2.11 Phase Shifter Networks and the Hilbert Transform

Suppose the input to an LTI system is $x(t)=A\cos(2\pi f_o t)$ where f_o can take any value between 0 and ∞. If, for any f_o, the output is $y(t) = A\cos(2\pi f_o t+\phi)$, where ϕ is independent of the value of the frequency, then the system is a ϕ phase shift network.

The frequency response of a phase shift network is

$$H(f) = \begin{cases} e^{j\phi} & f > 0 \\ e^{-j\phi} & f < 0 \end{cases} \tag{2.120}$$

The Hilbert transform is a phase shift network that shifts the phase of the input by $-\pi/2$. Therefore the frequency response of the Hilbert transform is:

$$H_{HT}(f) = \begin{cases} e^{-j\frac{\pi}{2}} & f > 0 \\ e^{j\frac{\pi}{2}} & f < 0 \end{cases} \tag{2.121}$$

We can rewrite (2.121) as:

$$H_{HT}(f) = \begin{cases} -j & f > 0 \\ j & f < 0 \end{cases} = -j\text{sgn}(f) \tag{2.122}$$

We can also find the impulse response of a Hilbert transformer by taking the inverse Fourier transform of (2.122). It is given by:

$$h_{HT}(t) = \frac{1}{\pi t} \tag{2.123}$$

Let $x_h(t)$ be the Hilbert transform of $x(t)$. Then

$$x_h(t) = x(t) * \frac{1}{\pi t} \tag{2.124}$$

and

$$X_h(f) = -j\text{sgn}(f)X(f) \tag{2.125}$$

Often we can find $x_h(t)$ from the inverse Fourier transform of $X_h(f)$.

Exercise 2.31

Show that $h_{HT}(t) = 1/\pi t$.

Example 2.18

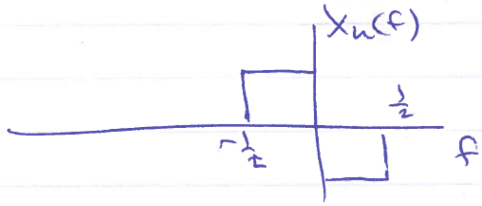
Find the Hilbert transforms of $x(t) = A\text{sinc}(t)$ and $y(t) = Az(t)\cos(2\pi f_o t)$ where f_o is larger than the bandwidth of $z(t)$.

Solution

$$X(f) = \Pi(f)$$

$$X_h(f) = -j\,\text{sgn}(f)\,\Pi(f)$$

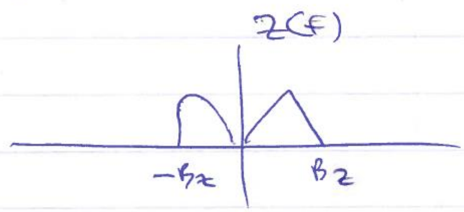
$$= \begin{cases} j & -\frac{1}{2} \le f \le 0 \\ -j & 0 \le f \le \frac{1}{2} \\ 0 & \text{otherwise/autrement} \end{cases}$$

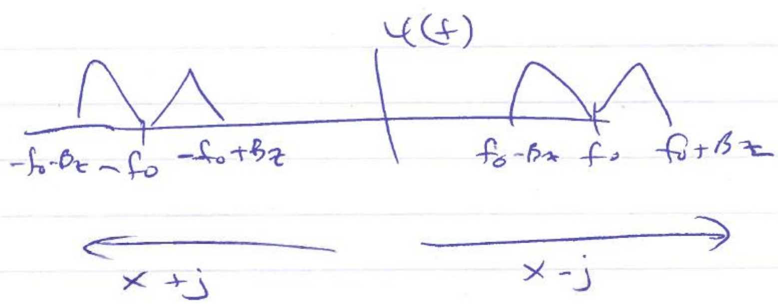


$$X_h(f) = -j\,\Pi\left(2\left(f-\tfrac{1}{4}\right)\right) + j\,\Pi\left(2\left(f+\tfrac{1}{4}\right)\right)$$

$$x_h(t) = \mathcal{F}^{-1}\{X_h(f)\}$$
$$= \text{sinc}\left(\frac{t}{2}\right)\sin\left(\frac{\pi}{2}f\right)$$

$$y(t) = A\,z(t)\cos 2\pi f_0 t$$
$$Y(f) = \frac{A}{2}Z(f-f_0) + \frac{A}{2}Z(f+f_0)$$





$$Y_h(f) = -j\frac{A}{2}Z(f-f_0) + j\frac{A}{2}Z(f+f_0)$$

$$= \frac{A}{2j}Z(f-f_0) - \frac{A}{2j}Z(f+f_0)$$

$$y_h(t) = \mathcal{F}^{-1}\{Y_h(f)\}$$

$$= A\,z(t)\sin 2\pi f_0 t$$

2.12 The Preenvelope of a Signal

A signal, $x(t)$, can be described by its positive and negative preenvelopes, $x_+(t)$ and $x_-(t)$. The spectrum of the positive preenvelope, $X_+(f)$, is given by:

$$X_+(f) = \begin{cases} 2X(f), & f > 0 \\ X(f), & f = 0 \\ 0, & f < 0 \end{cases} \tag{2.126}$$

and the spectrum of the negative preenvelope is given by:

$$X_-(f) = \begin{cases} 0, & f > 0 \\ X(f), & f = 0 \\ 2X(f), & f < 0 \end{cases} \tag{2.127}$$

