulated signal is $s(t)$.

The principle model of a linear modulator is shown in Fig. 3.2.1. Modulating signal $m(t)$ and the carrier are multiplied in the multiplier in the figure, and the result of multiplication is

$$s'(t) = m(t)A\cos\omega_0 t \tag{3.2-2}$$

Then it passes through a bandpass filter with transfer function $H(f)$, and obtains modulated signal $s(t)$.

Now assume the modulating signal is an energy signal, its frequency spectral density is $M(f)$, and there is Fourier transform relationship between them, as well as use "$\Longleftrightarrow$" to express the Fourier transform, then we have

$$m(t) \Longleftrightarrow M(f) \tag{3.2-3}$$

$$m(t)A\cos\omega_0 t \Longleftrightarrow S'(f) \tag{3.2-4}$$

where $$S'(f) = \frac{A}{2}[M(f-f_0) + M(f+f_0)] \tag{3.2-5}$$

$S'(f)$ is the frequency spectral density of $s'(t)$.

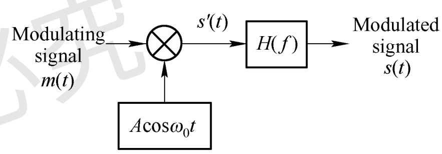
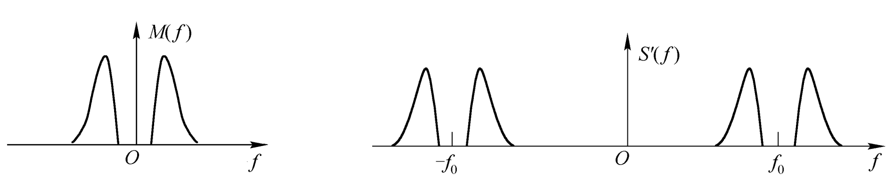


Figure 3.2.1 Principle model of linear modulator

As can be seen from eq.(3.2-2), the output signal $s'(t)$ of the multiplier is a cosinusoidal wave, the amplitude of which is proportional to $m(t)$, i.e., the amplitude of the carrier waveform is modulated. In addition, as can be seen from eq.(3.2-5), frequency spectral density $S'(f)$ of the output signal of the multiplier is the displacement result of frequency spectral density $M(f)$ of the modulating signal (there is a difference in a constant factor), as shown in Fig. 3.2.2. This is called linear modulation because the relationship between modulating signal $m(t)$ and output signal of the multiplier is linear.

The characteristic $H(f)$ of the bandpass filter in Fig. 3.2.1 may have different designs, so different modulation modes are resulted. They will be introduced respectively in the following paragraphs.



(a) Frequency spectral density of input signal (b) Frequency spectral density of output signal

Figure 3.2.2 Frequency spectra of input signal and output signal of the multiplier

3.2.1 Amplitude Modulation (AM)

Assume modulating signal $m(t)$ contains D. C. component, and its expression is written as $[1 + m'(t)]$, where $m'(t)$ is the A. C. component in the modulating signal, and $|m'(t)| \leqslant 1$. The maximum of $|m'(t)|$ is called modulation index m, and $m \leqslant 1$. Thus, eq. (3.2-2) of the multiflier output signal may be rewritten as

$$s'(t) = [1 + m'(t)]A\cos\omega_0 t \tag{3.2-6}$$

As can be seen from the above equation, the envelope of $s'(t)$ contains a D. C. component A, and an A. C. component $m'(t)A$ which is superposed on the basis of A. The envelope of $s'(t)$ isn't less than 0, i. e., the envelope cannot be negative (see Fig. 3.2.3) because the absolute value of $m'(t)$ isn't larger than 1. Here, if transfer function $H(f)$ of the filter allows frequency spectral density $S'(f)$ of $s'(t)$ to pass through without distortion, then output signal $s(t)$ of the modulator is just an amplitude modulation signal. For modulating signal without D. C. component, in order to obtain amplitude modulation, other simpler modulator circuit is often used, and the method of adding D. C. component is not used.

There are discrete carrier components in the frequency spectral density of a amplitude modulated signal, and they are shown in Fig. 3.2.3 by arrows. Now let us consider the ratio of carrier power to sideband power in the modulated signal. If modulating signal $m'(t)$ is a cosinusoidal wave $\cos\Omega t$, then it is not difficult to prove that when modulation index m is maximum (equals 100%), the sum of the powers of two sidebands of the modulated signal equals one half of the carrier power. (left as exercise) That is to say, most part of the power in such a modulated signal is occupied by the carrier, and the carrier itself doesn't contain the information of the baseband signal. Therefore, the carrier is not necessary to be transmitted. Hence the double-sideband modulation is resulted which will be discussed in section 3.2.2.

