DRAFT
ELEC 350, fall 2015
Signals and Modulation

1 IQ Signals

1.1 Carrier waves, amplitude and phase

Learning objectives
A communications signal carries a message from point A to point B.
In this section, we learn the mathematics of a general communications signal waveform comprising a carrier wave whose amplitude and phase is modified (modulated) in step with a message signal. The communications signal may be written as the real part of a complex waveform with real and imaginary parts.

Radio waves are used to carry a message over a distance determined by the link budget. The radio wave (called a carrier wave) is "modulated" (modified) by the message signal $m(t)$; in other words the amplitude and/or phase of the carrier wave is modified so as to include the information stored within the message.

The general form of a radio signal (or any communications signal) is as follows:

$$s(t) = \mathrm{Re}\{a(t)e^{i\theta(t)}\} = a(t)\cos\theta(t),$$

where $a(t)$ represents the amplitude of the signal after modulation and $\theta(t)$ is the phase of the carrier wave. The message is contained within $a(t)$ and $\theta(t)$ in a manner to be described later. The message may be analog or digital.

Note that the instantaneous frequency $f_i(t)$ of the signal is related to the phase $\theta(t)$ via

$$f_i(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt}$$

In the special case where there is no modulation applied to the carrier signal (ie: no message sent), then we say the carrier wave is "unmodulated" and we write $s(t) = c(t)$ is a carrier wave oscillating at a frequency of f_c and scaled by the carrier amplitude coefficient A_c.

In this special case, the amplitude $a(t) = A_c$ is a constant, and the phase increases linearly with time such that $\theta(t) = 2\pi f_c t$. The instantaneous frequency is exactly the carrier frequency

$f_i(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt} = f_c$ as expected.

In this case, we write

$a(t) = A_c$ (constant independent of time)

$\theta(t) = 2\pi f_c t$ (linear increase of phase with time, 2π radians every $1/f_c$ seconds.

therefore

$c(t) = A_c \cos(2\pi f_c t) = A_c \cos(\omega_c t)$

This signal is called the *carrier wave* $c(t)$ and can also be visualized as a phasor with angular frequency $\omega_c = 2\pi f_c$ and with period $T = 1/f_c$.

In general, when a message is sent and the carrier wave is modulated, the amplitude $a(t)$ of the carrier wave may be time varying, and the phase of the carrier $\theta(t)$ may have a time varying phase component $\varphi(t)$ that is added to the linear phase, thus

$$s(t) = \mathrm{Re}\{a(t)e^{i\theta(t)}\} = a(t)\cos(\theta(t))$$
$$\theta(t) = 2\pi f_c t + \varphi(t)$$
$$s(t) = a(t)\cos(2\pi f_c t + \phi(t)) = \mathrm{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\},$$

The radio signal $s(t)$ is a cosine wave at frequency f_c with time-varying amplitude and phase $a(t), \phi(t)$ It is useful to write the radio signal as the real part of a complex waveform. In the equation above, the complex exponentials are written in polar form showing amplitude and phase.

We define the *complex envelope* $a(t)e^{j\phi(t)}$

To encode a message $m(t)$ on the carrier wave we vary $a(t)$ and/or $\phi(t)$ in step with the message $m(t)$. Thus $a(t), \phi(t)$ are specified as a function of the message $m(t)$. These functions will be specified later when we discuss specific modulation types.

