9.1 Amplitude Modulation

We consider amplitude modulation (AM). One of its prominent applications is radio broadcast.

9.1.1 Frequency Translation

Consider the transmission of a baseband message signal $m(t)$. For example, $m(t)$ can represent a piece of music. The frequency content of $m(t)$, represented by its Fourier transform $M(f)$, may contain only low frequencies up to about 20 kHz. However, in order to efficiently transmit the message through the atmosphere, it is desirable to alter the frequency content of the signal so that it centers around a much higher frequency. An intuitive approach is to modulate the message signal $m(t)$ onto a carrier of high frequency ($f_c >> W$) to give the modulated signal

$$s(t) = Am(t) \cos(2\pi f_c t).$$

The Fourier transform of $s(t)$ is given by

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

Notice that the frequency content of $s(t)$ centers around $f_c$ and $-f_c$ as shown in Fig. 9.2.
9.1.2 Amplitude Modulation

Consider the time domain again. If \( m(t) \) is always positive, then it can be easily recovered from the envelope of \( s(t) \) (Fig. 9.3).

![Figure 9.3: Positive \( m(t) \) and the corresponding \( s(t) \)](image)

However, if \( m(t) \) can be positive and negative, then the recovery is not so easy (Fig. 9.4).

![Figure 9.4: Arbitrary \( m(t) \) and the corresponding \( s(t) \)](image)
In order to guarantee that the signal is always positive before being multiplied onto the carrier, we devise the following modulation

\[ s(t) = A[1 + am(t)] \cos(2\pi f_c t) \]  

(9.1)

where the modulation index \( a < 1 \), and the message signal \( m(t) \) is normalized so that its maximum magnitude is 1. It is the commonly used amplitude modulation (AM).

### 9.1.3 Common Modulators and Demodulators

The main advantage of AM, compared with other modulation methods, is the availability of simple demodulators.

**Switching modulator**

Consider modulator shown in Fig. 9.5. We assume that the diode is ideal.

![Switching Modulator Diagram](image)

**Figure 9.5: Switching Modulator**

Then the output of the modulator is given by

\[ v(t) = [A \cos(2\pi f_c t) + m(t)]_+ \]

where the operator \([·]_+\) is defined by

\[ [x(t)]_+ = \begin{cases} 
  x(t) & \text{if } x(t) \geq 0, \\
  0 & \text{if } x(t) < 0. 
\end{cases} \]
We further assume that $|m(t)| << A$ for all $t$. Then the operation of the ideal diode can be approximated by the multiplication of the switching function

$$
c(t) = \begin{cases} 
1 & \text{if } nT_c - \frac{T_c}{4} < t < nT_c + \frac{T_c}{4}, \\
0 & \text{else,}
\end{cases}
$$

where $T_c = 1/f_c$ is the period of the carrier.

The switching function $c(t)$ is itself periodic with period $T_c$. Therefore, it can be represented by a Fourier series.

$$
c(t) = \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} \cos[2\pi f_c t (2n-1)]
$$

The output of the switching modulator $v(t)$ is given by

$$
v(t) = \frac{A}{\pi} + \frac{m(t)}{2} + \frac{A}{2} \cos(2\pi f_c t) + \frac{2}{\pi} m(t) \cos(2\pi f_c t) + \text{double and higher frequency terms}
$$

Finally, the desired AM component can be obtained by passing $v(t)$ through a bandpass filter with center frequency $f_c$.

**Envelope detector**

The envelope detector is shown in Fig. 9.7.
The diode cuts off negative cycles while the R-C pair acts as a low pass filter to obtain the envelope. Notice that the time constant of the R-C pair should be chosen so that

\[ \frac{1}{f_c} << RC << \frac{1}{W} \]

where \( W \) is the message bandwidth. In this case, the R-C pair would respond fast enough to track the message (the envelope), but slow enough to ignore the variations due to the carrier.

### 9.1.4 Disadvantages of AM

The main disadvantages of AM are

- **Wasteful of power:**

  Consider the spectrum of the AM signal. Notice the presence of the impulses at \( \pm f_c \). Power is wasted to transmit the unmodulated carrier. Let \( P_m \) be the power of the message \( m(t) \), i.e,

  \[ P_m = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} m^2(t) dt \]

  The AM signal \( s(t) \) is given by

  \[ s(t) = A[1 + am(t)] \cos(2\pi f_c t) \]

  Then the transmitted power of the AM signal is given by

  \[
  P_s = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} A^2 [1 + am(t)]^2 \cos^2(2\pi f_c t) dt \\
  = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} A^2 \left[ \frac{1 + am(t)^2}{2} \right] dt \\
  = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} A^2 \left(1 + a^2 m^2(t)\right) dt \\
  = \frac{A^2}{2} + \frac{A^2 a^2}{2} P_m
  \]
The second line is justified when \( f_c \gg W \), the highest frequency of \( m(t) \). The third line is justified when the time average of \( m(t) \) is zero, which implies the message contains no power at DC. The power to transmit the message without the carrier, i.e, to transmit

\[
\tilde{s}(t) = Aam(t) \cos(2\pi f_c t),
\]

is given by

\[
P_{\tilde{s}} = \frac{A^2 a^2}{2} P_m.
\]

Therefore, the amount \( A^2/2 \) is wasted.

E.g., The message is a single tone, i.e., \( m(t) = \cos(2\pi f_0 t) \) where \( f_0 \ll f_c \). Then \( P_m = 1/2 \).

The power to transmit the AM signal is \( A^2 + \frac{A^2 a^2}{4} \). The power to transmit the message without the carrier is \( A^2 a^2/4 \). Therefore the ratio of useful power to total power is \( \frac{a^2}{a^2+2} \). Since \( a < 1 \), at most one third of the transmitted power is useful.