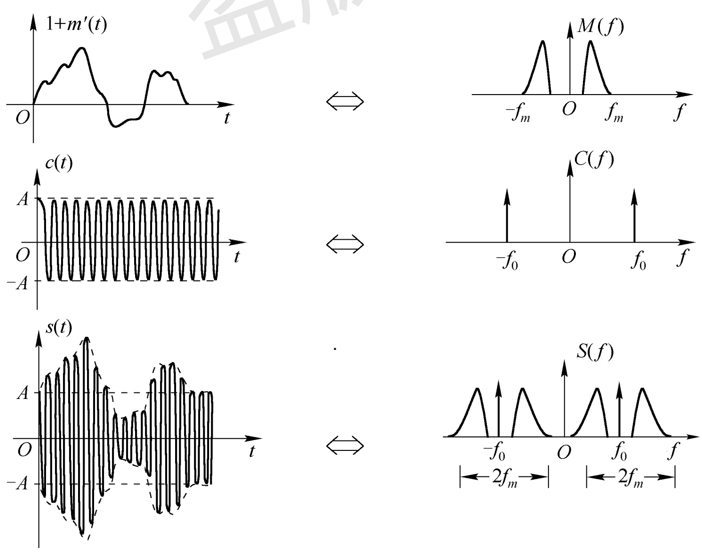


Figure 3.2.3 Waveform and spectrum of amplitude modulated signal

It is not difficult to know from the waveform of a amplitude modulated signal (AM signal, for short), the shape of the envelope of the modulated signal is the same as the shape of the modulating signal. So, envelope detection method can be used to restore the original modulating signal during demodulation at the receiver. Envelope detector can consist of a rectifier and a low-pass filter, as shown in Fig. 3. 2. 4. Related waveforms are also drawn in this figure. There is a blocking D. C. circuit (expressed by a capacitor) at the output to separate the D. C. component in the output of the rectifier because the D. C. component can pass through the low-pass filter.

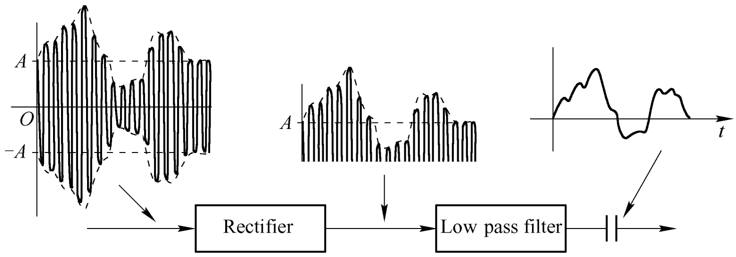


Figure 3. 2. 4 Envelope detector

Now performance of AM signal demodulated by envelope detector will be discussed. Assume the input voltage of the envelope detector is

$$y(t) = \{[1 + m'(t)]A + n_c(t)\}\cos\omega_0 t - n_s(t)\sin\omega_0 t \tag{3.2-7}$$

where $n_c(t)\cos\omega_0 t - n_s(t)\sin\omega_0 t$ is the input noise voltage of the detector (see eq.(2. 8-2)).

Hence, the envelope of $y(t)$ is

$$V_y(t) = \sqrt{\{[1 + m'(t)]A + n_c(t)\}^2 + n_s^2(t)} \tag{3.2-8}$$

Under the condition of large signal to noise ratio, the above equation may approximately be written as

$$V_y(t) \approx [1 + m'(t)]A + n_c(t) \tag{3.2-9}$$

Therefore the output voltage of the AM signal after detection is (D. C. component has been blocked in the following equation)

$$v(t) = m'(t)A + n_c(t) \tag{3.2-10}$$

Here, the signal to noise power ratio of the output is

$$r_o = E[m'^2(t)A^2/n_c^2(t)] \tag{3.2-11}$$

And the signal to noise power ratio before detection is

$$r_i = E\left\{\frac{1}{2}[1 + m'(t)]^2 A^2/n^2(t)\right\} \tag{3.2-12}$$

Hence the ratio of signal to noise power ratios before and after detection is

$$\frac{r_o}{r_i} = E\left\{\frac{m'^2(t)A^2/n_c^2(t)}{\frac{1}{2}[1 + m'(t)]^2 A^2/n^2(t)}\right\} = E\left[\frac{2m'^2(t)}{[1 + m'(t)]^2}\right]$$

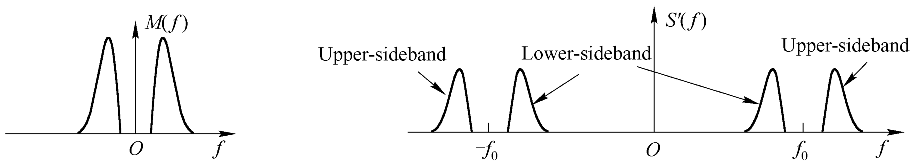
During calculation of the above equation, the characteristic of narrow band random process described in Section 2. 8. 2 has been used: $E[n_c^2(t)] = E[n^2(t)]$.

Obviously, r_0/r_i is less than 1 for $m'(t) \leqslant 1$, i. e., signal to noise ratio has decreased after detection. This is because most part of power in signal before detection is occupied by the carrier, and it makes no contribution to the useful detected signal.

Although there are better demodulation methods than envelope detection, for instance coherent demodulation method introduced in Chapter 6, envelope detector is often used for detection of AM signal due to the fact that it's simple and cheap.