We can write the radio signal in a form where the complex envelope is separated into its real and imaginary parts $a(t)e^{j\phi(t)} = I(t) + jQ(t)$

Exercise: write $a(t)$ and $\phi(t)$ as a function of $I(t), Q(t)$

Using the identity $\cos(A+B) = \cos A \cos B - \sin A \sin B$

the general radio signal $s(t) = a(t)\cos(2\pi f_c t + \phi(t)) = \mathrm{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\}$ may also be written

$$s(t) = a(t)\cos 2\pi f_c t \cos\phi(t) - a(t)\sin 2\pi f_c t \sin\phi(t)$$
$$= I(t)\cos 2\pi f_c t - Q(t)\sin 2\pi f_c t$$

where

$$I(t) = a(t)\cos\phi(t) = \mathrm{Re}\{a(t)e^{j\phi(t)}\},$$
$$Q(t) = a(t) sin\phi(t) = \mathrm{Im}\{a(t)e^{j\phi(t)}\}$$

are called the "in-phase" and "quadrature" components, respectively,

and thus $I(t) + jQ(t) = a(t)e^{j\phi(t)}$

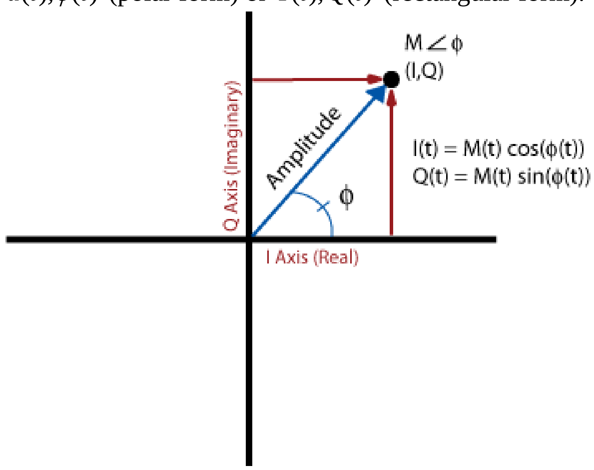
Thus we can write

$$s(t) = \mathrm{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\} =$$
$$\mathrm{Re}\{[I(t) + jQ(t)][\cos 2\pi f_c t + j sin 2\pi f_c t]\}$$
$$= I(t)\cos 2\pi f_c t - Q(t) sin 2\pi f_c t$$
$$= a(t)\cos[2\pi f_c t + \phi(t)]$$

The general radio signal $s(t)$ must be a real signal that we can view on an oscilloscope. We can write $s(t)$ as the real part of a complex signal. From the above equation, we see that $s(t)$ can be described as either

- a cosine wave with amplitude $a(t) \geq 0$ and phase $\phi(t)$, or
- the sum of a cosine wave with amplitude $I(t)$ and a sine wave with amplitude $Q(t)$, where $I(t), Q(t)$ can be greater or less than zero.

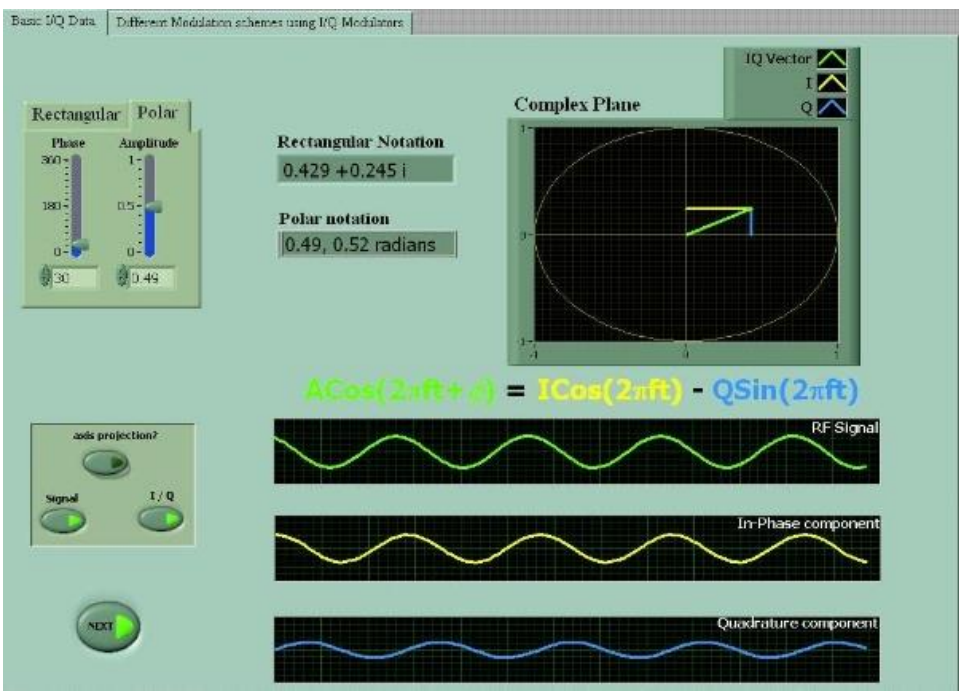
In both cases, the message is a two-dimensional (complex) signal represented using either $a(t), \phi(t)$ (polar form) or $I(t), Q(t)$ (rectangular form). In the figure, $M(t) = a(t)$.