It is clear from (2.126) and (2.127) that $X_+(f) + X_-(f) = 2X(f)$. Therefore

$$x(t) = \tfrac{1}{2}x_+(t) + \tfrac{1}{2}x_-(t). \tag{2.128}$$

In order to define $x_+(t)$ and $x_-(t)$ in the time domain, we need to take the inverse Fourier transform of (2.126) and (2.127). However, we can simplify them by noting that (2.126) can be written as:

$$X_+(f) = X(f) + \mathrm{sgn}(f)X(f) \tag{2.129}$$

Also, since $(-j)(j) = 1$, we can rewrite (2.129) as:

$$X_+(f) = X(f) + j(-j\mathrm{sgn}(f)X(f)) \tag{2.130}$$

We can recognize that the term in parentheses in (2.130) is the Fourier transform of the Hilbert transform of $x(t)$. Therefore:

$$X_+(f) = X(f) + jX_h(f) \tag{2.131}$$

and

$$x_+(t) = x(t) + jx_h(t) \tag{2.132}$$

If $x(t)$ is a real-valued signal, its Hilbert transform is also real-valued. Therefore when $x(t)$ is a real-valued signal:

$$x(t) = \mathrm{Re}\{x_+(t)\} \tag{2.133}$$

From (2.128) we can deduce that $x_-(t)$ is given by:

$$x_-(t) = x(t) - jx_h(t) \tag{2.134}$$

Example 2.19

Find the positive preenvelope of $x(t) = \mathrm{sinc}(t)$.

Solution

The Fourier transforms of $x(t)$ and $x_+(t)$ are shown in Figure 2.17. We can see from Fig. 2.17(b) that $X_+(f) = 2\Pi(2(f-0.25))$.

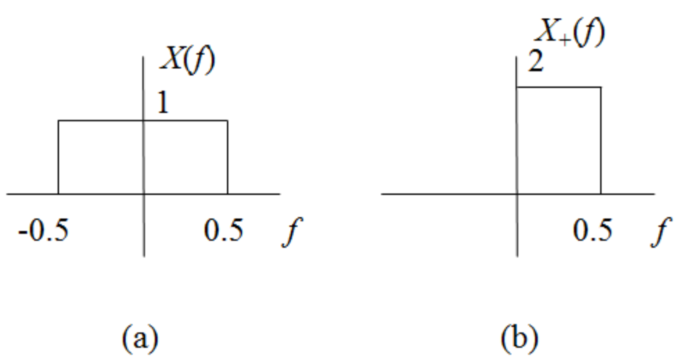


Figure 2.17: Fourier transform of (a) $x(t)$=sinc(t) and (b) $x_+(t)$.

The preenvelope of $x(t)$=sinc(t) is therefore:

$$x_+(t) = \mathcal{F}^{-1}\left\{2\Pi\left[2\left(f-\tfrac{1}{4}\right)\right]\right\} = \text{sinc}\left(\tfrac{t}{2}\right)e^{j\frac{\pi}{2}t}$$

2.12 Bandpass Signals and Their Complex Envelopes (Lowpass Equivalent)

A bandpass signal is one that occupies a range of frequencies not adjacent to 0 and has a Fourier transform that resembles one that would be at the output of a bandpass filter. Figure 2.18 demonstrates the spectrum of both a baseband and bandpass signal. The bandwidth of the baseband signal is W while the bandwidth of the bandpass signal is $2B$.

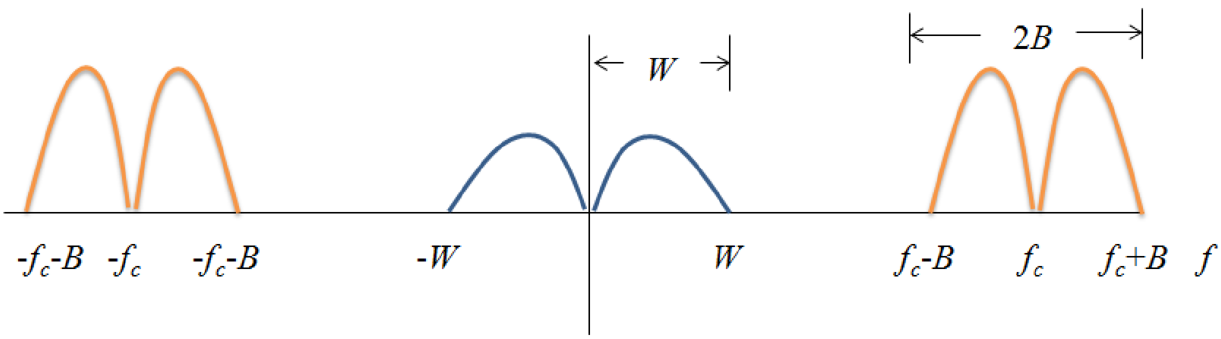


Figure 2.18: Spectrums of a baseband (blue) and bandpass (red) signal.