3. 2. 2 Double-sideband Modulation (DSB)

If modulating signal $m(t)$ in the linear modulator (Fig. 3. 2. 1) has no D. C. component, then there will be no carrier in the output signal of the multiplier. The frequency spectrum of the modulated signal now is shown in Fig. 3. 2. 5(b). This kind of modulation is called double-sideband modulation (DSB); the reason is that its frequency spectrum contains two sidebands, and same information is contained in the two sidebands. Its full term is double-sideband suppressed carrier AM. The two sidebands are called upper-sideband and lower-sideband respectively (see Fig. 3. 2. 5); the sideband located higher than carrier frequency is called upper-sideband, and the one lower than carrier frequency is called lower-sideband.



(a) Frequency spcctral density of modulating signal (b) Frequency spectral density of modulated signal

Figure 3. 2. 5 Spectrum of double-sideband modulation signal

Since the carrier of the DSB signal is not transmitted, power of the carrier may be saved. However, it is necessary to have a carrier in the receiver for demodulation, and the frequency and phase of the carrier generated in the receiver should be the same as that in the transmitter. Hence, the receiver circuit will be rather complex. Fig. 3. 2. 6 shows a block diagram of this kind of demodulator.

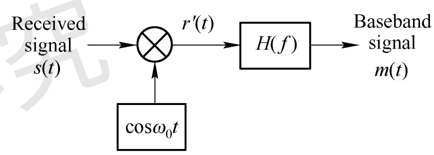


Figure 3. 2. 6 Block diagram of DSB signal demodulator

Assume the received DSB signal is $m'(t)\cos\omega_0 t$, and the frequency and the phase of the local carrier at the receiver both have certain error, i. e., assume its expression is $\cos[(\omega_0+\Delta\omega)t+\varphi]$, then the product of the received signal and the local carrier is

$$\begin{aligned} r'(t) &= m'(t)\cos\omega_0 t\cos[(\omega_0+\Delta\omega)t+\varphi] \\ &= \frac{1}{2}m'(t)\{\cos(\Delta\omega t+\varphi)+\cos[(2\omega_0+\Delta\omega)t+\varphi]\} \end{aligned} \quad (3.\ 2\text{-}13)$$

The second term in the above equation is the component with the frequency $(2\omega_0+\Delta\omega)$, and can be filtered out by the low-pass filter $H(f)$. So, the demodulated output signal is $\frac{1}{2}m'(t)\cos(\Delta\omega t+\varphi)$.

The output signal equals $m'(t)/2$, only if the local carrier has no frequency and phase error, i. e., $\Delta\omega=\varphi=0$. Under this condition, demodulated output signal has no distortion, and the difference in comparison with the modulating signal is only one constant factor (1/2).

3. 2. 3 Single-Sideband Modulation (SSB)

Two sidebands in DSB modulation contain the same information, so there is no need to transmit two sidebands. Therefore, one sideband among them may be filtered out by the filter in the linear modulator, and only the other sideband is transmitted. This is called the single-sideband modulation. In order that the filter can practically separate the upper-sideband and the lower-sideband, the spectrum of a modulating signal is not allowed to have too low frequency components, since the filter cannot have sharp cutoff characteristic. The spectrum sketch of a single-sideband signal is shown in Fig. 3. 2. 7. A high-pass filter with transfer characteristic $H_H(f)$ used in the modulator can give upper-sideband signal; and a low-pass filter with transfer characteristic $H_L(f)$ used can give lower-sideband signal.

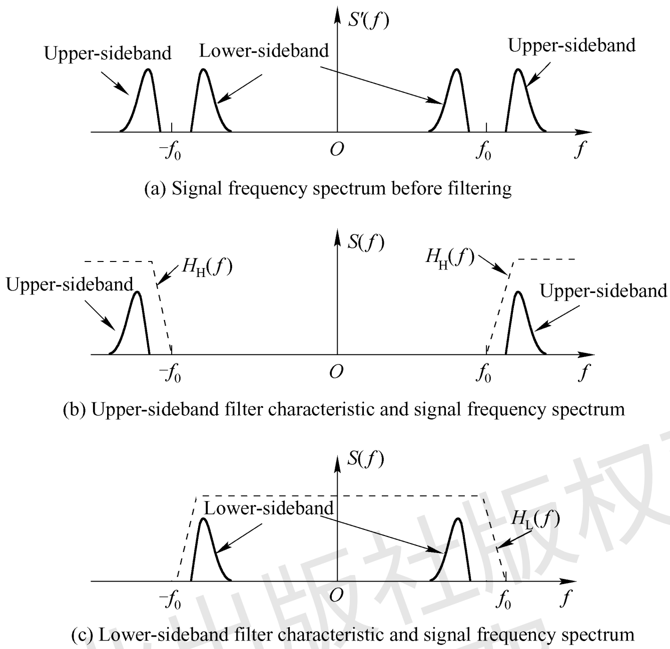


Figure 3. 2. 7 Frequency spectrum of single-sideband signal

It is necessary to add a carrier during the demodulation of SSB signals. SSB signal multiplies a carrier, and then the original baseband modulating signal is restored. A brief illumination is given as follows.

