2 Frequency conversion

Very often, the signal with the information content (the 'message') is a *baseband signal*, having the significant part of its spectrum around the zero frequency. The purpose of frequency conversion is to shift the message to a different frequency band, thus turning it into a *pass-band* signal. This is done to adapt the transmission to the type of available channel, and/or to place various transmissions simultaneously on the same band (frequency-division multiplexing, or FDM). For example, the filtered audio signal used in an AM radio station (an analog communication system) has a bandwidth of approximately 5 kHz (high-quality audio has a 20-kHz bandwidth), whereas the frequency band assigned to AM broadcasting, which ranges from 550 kHz to 1.6 MHz, must accommodate multiple stations.

The receiver will try to retrieve as accurate as possible a replica of the original message, for which it will be necessary to reverse the process by means of a demodulator.

The simplest frequency conversion process is *amplitude modulation (AM),* also known as Double Side Band (DSB) modulation, with or without carrier, as described below.

2.1 AM with Large Carrier (DSB-LC)

In this type of modulation, the **modulating signal or message** $m(t)$ is multiplied by the **carrier** $A_c \cos(2\pi f_c t+\phi_c)$, and a carrier component (with amplitude A_p) is added to the previous result. We assume that A_p and A_c are positive.

The **modulated (and transmitted) signal** is then given by

$$x(t) = A_c m(t) \cos(2\pi f_c t + \phi_c) + A_p \cos(2\pi f_c t + \phi_c). \tag{2.1}$$

From now on, we will generally assume that $m(t)$ is a low-pass signal or a realization of a low-pass random process, with zero mean and normalized in amplitude ($|m(t)| \leq 1$ for all t).

Figure 2.1 shows an ideal DSB-LC modulator scheme which, as we will see later, is implemented in practice by a circuit known as a *mixer*. In the frequency domain, expression (2.1) translates into:

$$X(f) = \frac{A_c}{2}\left[e^{j\phi_c}M(f-f_c) + e^{-j\phi_c}M(f+f_c)\right] + \frac{A_p}{2}\left[e^{j\phi_c}\delta(f-f_c) + e^{-j\phi_c}\delta(f+f_c)\right]. \tag{2.2}$$

Figure 2.2 illustrates the relation between the modulating and modulated signal spectra, where you can see the two replicas of the original spectrum, centered at $f = \pm f_c$, and the two spectral lines (deltas) at those frequencies. These lines are due to the added carrier term in (2.1), which is sometimes referred to as *additional carrier* or *pilot tone.*

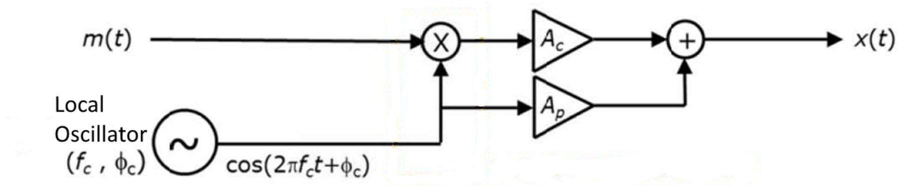

Figure 2.1: Block diagram of a DSB-LC modulator.

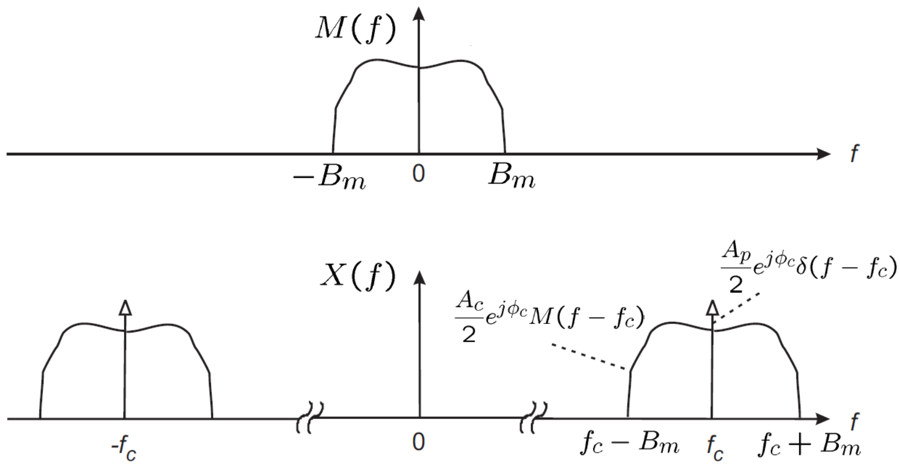

Figure 2.2: An example spectrum of a DSB-LC signal.