Note that the positive and negative portions of a signal`s spectrum and very close to one another in the case of a baseband signal, while for bandpass signals, they are separated by a very wide range of frequencies. A baseband signal`s spectrum does not need to contain frequency components very close to 0 but generally the highest non-zero frequency component (W) is very large compared to the lowest non-zero frequency component. For bandpass signals, the lowest non-zero frequency component (f_c-B) is comparable in magnitude to the highest non-zero frequency component (f_c+B) in its spectrum. This is because f_c (referred to as the center or carrier frequency) is large compared to B. Bandpass signals in communication systems are usually modulated so as to multiplex different signals on a common channel. Other signals will be moved in the spectrum to occupy frequencies that are unoccupied by the bandpass signal of Figure 2.18. For that reason, we do not include frequencies below f_c-B in its bandwidth, because other signals occupy those frequencies. Therefore its bandwidth is $f_c + B - (f_c - B) = 2B$. In the case of baseband signals, we specify their bandwidth simply as its highest non-zero frequency component since these signals are not meant to be transmitted over a shared channel, therefore even if there are some unused frequencies at the low end of its spectrum, we do not remove them from the signal`s bandwidth calculation. Other reasons for this will become evident in the following chapters on analog modulation techniques.

Let us consider a bandpass signal, $x(t)$, whose Fourier transform, $X(f)$, is shown in Figure 2.19(a). The Fourier transform of its pre-envelope, $X_+(f)$, is shown in Figure 2.19(b).

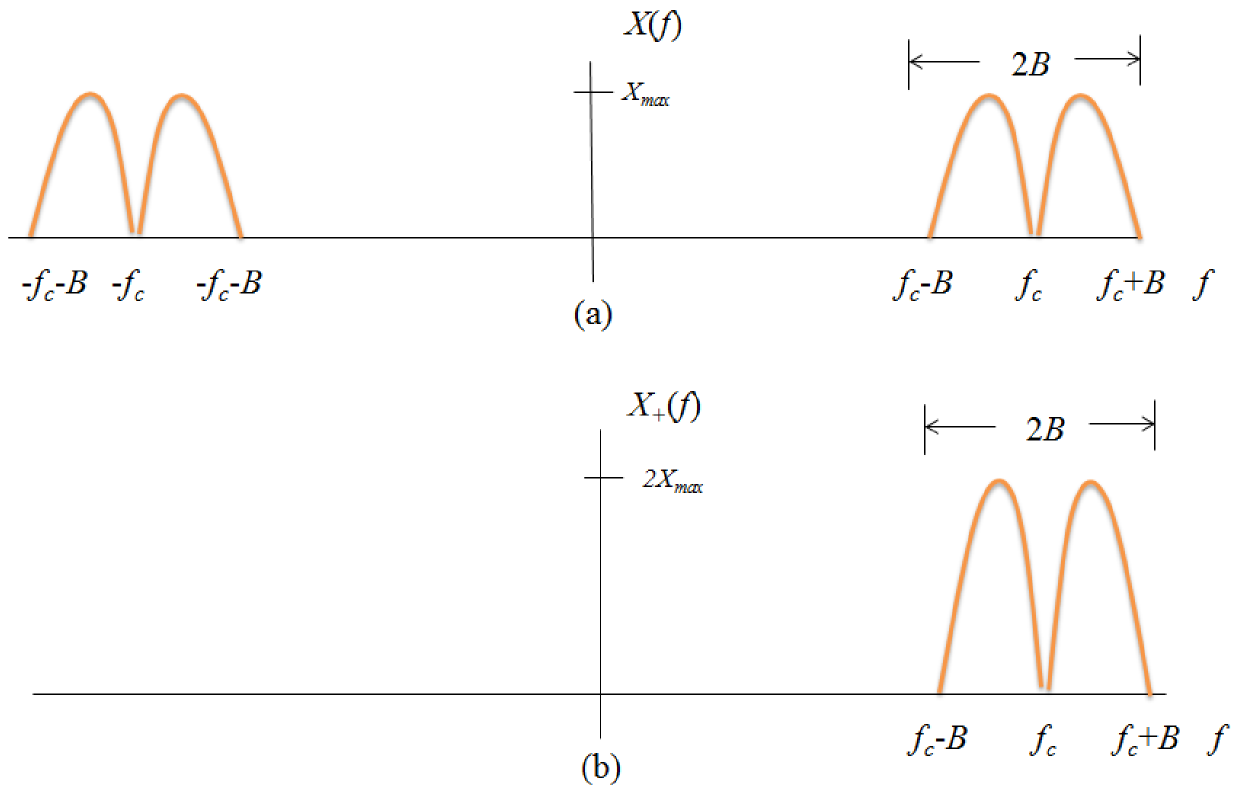


Figure 2.19 (a) Fourier Transform of a Bandpass Signal and (b) Fourier Transform of its Preenvelope.

We now define the complex envelope, $\tilde{x}(t)$, of $x(t)$ to be defined as having a Fourier transform that resembles that of a baseband signal. In other words, it has the Fourier transform of $X_+(f)$ centred at 0 rather than f_c. Therefore:

$$\tilde{X}(f) = X_+(f + f_c) \tag{2.135}$$

and therefore

$$\tilde{x}(t) = x_+(t)e^{-j2\pi f_c t} \tag{2.136}$$

From (2.136) we can express $x_+(t)$ in terms of $\tilde{x}(t)$ as:

$$x_+(t) = \tilde{x}(t)e^{j2\pi f_c t} \tag{2.137}$$

If $x(t)$ is a real-valued signal, then we know from (2.133) that $x(t)$ is the real part of $x_+(t)$. Therefore we can express $x(t)$ in terms of $\tilde{x}(t)$ by:

$$x(t) = \mathrm{Re}\left\{\tilde{x}(t)e^{j2\pi f_c t}\right\} \tag{2.138}$$

Expanding (2.138) we can express $x(t)$ as:

$$x(t) = \mathrm{Re}\{\tilde{x}(t)\}\cos(2\pi f_c t) - \mathrm{Im}\{\tilde{x}(t)\}\sin(2\pi f_c t) \tag{2.139}$$