We have proven in Sec. 2. 10. 2 that if two time functions are convoluted:

$$z(t) = x(t) * y(t)$$

then their Fourier transforms have multiplication relationship:

$$Z(\omega) = X(\omega) Y(\omega)$$

Similarly, we can prove (left as exercise) that if two time functions are multiplied:

$$z(t) = x(t) y(t)$$

then their Fourier transforms have convolution relationship:

$$Z(\omega) = X(\omega) * Y(\omega)$$

During the demodulation of SSB signal, carrier $\cos\omega_0 t$ multiplying the received signal is equivalent to the convolution of the carrier frequency spectrum with the signal frequency spectrum in the frequency domain. Now we use demodulation of the upper-sideband signal as an example to draw the convolution of the frequency spectrum in Fig. 3. 2. 8. Among these figures, Fig. 3. 2. 8(a) is the frequency spectrum of the carrier; Fig. 3. 2. 8(b) is the frequency spectrum of the upper-sideband signal; Fig. 3. 2. 8(c) is the convolution result which is then filtered by a low-pass filter $H_L(f)$ to obtain the required frequency spectrum $M(f)$ of the demodulated baseband signal.

SSB modulation can further save transmitting power and the occupied frequency band; hence it is an extensively-used transmission system in analog communication.

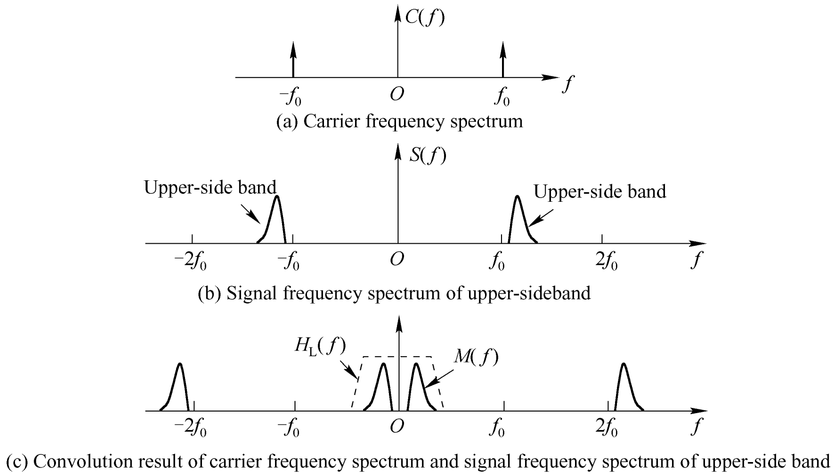


Figure 3.2.8 Demodulation of SSB signal

3.2.4 Vestigial Sideband Modulation (VSB)

The SSB signal introduced above has the advantages in the utilization of power and frequency band, but during demodulation in the receiver, it is necessary for the local carrier to possess the same frequency and to be in-phase with that in the transmitter; and herewith the frequency spectrum of SSB signal can be moved to the correct baseband location. In addition, in order to filter the SSB signal in the transmitter, sharp cut-off of the filter is required. However, sometimes it is difficult to attain the purpose. Frequency spectrum width of VSB signal is between that of DSB and SSB signals; and its frequency spectrum contains carrier component. Hence, it can counterbalance the disadvantage of the SSB modulation mentioned above. It is particularly suitable for the baseband signal containing D. C. and very low frequency components. VSB modulation is currently used in analog TV broadcasting system extensively. It will be briefly introduced as follows.

Vestigial sideband modulation belongs to linear modulation. The block diagram of a linear modulator in Fig. 3.2.1 is still applicable, only the characteristic of the filter should be modified correspondingly. Frequency spectrum expression of the output signal of the multiplier in Fig. 3.2.1 is

$$S'(f)=\frac{A}{2}[M(f-f_0)+M(f+f_0)]$$

Assume the transfer characteristic of the filter producing VSB signal is $H(f)$. Frequency spectrum of VSB signal $s(t)$ after being filtered should be

$$S(f)=\frac{A}{2}[M(f-f_0)+M(f+f_0)]H(f) \tag{3.2-14}$$

Now let's find the condition which should be satisfied by transfer function $H(f)$ of the filter in VSB signal modulator. If the demodulation method in Fig. 3.2.6 is still used, then after multiplication of signal $s(t)$ and local carrier $\cos\omega_0 t$, the frequency spectrum of product $r'(t)$ will be the result of the displacement f_0 of $S(f)$, i. e., the frequency spectrum of $r'(t)$ is

$$\frac{1}{2}[S(f+f_0)+S(f-f_0)] \tag{3.2-15}$$

Substituting eq.(3.2-14) into eq.(3.2-15), obtains the frequency spectrum of:

$$\frac{A}{4}\{[M(f+2f_0)+M(f)]H(f+f_0)+[M(f-2f_0)+M(f)]H(f-f_0)\} \tag{3.2-16}$$

The two terms $M(f+2f_0)$ and $M(f-2f_0)$ in the above equation can be filtered out by low-pass filter. Hence, the output demodulated signal after being filtered is

$$\frac{A}{4}M(f)[H(f+f_0)+H(f-f_0)] \tag{3.2-17}$$

For distortionless transmission, require

$$H(f+f_0)+H(f-f_0)=C \tag{3.2-18}$$

where $C=$ constant.

Since $M(f)$ is the frequency spectral density function of the baseband modulating signal, as shown in Fig. 3.2.3, its highest frequency component is f_m, i.e.,

$$M(f)=0 \qquad |f|>f_m \tag{3.2-19}$$

So, eq.(3.2-18) can be written as

$$H(f+f_0)+H(f-f_0)=C \quad |f|>f_m \tag{3.2-20}$$

The above equation is a prerequisite for the filter characteristic to produce VSB signal. The required condition is drawn in Fig. 3.2.9, i.e., the cut-off characteristic of the filter is complementary symmetry with respect to the carrier frequency f_0.