The phase ϕ_c represents the initial phase of the oscillator, which is a parameter that is not directly controllable so, in practice, it is usually modelled as a random variable uniformly distributed in the interval $[-\pi, \pi]$. However, in order to simplify the expressions, sometimes we will consider that $\phi_c = 0$, in which case (2.2) reduces to:

$$X(f) = \frac{A_c}{2}\left[M(f - f_c) + M(f + f_c)\right] + \frac{A_p}{2}\left[\delta(f - f_c) + \delta(f + f_c)\right]. \tag{2.3}$$

Defining the **modulation index** as $a = A_c/A_p$ $(a > 0)$, expression (2.1) can be rewritten as:

$$x(t) = A_p[1 + a \cdot m(t)] \cos(2\pi f_c t + \phi_c). \tag{2.4}$$

It can be shown that if $m(t)$ is a realization of a wide-sense stationary (WSS) random process, $M(t)$, with zero mean and autocorrelation $R_M(\tau)$, and ϕ_c is a uniform random variable in the interval $[-\pi, \pi]$, then $x(t)$ is a realization of the random process $X(t) = A_p[1+a\cdot M(t)] \cos(2\pi f_c t+ \phi_c)$, which is also WSS with zero mean and autocorrelation function given by

$$R_X(\tau) = \frac{A_p^2}{2}[1 + a^2 R_M(\tau)] \cos(2\pi f_c \tau). \tag{2.5}$$

The power spectral density (PSD) of $X(t)$ is obtained by taking the Fourier transform of (2.5):

$$\begin{aligned} S_X(f) &= \frac{A_p^2}{4}[\delta(f) + a^2 S_M(f)] \star [\delta(f-f_c) + \delta(f+f_c)] \\ &= \frac{A_p^2}{4}\left[\delta(f-f_c) + \delta(f+f_c) + a^2 S_M(f-f_c) + a^2 S_M(f+f_c)\right], \end{aligned} \tag{2.6}$$

As a consequence, the average power of the transmitted process $X(t)$ is

$$P_X = R_X(0) = \int_{-\infty}^{\infty} S_X(f)df \;=\; \frac{A_p^2}{2}[1 + a^2 P_M] \;=\; \frac{A_p^2}{2} + \frac{A_p^2}{2} a^2 P_M. \tag{2.7}$$

The first term of the previous expression, $\frac{A_p^2}{2}$, represents the power dedicated to just transmitting the carrier term (which does not carry 'useful' information), while the second corresponds to the useful power dedicated to transmitting the message. The **power efficiency** of the modulation is defined as the ratio between the useful and the total transmitted powers:

$$\eta = \frac{\frac{A_p^2}{2} a^2 P_M}{\frac{A_p^2}{2}[1 + a^2 P_M]} = \frac{a^2 P_M}{1 + a^2 P_M}. \tag{2.8}$$

The fact that a significant fraction of the power is 'wasted' in transmitting the carrier tone gives the benefit of making the demodulation process much simpler, by enabling the use of simple device known as **envelope detector,** discussed next.

Envelope detection

An envelope detector is an *incoherent* type of demodulator, which means that it does not require precise knowledge of either the frequency or phase of the carrier.