- **Wasteful of bandwidth**

  Consider a baseband message signal \( m(t) \) with bandwidth \( W \). When it is modulated with AM, the transmitted AM signal occupies the frequency band from \( f_c - W \) to \( f_c + W \) with a bandwidth of \( 2W \). The additional bandwidth turns out to be unnecessary.

## 9.2 Double Sideband Suppressed Carrier

Power is wasted in sending the unmodulated carrier in AM. In order to save power, we can suppress the carrier. The resulting modulation is called double sideband suppressed carrier (DSB-SC) modulation.

\[
s(t) = Am(t) \cos(2\pi f_c t)
\]

(9.2)

Notice that the result is the signal we have right from the beginning — a product of the message signal and the carrier. To generate the DSB-SC signal, we can use a product modulator (or mixer).

### 9.2.1 Detection of DSBSC Signals

As we have noted before, the message cannot be directly recovered from the envelope of the DSB-SC signal. To demodulate the signal, we consider the receiver in Fig. 9.9.
Suppose that the local oscillator is tuned at frequency $f_c$ and phase-aligned with the received signal. Then the output of the product modulator is

$$Am(t) \cos^2(2\pi f_c t) = \frac{A}{2} m(t) + \frac{A}{2} m(t) \cos(2\pi 2f_c t).$$

If $W < f_c$ (which is almost always the case), then a low-pass filter (LPF) with cutoff frequency at $W$ will remove the double carrier term. The result is the message signal.

With this approach, we need to generate a local reference of the carrier with the same frequency and the same phase. This is called coherent detection. Consider that we generate a local carrier with a wrong phase. As an extreme example, we generate $\sin(2\pi f_c t)$ as our local reference. Then the output of the product modulator is

$$\frac{A}{2} m(t) \sin(2\pi 2f_c t)$$

which will be filtered away by the LPF. The output will be zero and the demodulation fails.

The problem becomes complicated because of the following reasons:

- Unknown delay is introduced due to time needed for the signal to propagate from the transmitter to the receiver. The delay manifests itself as a phase shift.
- Phase shift may also introduced by the channel. For example, when a signal is reflected by an object, phase shift is often introduced.
- Even at the transmitter, the phase of the carrier may be unknown.

In general, the received signal is of the form

$$Am(t) \cos(2\pi f_c t + \theta).$$
(Actually, θ may also be varying slowly due to changes in channel conditions.) Therefore, we have to track the phase of the carrier. One approach is to use the Costas loop in Fig. 9.10.

![Costas Loop Diagram](image)

Figure 9.10: Costas Loop

It tracks the phase, as well as small offsets of the carrier frequency (which is a form of slow-varying phase).

### 9.2.2 Quadrature carrier multiplexing

Notice from the above discussion that if the received carrier is \( \cos(2\pi f_c t) \) and the local reference is \( \sin(2\pi f_c t) \), then the output is zero. Similarly, if the received carrier is \( \sin(2\pi f_c t) \) and the local reference is \( \cos(2\pi f_c t) \), then the output is also zero. Therefore, we can send two message signals at the same time by modulating them with sine and cosine respectively.

\[
Am_1(t) \cos(2\pi f_c t) + Am_2 \sin(2\pi f_c t)
\]  

(9.3)

These two carriers are called quadrature carriers. We can recover \( m_1(t) \) and \( m_2(t) \) separately with two coherent receivers (or a combined receiver). Notice that we are sending twice as much information without using more bandwidth.
9.3 The Superheterodyne Receiver

A practical receiver in a broadcast system should be equipped with the following functions:

- **Carrier frequency tuning**
  
  It should allow the user to change the carrier frequency to select the desired signal.

- **Filtering**
  
  It should separate the desired signal from other unwanted signals.

- **Amplification**
  
  It should provide amplification to compensate for the power loss during the course of transmission.

- **Demodulation**
  
  It should demodulate the modulated signal to give the message signal.

A conceptually simple approach is shown in Fig. 9.11.

The radio frequency (RF) section is a tunable high-gain filter that can effectively reject adjacent channels at RF frequencies. However, such a filter is difficult to build.
To overcome this problem, the superheterodyne receiver is commonly used. A block diagram of the superhet is shown in Fig. 9.12.

![Block Diagram of Superhet Receiver]

Its main feature is that the incoming signal is translated from the carrier frequency to a predetermined intermediate frequency (IF) before demodulation is performed. The RF section is a BPF that allows tuning, provides high gain, and, to a certain degree, suppresses unwanted channels. Suppose that the carrier frequency of the desired signal is $f_{RF}$. The local oscillator generates a local carrier at frequency $f_{RF} - f_{IF}$. The output of the mixer consists of two copies of the desired signal – one centers at $f_{IF}$ and one centers at $2f_{RF} - f_{IF}$. The IF section consists of a BPF filter with sharp cutoff that only allows passage of frequencies around $f_{IF}$. The IF section carries the main responsibility of adjacent channel rejection. Demodulation then follows.

Notice that a signal with carrier frequency $f_{RF} - 2f_{IF}$ will produce two copies at the output of the mixer – one centers at $f_{IF}$ and one centers at $2f_{RF} - 3f_{IF}$. Such a signal is called an image signal. The copy at $f_{IF}$ at the output of the mixer will interfere with the desired signal. To avoid this problem, the image signal must be filtered off at the RF section. Notice that if $f_{IF}$ is small, it may be difficult to build the RF section. On the other hand, if the $f_{IF}$ is large, it may be difficult to build the IF section. Therefore, a compromise has to be made.
Figure 9.13: Functions of RF and IF sections