Equation (2.139) is called the quadrature form of a bandpass signal, x(t) and it is written as:

$$x(t) = x_I(t)\cos(2\pi f_c t) - x_Q(t)\sin(2\pi f_c t) \tag{2.140}$$

where $x_I(t)$ = Re$\{\tilde{x}(t)\}$ is called the inphase component of $x(t)$ and $x_Q(t)$ = Im$\{\tilde{x}(t)\}$ is called the quadrature component of $x(t)$. We can express $\tilde{x}(t)$ as:

$$\tilde{x}(t) = x_I(t) + jx_Q(t) \tag{2.141}$$

Example 2.20

The signal $x(t)$ is a bandpass signal given by $x(t) = 3\cos(2\pi 500t) + 4\sin(2\pi 550t) + 4\sin(2\pi 600t)$.

(a) What is the bandwidth of $x(t)$?
(b) Express $x(t)$ as $x_I(t)\cos(2\pi 550t)$-$x_Q(t)\sin(2\pi 550t)$.
(c) Express $x(t)$ as $x_I(t)\cos(2\pi 500t)$-$x_Q(t)\sin(2\pi 500t)$.

Solution

Exercise 2.32

The signal $s(t) = 3\sin(2\pi 200t)+2\cos(2\pi 20t)\cos(2\pi 210t)$ is a bandpass signal.

(a) Find its Fourier transform.
(b) Express s(t) in its quadrature form for f_c = 210Hz.

Chapter 3 Amplitude Modulation Techniques

3.1 Introduction

Before an information-bearing signal can be transmitted over a channel, some type of modulation process is used to transform the information-bearing signal into a signal that can be easily accommodated by the channel but from with the information can be easily extracted. The process of converting the information-bearing signal into a signal for transmission over the channel is called modulation. The process of extracting the information from the modulated signal received from the channel is called demodulation.

There are two types of modulation schemes: Analog modulation and digital modulation. In analog modulation, the analog signal is used to modulate a carrier or pulse. In digital modulation, either the source produces digital symbols that are discrete in value and in time and these symbols are used to modulate a carrier or pulse or an analog signal is converted to a digital one before modulation. In digitally modulated signals, the demodulation process typically uses a correlator to determine the symbol value.

The two basic types of analog modulation are continuous-wave (CW) modulation and pulse modulation. CW modulation uses a high frequency carrier wave and is typically used in wireless communications. Pulse modulation uses a baseband pulse and is used to carry the values of the samples of a sampled analog signal. This technique is typically used in wireline communications.

The instantaneous value of the information-bearing signal can be carried by the amplitude of the carrier signal in CW modulation. In pulse modulation, the amplitude of the pulse can carry the value of the information-bearing signal's sample. If the amplitude of the carrier or pulse is linearly related to the information bearing signal, then the result is linear modulation.

If the information-bearing signal is carried in some other parameter, such as the frequency or phase of the carrier or in the width or position of the pulse, then the result is nonlinear modulation. In CW modulation, when the information-bearing signal is carried in the frequency or phase of the carrier then the modulation technique is called angle modulation. The two types of angle modulation are frequency modulation (FM) and phase modulation (PM).

The chapter is organized as follows: In section 3.2 we explain why modulation is used. In section 3.3 we introduce the concept of amplitude modulation (AM) by studying double sideband (DSB) modulation as well as double sideband suppressed carrier (DBS-SC) modulation. We also consider their demodulation. In section 3.4 we address the shortcomings of DSB and DSB-SC demodulation and introduce the conventional AM technique which is still used in AM radio today. Conventional AM was created to enable a simple demodulation but it comes at the cost of inefficient transmission power usage. In section 3.5 we discuss single sideband (SSB) modulation. Analog quadrature amplitude modulation (QAM) is discussed in 3.6.

3.2 Why do we use modulation?

Modulation is used to transmit a signal for the following reasons:

1) The spectrum of $m(t)$ is outside the range of frequencies of the channel's passband. By employing a carrier based modulation scheme, we can move the spectrum of the transmitted signal to the appropriate frequency range.
2) For ease of transmission, higher frequency signals must be used. For example, antenna length is proportional to the signal's wavelength, which is inversely proportional to frequency. By moving the transmitted signal to a higher frequency range by employing some modulation scheme, we can use smaller antennas.
3) The use of modulation allows for multiple users to access a shared channel, by providing some means of signal separation. For example, in amplitude modulation (AM), different signals are assigned different, sufficiently space, carrier frequencies. In doing so, the different signals occupy different frequency ranges in the spectrum, and each signal is uniquely recoverable. This is known as frequency division multiple access (FDMA), or frequency division multiplexing (FDM).