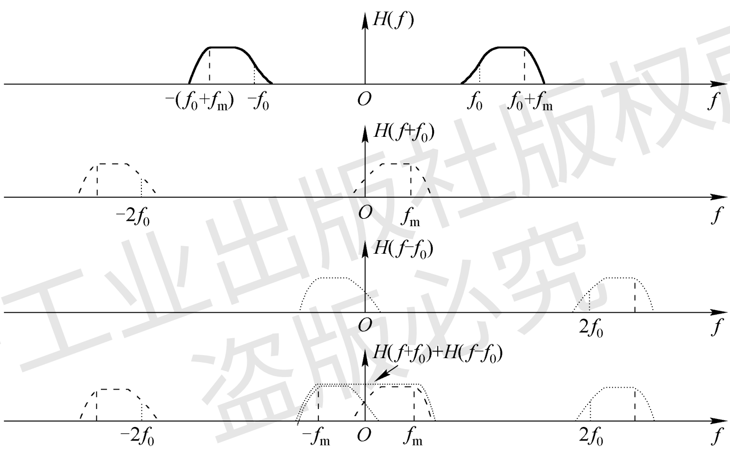


Figure 3.2.9 Filter characteristic for producing VSB signal

As can be seen from Fig. 3.2.9, in addition to the whole frequency spectrum of SSB signal, a part of the carrier frequency component and a small part of another sideband frequency spectrum are reserved in the frequency spectrum of VSB modulated signal. In this way, frequency error produced in the receiver during extraction of the transmitted carrier frequency is avoided; and the filter in the modulator of the transmitter can easily be manufactured. The main disadvantage is that the occupied frequency band is a little wider than that of the SSB signal.

The above description of VSB modulation is for the reservation of the whole lower-side band and a part of the upper-side band. Similarly, VSB modulation may reserve the whole upper-side band and a part of the lower-side band.

3.3 Nonlinear Modulation

3.3.1 Basic Principles

In the discussion of linear modulation, we were already familiar with the concept of the carrier.

Modulating signal is carried on the amplitude of the carrier in linear modulation. Nonlinear modulation is also called angle modulation, where modulating signal is carried on the phase of the carrier. In mathematical definition, carrier is a sinusoidal (or cosinusoidal) wave with constant amplitude, constant frequency, and constant phase, as well as infinitely extended in time from negative infinity to positive infinity. Therefore, it has single frequency component in frequency domain. The frequency spectrum of the carrier no longer has single frequency component after being modulated or truncated; and has many discrete or continuous frequency components as well as occupies certain frequency bandwidth. That is to say, angle modulation enables frequency and phase of the carrier to vary with the modulating signal. Here the concept of "instantaneous frequency" is already introduced because in its strict mathematical meaning, frequency of a carrier is a constant. Now let us define instantaneous frequency.

Assume a carrier can be expressed as

$$c(t) = A\cos\varphi(t) = A\cos(\omega_0 t + \varphi_0) \tag{3.3-1}$$

where φ_0 is the initial phase of the carrier; $\varphi(t) = \omega_0 t + \varphi_0$ is the instantaneous phase of the carrier; $\omega_0 = \mathrm{d}\varphi(t)/\mathrm{d}t$ is the angular frequency of the carrier.

Angular frequency ω_0 of the carrier is originally a constant. Here $\mathrm{d}\varphi(t)/\mathrm{d}t$ after being angular modulated is defined as instantaneous frequency $\omega_i(t)$, i. e.,

$$\omega_i(t) = \mathrm{d}\varphi(t)/\mathrm{d}t \tag{3.3-2}$$

It is a function of time.

From the above equation we can write

$$\varphi(t) = \int \omega_i(t)\,\mathrm{d}t + \varphi_0 \tag{3.3-3}$$

As can be seen from eq.(3.3-1), $\varphi(t)$ is the phase of the carrier. If it varies with modulating signal $m(t)$ according to certain mode, then it is called angle modulation. If phase $\varphi(t)$ linearly varies with $m(t)$, i. e., let

$$\varphi(t) = \omega_0 t + \varphi_0 + k_{\mathrm{p}} m(t) \tag{3.3-4}$$

where k_{p} is a constant, then it is called phase modulation. Thus, the expression of the modulated signal is

$$s_{\mathrm{p}}(t) = A\cos[\omega_0 t + \varphi_0 + k_{\mathrm{p}} m(t)] \tag{3.3-5}$$

Substituting eq.(3.3-4) into eq.(3.3-2), instantaneous frequency of the modulated carrier can be obtained as

$$\omega_i(t) = \omega_0 + k_{\mathrm{p}} \frac{\mathrm{d}}{\mathrm{d}t} m(t) \tag{3.3-6}$$

The above equation shows instantaneous frequency in phase modulation varies with the derivative of the modulating signal.