Given a signal of the form

$$x(t) = u(t)\cos(2\pi f_c t + \phi_c), \tag{2.9}$$

where the bandwidth of $u(t)$ is much smaller than f_c, its **envelope** is defined as

$$v_x(t) = |u(t)|, \tag{2.10}$$

which, in the particular case of DSB-LC, takes the form

$$v_x(t) = A_p|1 + a \cdot m(t)|. \tag{2.11}$$

If the message satisfies $1 + a \cdot m(t) \geq 0 \;\forall t$, then $v_x(t) = A_p(1 + a \cdot m(t)) = A_p + A_c m(t)$, and it is possible to recover $m(t)$ from the envelope $v_x(t)$ just by removing its DC component and rescaling the resulting signal. Figure 2.3 shows a possible implementation of an envelope detector, as well as its behaviour in terms of its time constant (product $R_1 \times C_1$). If the condition $1 + a \cdot m(t) \geq 0$ is not satisfied for all t, then **overmodulation** occurs, and the signal cannot be correctly demodulated with an envelope detector (although it could be by using a more complex *coherent* demodulator). If $|m(t)| \leq 1$ as we have assumed, then the condition to avoid overmodulation amounts to $0 < a \leq 1$. Figure 2.4 illustrates the effect of overmodulation.

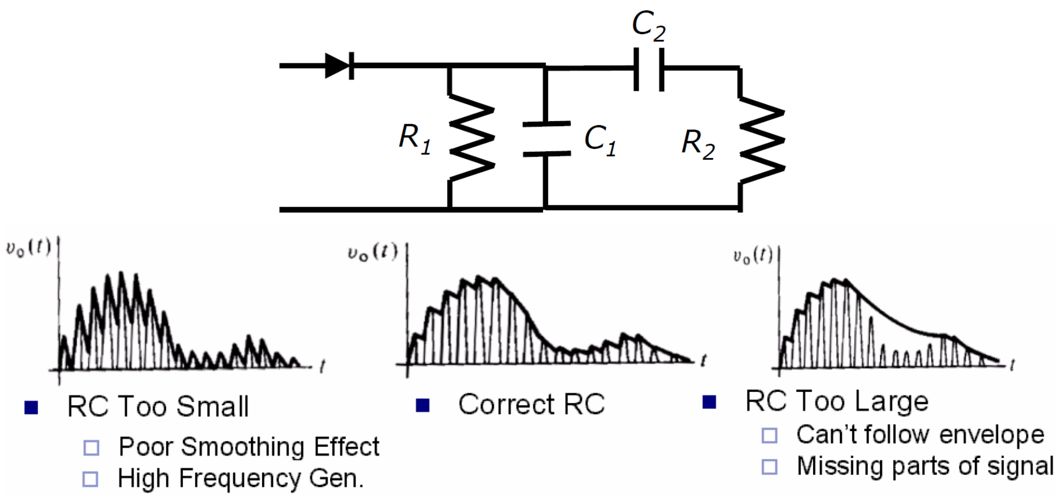

Figure 2.3: Envelope detector (rectifying diode followed by a lowpass filter $R_1||C_1$) plus DC suppressor (capacitor C_2). Influence of the constant R_1C_1 at the output of the lowpass filter.

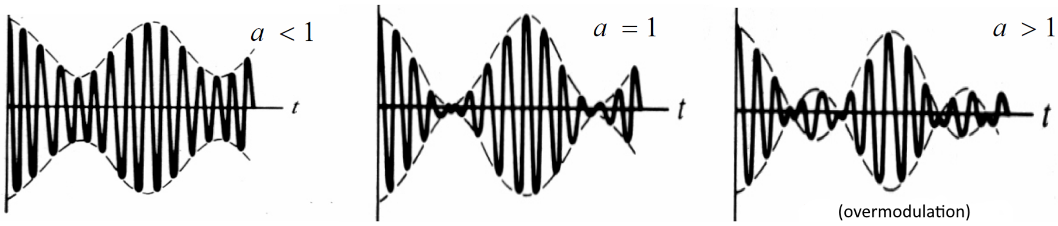

Figure 2.4: Recovered envelope as a function of the modulation index.

2.1.1 AM modulation and demodulation in Matlab

The code in `DSB_LC.m` shows an example where the message is the sum of two sinusoids of frequencies 20 and 70 Hz, and the carrier frequency (1 kHz) is much higher than the message bandwidth (1 kHz $\gg$ 70 Hz).

Run `DSB_LC.m` and observe the time signals corresponding to the message, the carrier, the modulated signal and the output of the envelope detector.