$I(t), Q(t)$ are functions of the message signal(s), where the exact function depends on the modulation type. A simple example is to consider the message signal to be a stereo (2-channel) music signal written as $m_L(t), m_R(t)$ and choose $m_L(t) = I(t), m_R(t) = Q(t)$

$I(t) + jQ(t) = a(t)e^{j\phi(t)} = \tilde{s}(t)$ is the so-called *complex baseband* signal that is a function of the message. We previously also called this the complex envelope. For a stereo music signal, the complex baseband signal is $m_L(t) + jm_R(t)$. The radio signal $s(t) = a(t)\cos[2\pi f_c t + \phi(t)]$ is the so-called *real passband* signal that contains the message modulated onto the carrier wave at frequency f_c

In the special case where $I(t), Q(t)$ are both constants I, Q, then $s(t)$ is the sum of a cos wave and a sin wave with different amplitudes, which is a cosine wave with constant amplitude and phase. The figure below shows the radio frequency (RF) signal $s(t) = I\cos 2\pi f_c t - Q\sin 2\pi f_c t = a\cos 2\pi f_c t + \phi$ where both a, ϕ are functions of I, Q The in-phase component $I\cos 2\pi f_c t$ and the quadrature component $Q\sin 2\pi f_c t$ are also cosine waves at frequency f_c with constant amplitude and phase.



<u>Question</u>: write an expression for both a, ϕ as a function of I, Q

<u>Summary</u>

A radio (or other communications) signal may be made up of a cosine wave at a carrier frequency f_c with time varying amplitude and phase. The signal may be written as a *real passband* signal $s(t) = a(t)\cos[2\pi f_c t + \phi(t)]$, where the message is contained in $a(t), \phi(t)$. The signal may also be written in *complex baseband* form as $\tilde{s}(t) = a(t)e^{j\phi(t)} = I(t) + jQ(t)$ and viewed in the complex plane. The complex baseband signal contains the amplitude and phase only (i.e. the message information) and does not explicitly include the carrier frequency f_c The real passband signal may be obtained from the complex baseband signal via $s(t) = \mathrm{Re}\{\tilde{s}(t)e^{j2\pi f_c t}\}$

A description of message types (analog and digital) and how they are contained in $a(t), \phi(t)$ is explained in detail a later section. A brief example: consider a digital message consisting of alternating 101010 bits. We can choose a modulation scheme as follows: assign $\phi(t) = 0$ during a 0 bit and $\phi(t) = \pi$ during a 1 bit and keep $a(t)$ constant for both 1 and 0 bits. We could also make different choices.

Exercise: sketch a graph of the passband signal waveform with this example modulation scheme.
Exercise: invent another (different) modulation scheme where we assign $a(t), \phi(t)$ as a function of the message, and sketch the passband signal waveform.
Exercise: invent a modulation scheme where we assign $I(t), Q(t)$ as a function of the message, and sketch the passband signal waveform.