If instantaneous frequency linearly varies with the modulating signal, then frequency modulation is obtained. Now instantaneous frequency is

$$\omega_i(t) = \omega_0 + k_f m(t) \tag{3.3-7}$$

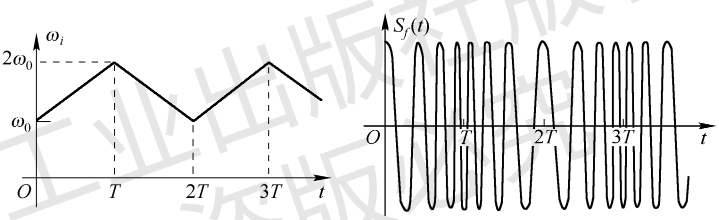
And from eq.(3.3-3) we have

$$\varphi(t) = \int \omega_i(t)\,\mathrm{d}t + \varphi_0 = \omega_0 t + k_f \int m(t)\,\mathrm{d}t + \varphi_0 \tag{3.3-8}$$

In this way, the expression of the modulated signal obtained is

$$s_f(t) = A\cos[\omega_0 t + \varphi_0 + k_f \int m(t)\,\mathrm{d}t] \tag{3.3-9}$$

As can be seen from the above discussion, the phase of the carrier in phase modulation varies linearly with modulating signal $m(t)$, and the phase of the carrier in frequency modulation varies linearly with the integration of modulating signal. There is no difference between them substantially. If modulating signal $m(t)$ is integrated first, and the phase of the carrier is modulated, then frequency modulated signal is obtained. Similarly, if modulating signal $m(t)$ is differentiated first, and frequency of the carrier is modulated, then phase modulated signal is obtained. Whether it is frequency modulation or phase modulation, the amplitude of a modulated signal is constant. And it is impossible to distinguish them by modulated signal waveforms. The difference between them is only the relationship between modulated signal and modulating signal. Fig. 3.3.1 is an example of waveform of angle modulation. Fig. 3.3.1(a) shows the relationship between instantaneous frequency ω_i of the modulated signal and time; it linearly varies between ω_0 and $2\omega_0$. Fig. 3.3.1(b) shows the waveform of a modulated signal. If the waveform of modulating signal $m(t)$ varies linearly as shown in Fig. 3.3.1(a), i.e., $m(t)$ varies along a straight line, then the modulated signal is frequency modulated signal. If $m(t)$ varies with t^2, then modulated signal is phase modulated signal, because from eq.(3.3-6) the instantaneous frequency of the modulated signal also varies linearly. Hence, we won't distinguish frequency modulation signal from phase modulation signal in the following discussion on the performance of angle modulation signal.



(a) Relation between instantaneous frequency and time

(b) Waveform of modulated signal

Figure 3.3.1 Waveform of angle modulation

3.3.2 Frequency Spectrum and Bandwidth of Modulated Signal

The bandwidths of various linear modulation signals discussed in Sec. 3.2 equal one to two times of the bandwidth of baseband signal. However, the bandwidth of angle modulated signal may be much larger than that of baseband modulated signal. As a simple example for analysis, a sinusoidal modulating signal for frequency modulation is used.

Assume the modulating signal $m(t)$ is a cosinusoidal wave:

$$m(t) = \cos\omega_m t \tag{3.3-10}$$

As can be seen from eq.(3.3-7), instantaneous angular frequency obtained from frequency modulation by this modulating signal is

$$\omega_i(t) = \omega_0 + k_f m(t) = \omega_0 + k_f \cos\omega_m t \tag{3.3-11}$$

The above equation shows that the largest frequency deviation $\Delta\omega$ of the carrier angular frequency equals

$$\Delta\omega = k_f \quad (\text{rad/s}) \tag{3.3-12}$$

Without losing generality, we let $\varphi_0 = 0$ in eq.(3.3-9), then the expression of modulated signal is

$$s_f(t) = A\cos\left[\omega_0 t + k_f\int\cos\omega_m t\mathrm{d}t\right] = A\cos[\omega_0 t + (\Delta\omega/\omega_m)\sin\omega_m t] \qquad (3.3\text{-}13)$$

where $\Delta\omega/\omega_m = \Delta f/f_m$ is the ratio of the largest frequency deviation to the frequency of the baseband signal; and is called frequency modulation index m_f:

$$m_f = \Delta f/f_m = \Delta\omega/\omega_m = k_f/\omega_m \qquad (3.3\text{-}14)$$

Eq.(3.3-13) is the expression of a frequency modulated signal, and there is a cosine function containing a sine function in the equation. It can be proven that the equation may be expanded as the following infinite series:

$$\begin{aligned} s_f(t) = A\{&J_0(m_f)\cos\omega_0 t + J_1(m_f)[\cos(\omega_0+\omega_m)t - \cos(\omega_0-\omega_m)t] + \\ &J_2(m_f)[\cos(\omega_0+2\omega_m)t + \cos(\omega_0-2\omega_m)t] + \\ &J_3(m_f)[\cos(\omega_0+3\omega_m)t - \cos(\omega_0-3\omega_m)t] + \cdots\} \end{aligned} \qquad (3.3\text{-}15)$$

where $J_n(m_f)$ is called the Bessel function of the first kind of order n. Its value can be obtained by table-lookup method (see Appendix F), and also can be obtained from the curves in Fig. 3. 3. 2.