DSB_LC.m

```
time=0.4; Ts=1/10000;                      % sampling interval and time
t=Ts:Ts:time;                              % define a time vector
fc=1000; c=cos(2*pi*fc*t);                 % define carrier at freq fc
m=.7*sin(2*pi*20*t)+.3*cos(2*pi*70*t);     % create "message"
x = 0.6*c .* m + c;                        % modulate with large carrier
fbe=[0 .05 .25 1]; damps=[1 1 0 0]; L=20;  % low pass filter design
h=firpm(L,fbe,damps);                      % impulse response of LPF
envv=(pi/2)*filter(h,1,abs(x));            % find envelope
```

□ Question 1

a) Using `plotspec.m`, observe the spectra of the message $m(t)$, the carrier $c(t)$, and the modulated signal $x(t)$. Do your observations correspond to what you would expect?

b) Plot the detected envelope in the time domain and also in the frequency domain. Has the original message been recovered?

c) If the message bandwidth is B Hz, which is the bandwidth of the modulated signal $x(t)$?

d) Which is the *modulation index* in this case? (Note that the message does not take amplitudes greater than 1 in absolute value).

e) Change the modulation index to 0.9, 1.0 and 1.5 and explain what happens in each case.

■

□ Question 2

DSB-LC has the advantage that the demodulation stage does not need to know the frequency and phase of the received modulated signal accurately. Let us verify this by modifying `DSB_LC.m`.

a) Add a constant phase term to the carrier: use `c=cos(2*pi*fc*t+phase)`, with `phase = 0.1, 0.5, pi/3, pi/2, pi`. How is the output of the envelope detector affected?

b) Change the carrier frequency: `c=cos(2*pi*(fc+g)*t)`, with `g = 10, -10, 100, -100`. Again, how is the output of the envelope detector affected?

■

□ Question 3

a) According to the vector `fbe`, what are the frequencies, in Hz, that indicate the edges of the pass-band, transition band, and stop-band of the implemented low-pass filter? Observe the transfer function of the filter by means of `freqz.m` or `fvtool.m` and verify your answer.

b) What is the time delay introduced by this low-pass filter?

c) Plot the spectrum of the rectified signal $|x(t)|$ using `plotspec.m` . Does it match what you would expect? Do you consider the filter adequate?

d) `DSB_LC.m` simulates a full-wave rectifier. Copy the code to another file `DSB_LC2.m` and modify it to implement an envelope detector with a half-wave rectifier. Note that the half-wave rectified signal can be written as

$$y(t) = \frac{1}{2}[x(t) + |x(t)|] = \begin{cases} x(t), & \text{if } x(t) \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Also replace the gain `pi/2` by `pi`. Does the half-wave envelope detector work properly with this simple change?

e) Using `plotspec.m`, observe the spectrum of the rectified signal $y(t)$. Does it match what you would expect? Modify the code to make the detector work properly.

■

Once again, we must emphasize that DSB-LC has the advantage of using an envelope detector as demodulator (which does not need to know the carrier phase and frequency very precisely), but at the cost of increasing the transmitted power. Therefore, we observe that there is a clear tradeoff between the power efficiency of the modulation and the receiver cost. For example, in AM broadcasting this tradeoff makes sense, because we have a few transmitters (radio stations) and many receivers (listeners), and we want to make the receivers as cheap as possible to obtain a large customer base.

2.2 AM with suppressed carrier (DSB-SC)

In this type of modulation[1], the modulated signal is simply the message $m(t)$ multiplied by the carrier $A_c \cos(2\pi f_c t + \phi_c)$. Thus, the transmitted signal is:

$$x(t) = A_c m(t) \cos(2\pi f_c t + \phi_c) \quad \Rightarrow \quad X(f) = \frac{A_c}{2}\left[e^{j\phi_c} M(f - f_c) + e^{-j\phi_c} M(f + f_c)\right]. \quad (2.12)$$

Demodulation can be accomplished by multiplying $x(t)$ by another sinusoid with the same frequency and phase as the carrier, and then low-pass filtering the result.

However, the receiver will never match exactly either the phase of the carrier or its frequency, because even when the transmitter and receiver oscillators are tuned to a common nominal frequency, their actual frequency fluctuates around this nominal value (the cheaper the oscillator, the larger the fluctuation). In order to mitigate this problem, in practice, communication receivers include a *carrier recovery* block that estimates the frequency and phase of the carrier from the received signal with considerable accuracy.