3.3 Double Sideband Modulation

In CW modulation, a parameter of the carrier wave is modulated so that its instantaneous value is a linear function of the information that is being transmitted. Let $m(t)$ be the information-bearing signal that we refer to as the **message signal**. Let $c(t)$ be the carrier wave that is modulated by $m(t)$. The carrier wave is given by:

$$c(t) = A_c \cos(2\pi f_c t + \phi_c) \tag{3.1}$$

where A_c is the carrier amplitude, f_c is the carrier frequency and ϕ_c is the carrier phase. In amplitude modulation, the carrier phase is not as important as the phase difference between the carrier used by the transmitter and the one generated at the receiver for demodulation purposes. Therefore, for simplification purposes, most textbooks in communications theory ignore ϕ_c by assuming it to be 0. Therefore we can assume that $c(t)$ is given by:

$$c(t) = A_c \cos(2\pi f_c t) \tag{3.2}$$

Double sideband (DSB) modulation is an amplitude modulation (AM) technique where the transmitted signal's amplitude is proportional to the message signal. Let $s_{DSB}(t)$ be the DSB modulated signal. It is given by:

$$s_{DSB}(t) = A_c m(t) \cos(2\pi f_c t) \tag{3.3}$$

Equation (3.3) shows that DSB modulation is simply the multiplication of the carrier of (3.2) by the message signal. It is crucial that the carrier frequency, fc, be much larger than the bandwidth of $m(t)$. Here we denote the bandwidth of $m(t)$ by B_m. The spectrum of $s_{DSB}(t)$ is $S_{DSB}(f)$ which is given by:

$$S_{DSB}(f) = \frac{A_c}{2} M(f - f_c) + \frac{A_c}{2} M(f + f_c) \tag{3.4}$$

Figure 3.1(a) shows a message signal $m(t)$, while Fig. 3.1(b) shows a carrier wave with amplitude of 3V. Figure 3.1(c) shows the DSB signal which is the result of the multiplication of the message and the carrier.

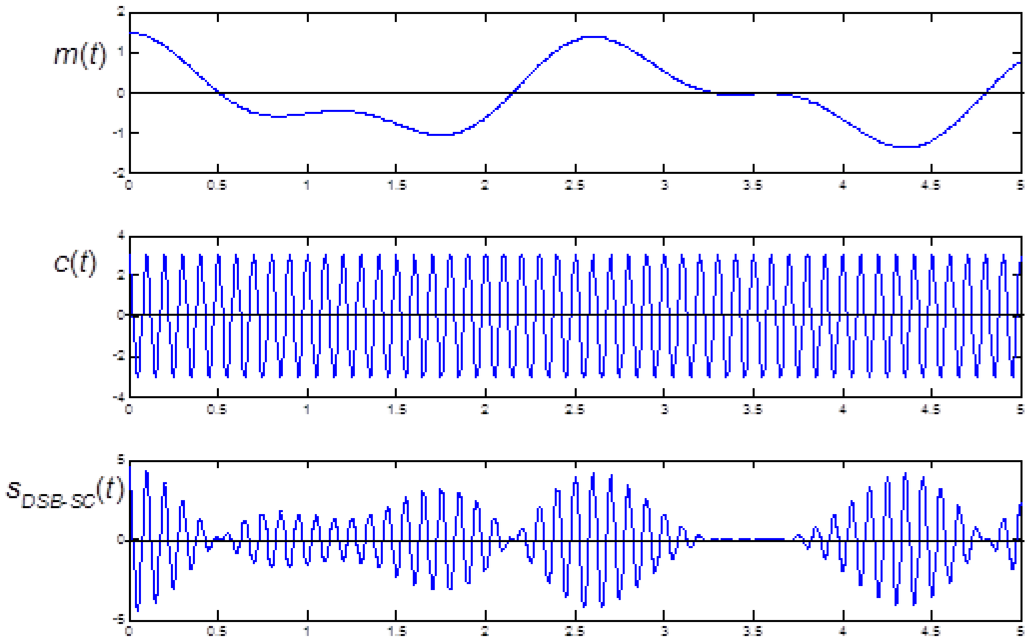


Figure 3.1: (a) Message signal, (b) carrier wave and (c) DSB signal.

Figure 3.2 shows the block diagram of a DSB modulator as well as the spectrum, $M(f)$, of a hypothetical message signal, $m(t)$, and the spectrum of the DSB signal. The bandwidth of the message signal is denoted by B_m. We can see from Fig. 3.2 that the bandwidth of the DSB signal is twice the bandwidth of the baseband signal m(t). If the bandwidth of the DSB signal is denoted by W, then

$$W = 2B_m \tag{3.5}$$

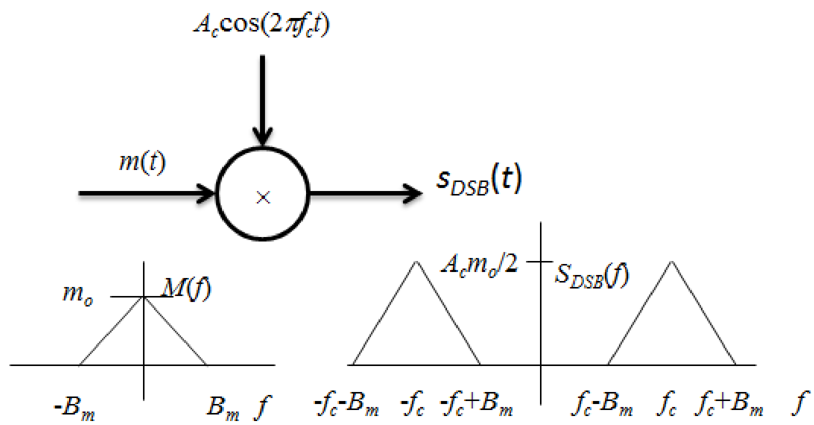


Figure 3.2: Block Diagram of a DSB Modulator and the Spectrum of the DSB Signal for a Given Message Signal Spectrum.