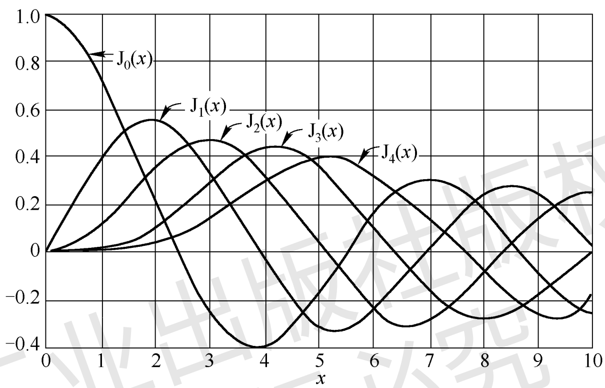


Figure 3. 3. 2 Bessel function

Since the Bessel function has the following properties:

$$\left.\begin{aligned} J_n(m_f) &= J_{-n}(m_f) \quad &\text{when } n \text{ is even} \\ J_n(m_f) &= -J_{-n}(m_f) \quad &\text{when } n \text{ is odd} \end{aligned}\right\} \qquad (3.3\text{-}16)$$

eq.(3.3-15) can be rewritten as

$$s_f(t) = A\sum_{n=-\infty}^{\infty} J_n(m_f)\cos(\omega_0 + n\omega_m)t \qquad (3.3\text{-}17)$$

It is known from eq.(3.3-15) and eq.(3.3-17), there are side frequencies in pair with angular frequencies $(\omega_0 \pm \omega_m)$, $(\omega_0 \pm 2\omega_m)$, $(\omega_0 \pm 3\omega_m)$, $\cdots$ appearing on two sides of the carrier frequency in the frequency spectrum, as shown in Fig. 3. 3. 3. It seems apparent from the figure that the bandwidth of modulated signal is infinity. However, actually most of the power concentrates within a limited bandwidth centered at the carrier frequency. For example, when frequency modulation index $m_f \ll 1$, as can be seen from Fig. 3. 3. 2 that all components may be neglected except $J_0(m_f)$ and $J_1(m_f)$. Now the bandwidth of the modulated signal is basically equal to that of amplitude modulation, i. e., $2\omega_m$. Frequency modulation with small modulation index is called narrowband frequency modulation. When the modulation index increases, bandwidth of the modulated signal also increases. Here modulation is called broadband frequency modulation. During broadband frequency modulation, if the side frequencies with amplitudes smaller than 1% of the unmodulated carrier amplitude are neg-

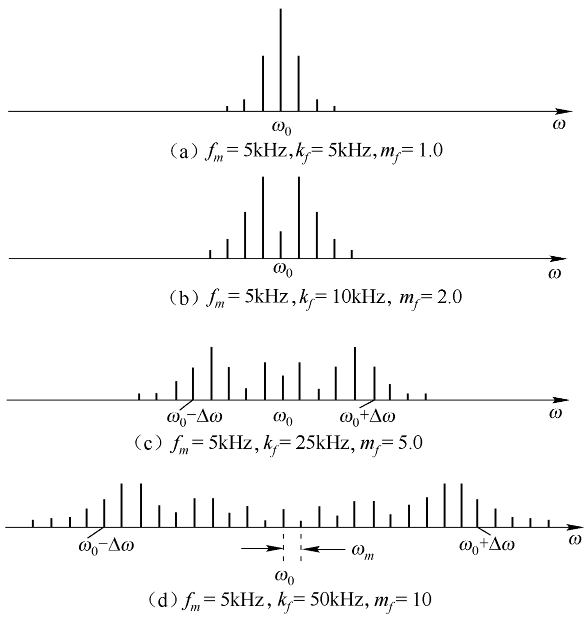


Figure 3. 3. 3 Example of frequency spectrum for frequency modulated signal

lected, then as can be seen from the curves of the Bessel function that the $J_n(m_f)$ with $n > m_f$ (n equals integer) may be neglected. Thus, the bandwidth B of the modulated signal for frequency modulation can approximately be

$$B \approx 2(\Delta\omega + \omega_m) \quad (\text{rad/s}) \tag{3. 3-18}$$

$$B \approx 2(\Delta f + f_m) \quad (\text{Hz}) \tag{3. 3-19}$$

where $\Delta f = \Delta\omega / 2\pi$ is frequency deviation; f_m is the frequency of the modulating signal.

The instance discussed above is the situation of single sinusoidal wave modulation. When a modulating signal has many frequency components, f_m in the above equation should be the frequency of the highest frequency component of the modulating signal.

3. 3. 3 Reception of Angular Modulated Signal

It is mentioned in Sec. 3. 3. 1 that the amplitude of an angular modulated signal is constant. Although the amplitude of the angular modulated signal after transmission over a random parameter channel fluctuates due to the rapid fading and the superimposed noise, there will be no loss of information by the change of the signal amplitude. The reason is that the amplitude of the angular modulated signal doesn't contain information of the modulating signal. The influence of the fading and the noise in the channel on the angle (frequency and phase) of the signal is less than that on the amplitude. Therefore anti-jamming ability of angular modulated signal is strong. For eliminating the influence of the fading and the noise on angular modulated signal, generally an amplitude limiter is used before demodulation in the receiver to remove amplitude variation. The received signal after the amplitude limiter becomes a signal with constant amplitude, and then demodulated by a frequency discriminator or a phase discriminator.