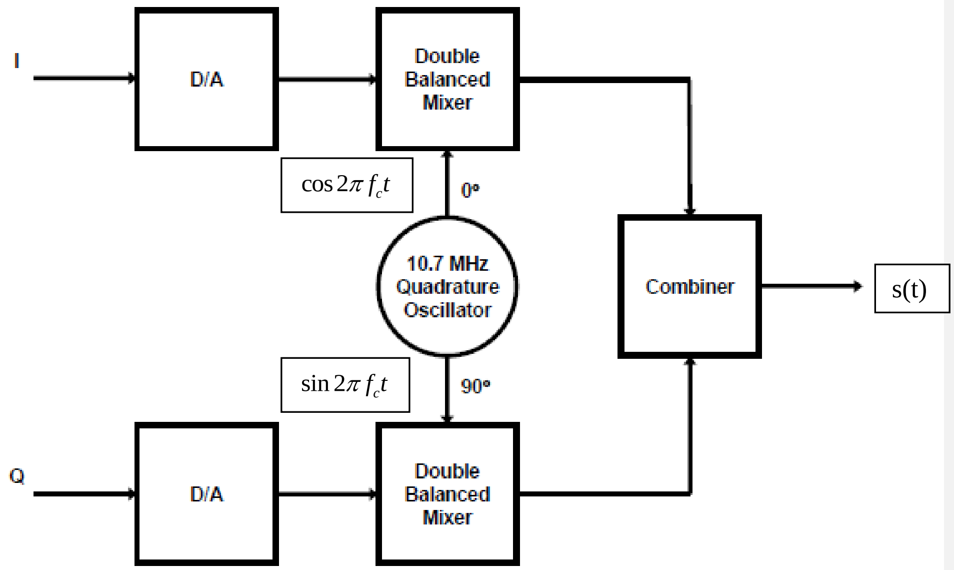
1.2 Block diagrams for real passband and complex baseband signals

In this section, we show block diagrams that implement the equations from the previous section.

1.2.1 Radio transmitter (modulator)

In the figure below, we show the following blocks

- *D/A* is a digital to analog converter,
- *double balanced mixer* is a multiplier
- *quadrature oscillator* has two outputs, a cos wave and a sin wave.
- *Combiner* is an adder



With the blocks defined as above, the figure shows a radio transmitter (or modulator) that produces the radio waveform $s(t) = I(t)\cos 2\pi f_c t - Q(t)\sin 2\pi f_c t$ from the message signals $I(t), Q(t)$, where in this example, $f_c = 10.7 MHz$

In this case, the blocks process analog signals. However, if the D/As are omitted, the remaining blocks process digital signals, using for example GNURadio or Matlab Simulink.

The radio transmitter (modulator) may also be described in complex form with a complex multiplication of the complex baseband message

$I(t) + jQ(t)$ with $\cos 2\pi f_c t + jsin2\pi f_c t$

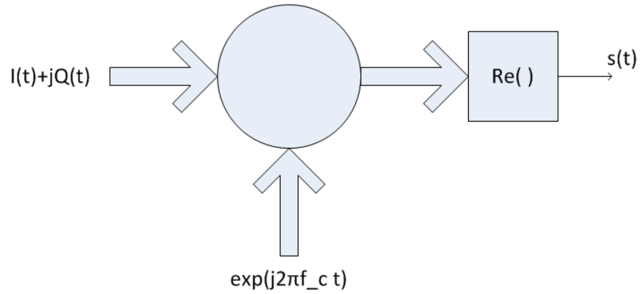
to yield a complex signal

$\hat{s}(t) = a(t)e^{j\phi(t)}e^{j2\pi f_c t} = [I(t) + jQ(t)][\cos 2\pi f_c t + jsin2\pi f_c t]$

whose real part is

$s(t) = I(t)\cos 2\pi f_c t - Q(t)sin2\pi f_c t$

The complex message $I(t), Q(t)$ s multiplied by the complex carrier wave
$e^{j2\pi f_c t|} = \cos 2\pi f_c t - jsin2\pi f_c t$
The radio signal $s(t)$ is the real part of this complex multiplication
$s(t) = I(t)\cos 2\pi f_c t - Q(t)\sin 2\pi f_c t$



The complex baseband diagram assumes the blocks are processing digital signals.