If the signal $m(t)$ does not have any spectral component at $f = 0$ (in other words, $M(0) = 0$), then the spectrum of $s_{DSB}(t)$ does not have a spectral component at the carrier frequency $\pm f_c$. Figure 3.3(a) shows the spectrum of a hypothetical signal $m(t)$ that does not have a spectral component at $f = 0$ and Figure 3.3(b) shows the spectrum of its DSB modulation signal. DSB signals that contain no spectral component at f = $\pm f_c$ are carrier double sideband suppressed carrier (DSB-SC) modulated signals.

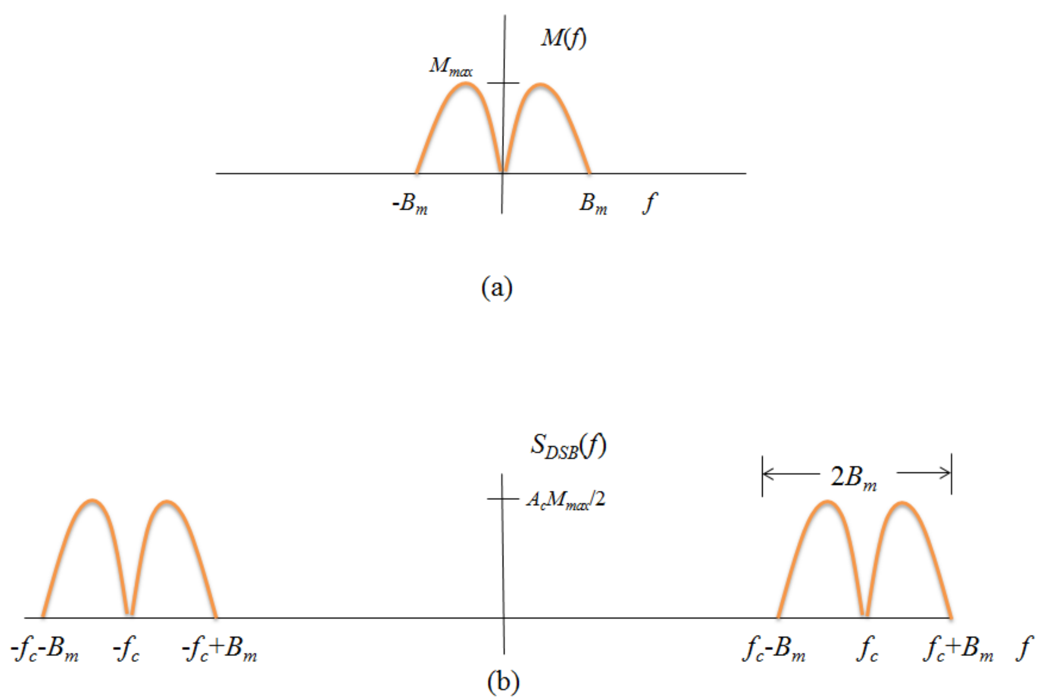


Figure 3.3: (a) Hypothetical spectrum of a message signal with no spectral component at $f = 0$ (b) Spectrum of its corresponding DSB-SC Signal.

The demodulation of a DSB signal is based on the trigonometric identity $2\cos^2(2\pi f_c t) = 1 + \cos(4\pi f_c t)$. By multiplying a DSB signal by $2\cos(2pf_c t)$ at the receiver, we get:

$$\begin{aligned} x(t) &= 2s_{DSB}(t)\cos(2\pi f_c t) \\ &= 2A_c m(t)\cos^2(2\pi f_c t) \\ &= A_c m(t) + A_c m(t)\cos(4\pi f_c t) \end{aligned} \tag{3.6}$$

Where the first term in the last line of (3.6) is a baseband signal and the second term is a new DSB signal (bandpass) centered at frequency $2f_c$. Assuming that $m(t)$ has the spectrum shown in Figure 3.3(a), then the spectrum of $x(t)$, shown in Figure 3.4, is given by:

$$X(f) = A_C M(f) + \frac{A_C}{2} M(f - 2f_c) + \frac{A_C}{2} M(f + 2f_c) \tag{3.7}$$

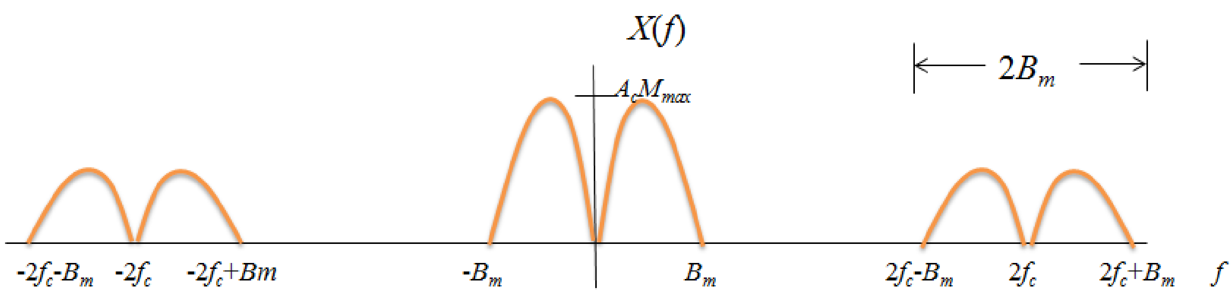


Figure 3.4: Spectrum of $x(t) = 2s_{DSB}(t)\cos(2\pi f_c t)$.