[1]DSB-SC: *Double SideBand – Supressed Carrier.*

Let us assume that the frequency and phase of the modulator carrier are f_c and zero respectively, while at the demodulator are $f_c + f_\Delta$ and ϕ. The demodulation process involves calculating

$$b(t) = x(t) \cdot \cos\left[2\pi(f_c + f_\Delta)t + \phi\right] \tag{2.13}$$

The estimate of the original message from the previous signal is given by $\hat{m}(t) = \text{LPF}\{b(t)\}$, where the operator $\text{LPF}\{\cdot\}$ denotes low-pass filtering. Assuming that the low-pass filter has a bandwidth a little wider than $m(t)$,

$$\hat{m}(t) = \frac{A_c}{2} m(t) \cos(2\pi f_\Delta t + \phi). \tag{2.14}$$

So, if $f_\Delta \neq 0$ a residual modulation remains in $\hat{m}(t)$. In digital communication systems, if this effect were not compensated, it would result in an excessively high error rate.

On the other hand, even when the receiver is frequency-locked with the transmitter ($f_\Delta = 0$), we get

$$\hat{m}(t) = \frac{A_c}{2} m(t) \cos(\phi). \tag{2.15}$$

Clearly, the factor $\cos(\phi)$ attenuates the signal (or may even cancel it if $\phi = \pm\pi/2$). If $\pi/2 < \phi < 3\pi/2$, it also changes signal polarity. Therefore, it also necessary to employ techniques that allow synchronizing the phases of the transmitter and receiver.

The mo/demodulation process is implemented by the code in `DSB_SC.m`, where it is possible to specify frequency and phase mismatches.

DSB_SC.m

```
time=0.2; Ts=1/10000;                      % sampling interval and time (s)
t=Ts:Ts:time;                              % define a time vector
fc=1000; c=cos(2*pi*fc*t);                 % define carrier at freq fc, Hz
m = .7*sin(2*pi*20*t)+.3*cos(2*pi*70*t);   % message
x = c.*m;                                  % modulation
fd = 0; phi = 0;                           % freq & phase offsets
c2 = cos(2*pi*(fc+fd)*t+phi);              % create cosine for demodulation
b = x.*c2;                                 % demodulate received signal
fbe=[0 0.05 0.25 1]; damps=[1 1 0 0]; L=20; % low pass filter design
h = firpm(L,fbe,damps);                    % impulse response of LPF
mest = 2*filter(h,1,b);                    % LPF the demodulated signal
```

☐ Question 4

a) Execute `DSB_SC.m` and observe the waveforms of the different signals. Notice the slight delay between the retrieved and original message. What is the reason for this?

   Plot and compare the spectra of the original message, the modulated signal, the demodulated signal and the retrieved message. Comment what you observe.

b) Change the carrier frequency to 2 kHz and repeat the previous process. Try again with $f_c = 4.9$ kHz. Can you explain what you observe?

c) With $f_c = 1$ kHz, test different phase mismatches: $\phi = 0,\ 0.4\pi,\ 0.5\pi$. How do they affect the estimated message? Can you predict the system behaviour for $\phi = 0.75\pi$ and $\phi = \pi$?

d) Try now with different frequency mismatches: $f_\Delta = 1.0,\ 10$ and 20 Hz. Determine the value of these frequency offsets in ppm (parts per million). Comment on the quality of the recovered message.

e) What conditions must the transfer function of the low-pass filter meet to ensure that, in the absence of phase and frequency mismatches, the receiver works properly? Do you consider the default low-pass filter adequate?

f) Add noise to the received signal changing the code line `x = c.*m;` by

```
x = c.*m + 0.2*randn(size(t));
```

What do you observe? Try other noise powers.

■

2.3 Quadrature modulation

Both in DSB-LC and DSB-SC modulations, the bandwidth of the transmitted signal is twice the bandwidth of the message (hence the term Double Side-Band). In case of a real-valued message, its spectrum is symmetric and, therefore, there are redundant spectral components in the modulated signal. By means of quadrature modulation it is possible to eliminate this loss of efficiency by transmitting *two* real-valued messages simultaneously in the same bandwidth of a single DSB signal. Actually, it is equivalent to considering a complex-valued message whose real and imaginary parts correspond to each of the two real-valued messages. Obviously, this results in a more efficient use of the available spectrum.