1.2.2 Radio receiver (demodulator)

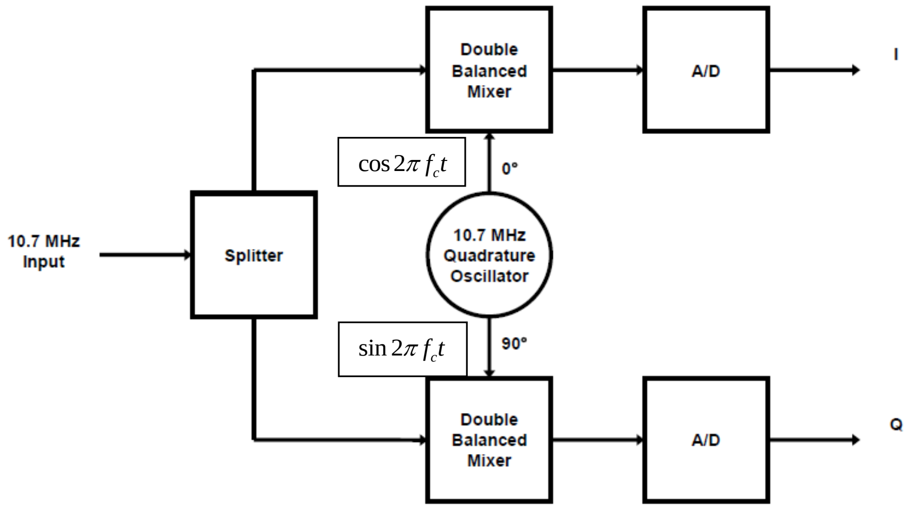
The signal $s(t)$ is transmitted over a distance via some channel (wired or wireless), attenuated by the path loss L_0 and picked up by a receiver in the form $r(t) = s(t)/L_0$
The receiver's task is to recover the message signals $I(t), Q(t)$ from the signal $r(t)$. This can be done using the receiver shown below.

In the figure below, we show the following blocks

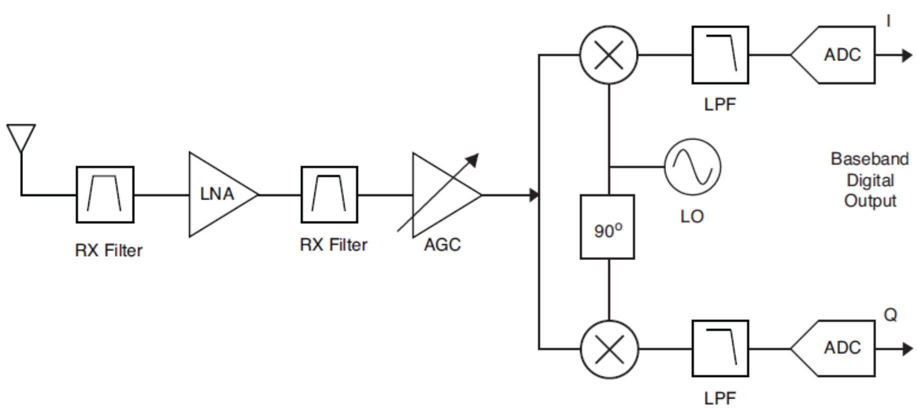
- *splitter* simply duplicates the input signal at the two outputs
- *double balanced mixer* is a multiplier

- quadrature oscillator has two outputs, a cos wave and a sin wave.
- A/D is an analog-to-digital converter

In this receiver, the input signal and all processing is analog, and $I(t), Q(t)$ are digitized by an analog-to-digital converter (A/D or ADC).