We can see from Figure 3.4 and (3.7) that by inputting $x(t)$ to a lowpass filter whose passband is slightly greater than B_m but completely rejects frequency components above $2f_c$-B_m, then the output of the lowpass filter will be $A_c m(t)$. Therefore the block diagram a demodulator for DSB modulation is given in Figure 3.5.

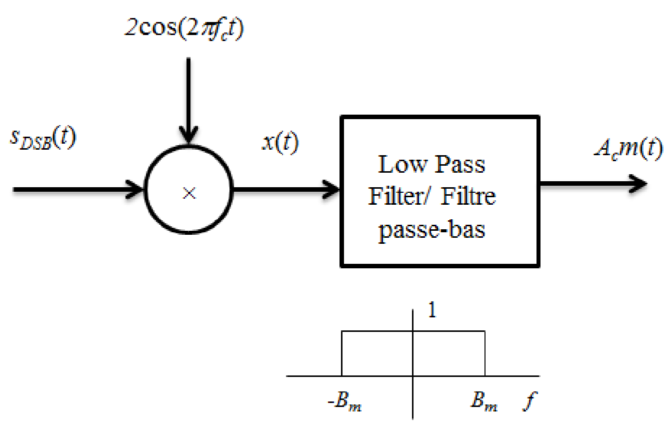


Figure 3.5: DSB Demodulator.

Example 3.1

The signal $m(t)$ is given by:

$$m(t) = 3\cos(2\pi 10t)+2\sin(2\pi 20t)$$

It is to be transmitted using DSB modulation using a carrier that has amplitude of 5V and a carrier frequency of 250Hz.

(a) Find the expression of $s_{DSB}(t)$ and sketch it.
(b) Find the spectrum of $s_{DSB}(t)$ ($S_{DSB}(f)$) and sketch it.
(c) What are the bandwidths of $m(t)$ and $s_{DSB}(t)$?
(d) Is $s_{DSB}(t)$ an energy or power signal? Find its corresponding energy or power.
(e) What is the autocorrelation function and corresponding spectral density of $s_{DBS}(t)$.
(f) Demonstrate the demodulation of the DSB signal by finding $x(t)$ and the corresponding output of the lowpass filter for the demodulator of figure 3.5. What is an appropriate cut-off frequency for the lowpass filter?

Solution

(a) $s_{DSB}(t) = 5(3\cos(2\pi 10t)+2\sin(2\pi 20t))\cos(2\pi 250t)$. The signal $m(t)$ is shown in Figure 3.6(a) while Figure 3.6(b) gives $s_{DSB}(t)$.

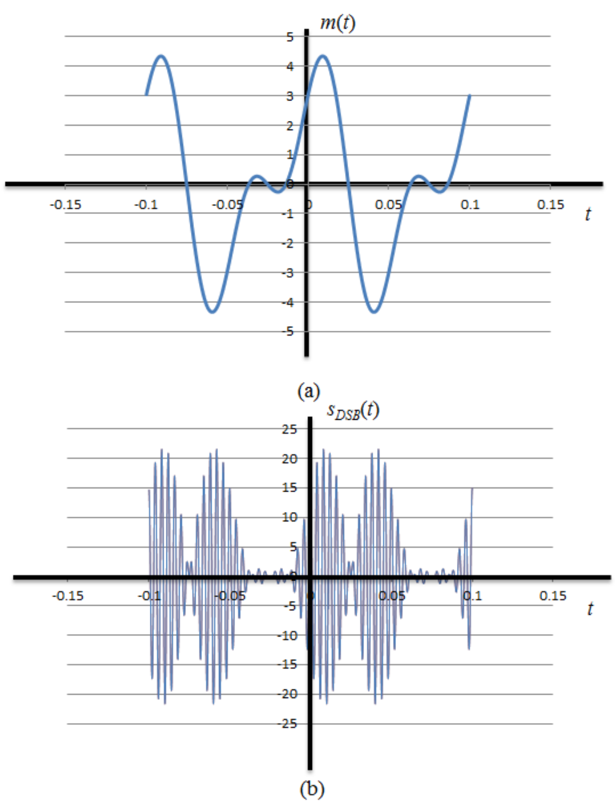


Figure 3.6(a) The signal $m(t)$ from Example 3.1 and (b) the corresponding DSB signal.

(b) $S_{DSB}(f) = (5/2)M(f-250)+(5/2)M(f+250)$, where $M(f)$ is given by:

$$M(f) = \frac{3}{2}\delta(f-10) + \frac{3}{2}\delta(f+10) + \frac{1}{j}\delta(f-20) - \frac{1}{j}\delta(f+20)$$

Therefore $S_{DSB}(f)$ is given by:

$$\begin{aligned} S_{DSB}(f) \quad = \quad & \frac{15}{4}\delta(f-260) + \frac{15}{4}\delta(f-240) + \frac{5}{2j}\delta(f-270) - \frac{5}{2j}\delta(f-230) \\ & + \frac{15}{4}\delta(f+240) + \frac{15}{4}\delta(f+260) + \frac{5}{2j}\delta(f+230) - \frac{5}{2j}\delta(f+270) \end{aligned}$$

Figure 3.7 (a) shows $|M(f)|$ and Figure 3.7(b) shows $|S_{DSB}(f)|$.