3.4 Brief Summary

Analog modulation system is briefly discussed in this chapter. Analog modulation is classified into two categories: linear modulation and nonlinear modulation. The structure of frequency spectrum of a linear modulated signal is the displacement of the frequency spectrum of the modulating signal, or the filtering-out of unnecessary frequency components after displacement. Linear modulation includes AM, SSB, DSB, and VSB. Nonlinear modulation is also called angle modulation. Frequency spectrum structure of a nonlinear modulated signal is much different from that of the modulating signal. Many new frequency components appear in the frequency spectrum of a modulated signal, hence frequency bandwidth of the signal also increases considerably. Nonlinear modulation includes frequency modulation and phase modulation.

In amplitude modulation system, ratio of the peak voltage of baseband modulating signal to that of the modulated carrier is called modulation index. The maximum value of modulation index is 100%. The envelope of modulated signal is the same as the waveform of baseband modulating signal, so envelope detection method can be used for the demodulation of amplitude modulated signal. The carrier component in the amplitude modulated signal doesn't carry information in baseband signal, but possesses a greater part of power in the signal, so transmission efficiency is low. If the carrier in amplitude modulated signal is eliminated, then we obtain DSB signal. In comparison with amplitude modulated signal, DSB signal can save a large part of transmitting power, but the carrier must be restored in the receiver. In this way, the complexity of the receiver increases. In DSB signal, upper-side band and lower-side band carry the same baseband information, and hence repeated transmission is formed. Therefore, we may transmit only upper-side band or lower-side band. Thus SSB signal is obtained. SSB signal has the superiority in the utilization of power and frequency band, but carrier frequency should be restored in the receiver during demodulation. In addition, for filtering SSB signal in the transmitter, sharp cut-off of the filter is required, which is sometimes difficult. Frequency spectrum of VSB signal is between that of DSB signal and that of SSB signal, and contains carrier component. Therefore the disadvantage of SSB signal mentioned above can be avoided. VSB signal is especially suitable for baseband signal containing D. C. component and very low frequency components.

Substantially, there is no difference between frequency modulation and phase modulation in angle modulation. They can't be distinguished from modulated signal waveforms; and difference exists only in the relationship between modulating signal and modulated signal. Generally speaking, angle modulation occupies broader frequency band. The amplitude of these signals doesn't contain information of modulating signal; hence in despite of the amplitude of the received signal random fluctuating due to transmission, information in the signal is unable to get lost. Angular modulated signal has strong anti-jamming ability, and is especially suitable for transmission in the fading channel.

Questions

3.1 What is the purpose of modulation?

3.2 What categories can be classified for analog modulation?

3.3 How is linear modulation categorized?

3.4 How is nonlinear modulation categorized?

3.5 What is the difference between amplitude modulation and DSB modulation? Are the bandwidths of these modulated signals equal?

3.6 Does the bandwidth of DSB speech signal equal twice that of SSB speech signal?

3.7 What is the requirement on the characteristic of the filter for producing the VSB signal?

3.8 What kind of baseband signal is suitable for VSB modulation?

3.9 Write the expressions of frequency modulated signal and phase modulated signal.

3.10 What is the frequency modulation index?

3.11 Write the approximate expression of the bandwidth for the frequency modulated signal.

3.12 Describe the main advantages of angle modulation.

Exercises

3.1 Assume the expression of a carrier is $c(t) = 5\cos(1000\pi t)$, and the expression of baseband modulating signal is $m(t) = 1 + \cos(200\pi t)$. Find the frequency spectrum of amplitude modulated signal, and plot frequency spectrum diagram.

3.2 In the above exercises, what are the amplitudes of carrier component and each sideband component respectively?

3.3 Assume the carrier frequency of a frequency modulated signal is 10 kHz, baseband modulating signal is a single sinusoidal wave with frequency 2 kHz, and modulation frequency deviation is 5 kHz. Find modulation index and the bandwidth of the modulated signal.

3.4 Prove that if a baseband cosinusoidal wave is used for amplitude modulation, then the maximum power of the sum of two sidebands of the amplitude modulated signal equals one half of the carrier power.

3.5 Prove that if two time functions are multiplied: $z(t) = x(t)y(t)$, then their Fourier transforms have convolution relationship: $Z(\omega) = X(\omega) * Y(\omega)$.

3.6 Let a baseband modulating signal be a sinusoidal wave with the frequency 10 kHz, and the amplitude 1 V. It modulates the phase of a carrier with frequency 10 MHz, and the maximum phase deviation of modulation is 10 rad. Calculate the approximate bandwidth of the phase modulated signal. If the frequency of the modulating signal is changed to 5 kHz, find its bandwidth.

3.7 If the modulating signal modulates the frequency of the carrier in the above exercise, and the maximum frequency deviation of modulation is 1 MHz, find the approximate bandwidth of the frequency modulated signal.

3.8 Let the expression of an angular modulated signal be

$$s(t) = 10\cos(2 \times 10^6 \pi t + 10\cos 2000\pi t)$$

Find: (1) the maximum frequency deviation of the modulated signal;

(2) the maximum phase deviation of the modulated signal;

(3) the bandwidth of the modulated signal.