A more complete drawing of the receiver adds some practical components



In this figure, we have added: RX filters needed to filter out undesired signals on nearby frequencies, a Low Noise Amplifier (LNA) to amplify $r(t)$ that is typically in the microvolt range to a level in the volt range suitable for ADC, Automatic Gain Control

(AGC) to adjust the gain to compensate for variations in the level $r(t)$ and low pass filters (LPF) before the ADC.

To see how this receiver works, we calculate the signals $x(t), y(t)$ at the two ADC inputs, and find that they are equal to $I(t), Q(t)$

Exercise: prove this. Several trigonometric identities will be needed:

$$\cos\alpha\cos\beta = [\cos(\alpha-\beta)+\cos(\alpha+\beta)]/2$$
$$\sin\alpha\cos\beta = [\sin(\alpha-\beta)+\sin(\alpha+\beta)]/2$$
$$\sin\alpha\sin\beta = [\cos(\alpha-\beta)-\cos(\alpha+\beta)]/2$$

Also, the double frequency terms (cosine waves at $2f_c$) are filtered out by the low pass filter (LPF).

The radio receiver (demodulator) may also be described in complex form with a complex multiplication of the complex passband signal $a(t)e^{j\phi(t)}e^{j2\pi f_c t}$. Recall the real passband signal $s(t) = a(t)\cos[2\pi f_c t + \phi(t)] = \mathrm{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\}$ which is the real part of the complex passband signal.

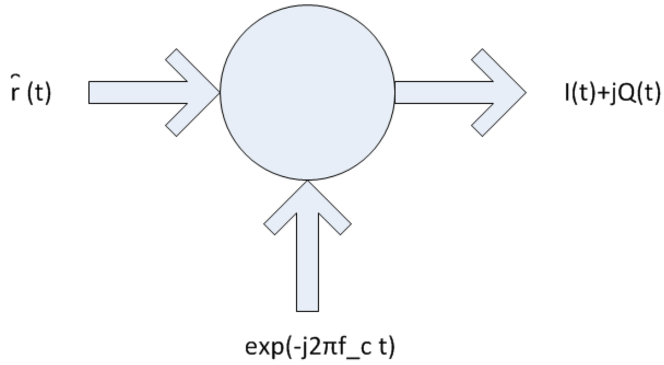
We write the received signal in complex passband form

$$\hat{r}(t) = L_0^{-1}a(t)e^{j\phi(t)}e^{j2\pi f_c t} =$$
$$L_0^{-1}[I(t)+jQ(t)][\cos 2\pi f_c t + jsin 2\pi f_c t]$$

The receiver demodulates the complex radio signal by multiplying $\hat{r}(t)$ by the complex local oscillator

$e^{-j2\pi f_c t|} = \cos 2\pi f_c t - jsin 2\pi f_c t$ to yield

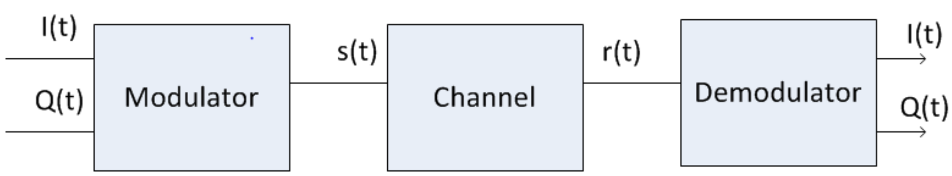
$$\hat{r}(t)e^{-j2\pi f_c t} = [L_0^{-1}a(t)e^{j\phi(t)}e^{j2\pi f_c t}]e^{-j2\pi f_c t} = L_0^{-1}a(t)e^{j\phi(t)} = L_0^{-1}[I(t)+jQ(t)]$$



1.2.3 Summary

We have created a communication system with message signals $I(t), Q(t)$ that are modulated onto a carrier wave at frequency f_c to create the radio signal $s(t) = I(t)\cos 2\pi f_c t - Q(t)\sin 2\pi f_c t$. $s(t)$ travels over a distance via a channel with path loss L_o. The receiver picks up the signal $r(t) = s(t)/L_0$ and recovers the messages $I(t), Q(t)$