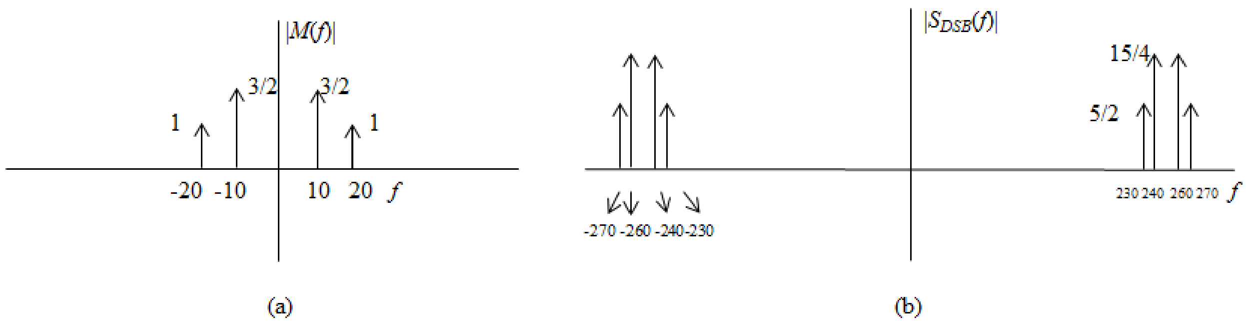


Figure 3.7: (a) $|M(f)|$, (b) $|S_{DSB}(f)|$.

(c) The signal $m(t)$ is a baseband signal. From Figure 3.7, we see that its bandwidth is 20Hz. The signal $s_{DSB}(t)$ is a bandpass signal. Its bandwidth is 270-230 = 40Hz.

(d) The signal $m(t)$ is a power signal with power P_m = 32/2+22/2 = 6.5W. Since $f_c > B_m$, using the properties of power signals (property 5) from chapter 2 then $s_{DSB}(t)$ is a power signal with power $P_s = (5^2/2)P_m$ = 81.25W.

(e) $R_s(\tau) = (A^2/2)R_m(\tau)$ and $R_m(\tau) = (9/2)\cos(2\pi 10\tau)+2\cos(2\pi 20\tau)$ (this was shown in Exercise 2.26). Therefore $R_s(\tau) = 56.25\cos(2\pi 10\tau)+25\cos(2\pi 20\tau)$.

(f) $x(t) = 2s_{DBS}(t)\cos(2\pi 250t) = 10m(t)\cos^2(2\pi 250t) = 5m(t) + 5m(t)\cos(2\pi 500t) = 5(3\cos(2\pi 10t)+2\sin(2\pi 20t)) + 5(3\cos(2\pi 10t)+2\sin(2\pi 20t))\cos(2\pi 500t)$. The Fourier transform of $x(t)$, $X(f) = 5M(f)+(5/2)M(f-500)+(5/2)M(f+500)$, is shown in Figure 3.8. We see that inputting this to a lowpass filter with a passband of at least 20Hz will pass the component corresponding to $5M(f)$ but reject the higher frequency parts as long as it rejects all frequency components above 480Hz. Therefore the output of the lowpass filter will be the inverse Fourier transform of $5M(f)$, which is $5m(t)$.

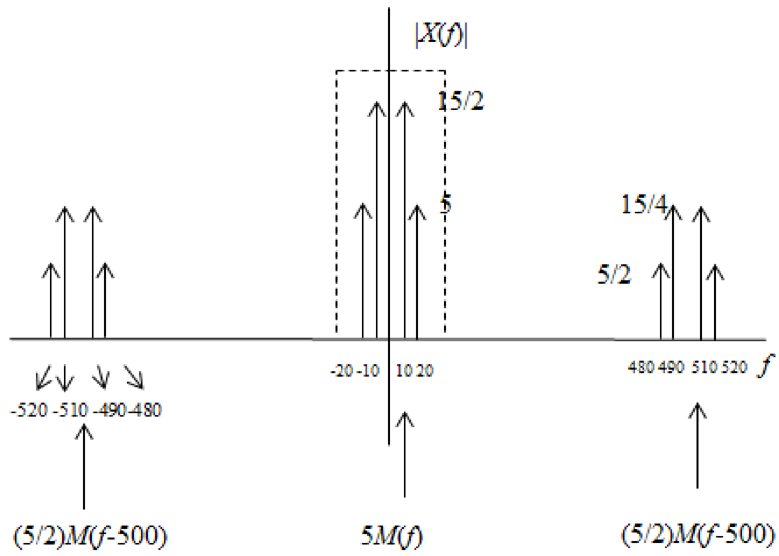


Figure 3.8: $|X(f)|$ of example 3.1.

Exercise 3.1

The signal $m(t) = 4\mathrm{sinc}^2(8t)\cos(2\pi 8t)$ is to be transmitted using DSB modulation.

(a) Find $M(f)$ and sketch it.
(b) If the carrier is $4\cos(2\pi 300t)$, find $s_{DSB}(t)$ and $s_{DSB}(f)$.
(c) What are the bandwidths of $m(t)$ and $s_{DSB}(t)$?
(d) What is the energy of $s_{DSB}(t)$?

Exercise 3.2

If $m(t)$ has Fourier transform $M(f) = 2\Pi((f-10)/8)+2\Pi((f+10)/8)$, then find the Fourier Transform of $s_{DSB}(t) = 3m(t)\cos(2\pi 400t)$. What are the bandwidths of $m(t)$ and $s_{DSB}(t)$?