If we denote the two messages by $m_1(t)$ and $m_2(t)$, the quadrature modulated signal is generated as:

$$x(t) = A_c \left[m_1(t) \cdot \cos(2\pi f_c t) - m_2(t) \cdot \sin(2\pi f_c t) \right]. \tag{2.16}$$

According to the previous expression, both messages are independently modulated using DSB-SC by means of two **orthogonal** carriers (i.e, with a phase shift of 90°). The transmitted signal is then obtained by subtracting both DSB-SC signals.

In order to retrieve both messages from $x(t)$, two orthogonal sinusoids are needed again that must be locked in frequency and phase to the carriers used to generate the modulated signal. If so, the demodulated signals are

$$b_1(t) = x(t) \cdot \cos(2\pi f_c t), \qquad b_2(t) = x(t) \cdot \sin(2\pi f_c t), \tag{2.17}$$

and after low-pass filtering, we get

$$\mathrm{LPF}\{b_1(t)\} = \frac{A_c}{2} m_1(t), \qquad \mathrm{LPF}\{b_2(t)\} = -\frac{A_c}{2} m_2(t), \tag{2.18}$$

and, thus, $m_1(t)$, $m_2(t)$ can be recovered (note that we need to change the sign of the second signal). Observe that we have assumed that the receiver oscillator is *perfectly synchronized* (in phase and frequency) with that of the transmitter. If this is not the case, there will be problems with more harmful effects than those in DSB-SC demodulation. Therefore, in a quadrature demodulator it is crucial to have a mechanism that maintains the oscillator of the receiver locked in frequency and phase to that of the transmitter.

☐ **Question 5**

a) Modify `DSB_SC.m` to simulate a quadrature modulation/demodulation system.

b) Introduce a phase shift ϕ in the receiver, so that $b_1(t) = x(t)\cos(2\pi f_c t+\phi)$, $b_2(t) = x(t)\sin(2\pi f_c t+\phi)$. Try several values of ϕ and examine their effect on the retrieved messages.

c) Repeat the previous point for a frequency shift in the receiver.

■

2.4 Heterodyne demodulator

Up to now we have only considered direct demodulation schemes: the band-pass signal, centered at f_c, is moved to baseband by multiplying by a sinusoid of the same frequency f_c. Nevertheless, another possibility, widely used in practice, is to demodulate the received signal in two steps:

1. First the modulated spectrum is moved and centered at a lower frequency, termed *intermediate frequency*, f_{IF}.

2. Then, the intermediate-frequency signal is moved to baseband by means of an ordinary demodulator.

The translation to intermediate frequency can be carried out by multiplying by a sinusoid of frequency $f_c - f_{\mathrm{IF}}$ (known as *low-side injection*) or of frequency $f_c + f_{\mathrm{IF}}$ (*high-side injection*). In both cases the resulting signal has to be filtered with a bandpass filter centered at f_{IF} to remove other unwanted frequency components.

☐ **Question 6**

a) Consider a transmission system that employs a DSB-SC modulation with carrier frequency $f_c = 2$ kHz. Write a demodulation routine by first mixing the modulated signal with a cosine of frequency 1.5 kHz and then with another one of frequency 0.5 kHz.

- Use as message the sum of two tones of 70 and 100 Hz with different amplitudes and phases.
- Pay attention to the necessary sampling rate in order to avoid aliasing along the simulation.
- Use an IF filter of order `L=80`. Design the filters to properly recover the message.

b) Let us now investigate the effect of interfering signals at the receiver input. Modify your routine to observe what happens when including an additive interfering tone of frequency (a) 2.5 kHz, (b) 1.6 kHz, (c) 0.99 kHz. Is the recovered message equally affected in all these cases? Explain the reason.

c) In order to mitigate the interference problem (image frequency) we must place a band-pass filter before downconverting to intermediate frequency. Determine the specifications of this filter.

■