IQ transmitter-receiver system



Any signal $s(t)$ can be written as a carrier wave at frequency f_c with time-varying amplitude and phase, i.e.

$$
\begin{aligned}
s(t) &= a(t)\cos[2\pi f_c t + \phi(t)] \\
&= \mathrm{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\} \\
&= \mathrm{Re}\{[I(t) + jQ(t)][\cos 2\pi f_c t + jsin2\pi f_c t]\} \\
&= I(t)\cos 2\pi f_c t - Q(t)sin2\pi f_c t \\
&= a(t)\cos[2\pi f_c t + \phi(t)]
\end{aligned}
$$

where

$$
\begin{aligned}
I(t) &= a(t)\cos\phi(t) = \mathrm{Re}\{a(t)e^{j\phi(t)}\}, \\
Q(t) &= a(t)sin\phi(t) = \mathrm{Im}\{a(t)e^{j\phi(t)}\}
\end{aligned}
$$

and

$$\tilde{s}(t) = I(t) + jQ(t) = a(t)e^{j\phi(t)}$$

is called the complex envelope of the signal and also called the complex baseband signal. The complex envelope contains two real waveforms $I(t), Q(t)$ that contain the information or message.

The real passband signal $s(t)$ is obtained by multiplying the complex envelope $\tilde{s}(t)$ with the complex carrier wave

$$e^{j2\pi f_c t} = cos2\pi f_c t + jsin2\pi f_c t$$

and taking the real part to yield

$$s(t) = \mathrm{Re}\{\tilde{s}(t)e^{j2\pi f_c t}\}$$

1.2.4 Review

Here we review the general I-Q receiver configuration that may be implemented for all types of signals using slightly different notation. Every type of signal consists of a standard configuration coupled with a “map” that is specific to the transmission type.

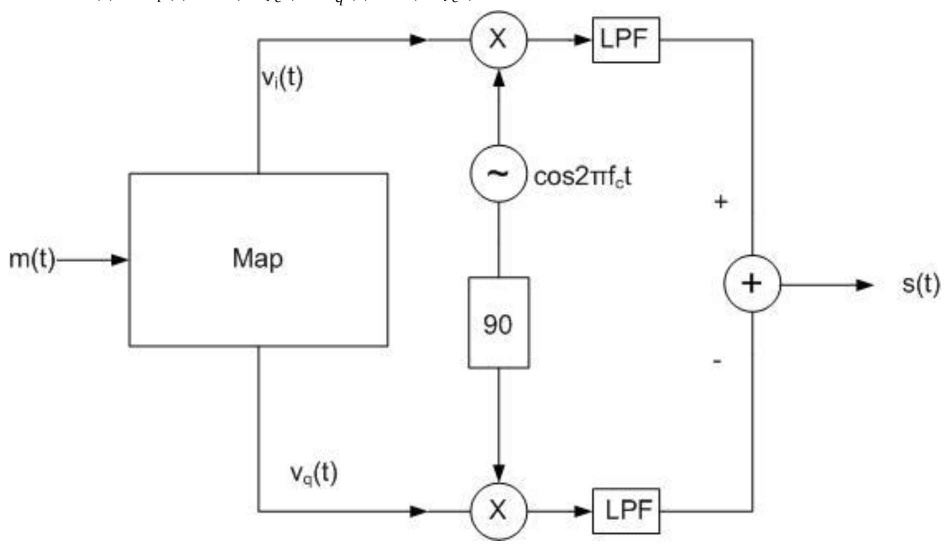
$$s(t) = a(t)\cos[2\pi f_c t + \psi(t)]$$
$$s(t) = a(t)[\cos(2\pi f_c t)\cos\psi(t) - \sin(2\pi f_c t)\sin\psi(t)]$$
$$s(t) = a(t)\cos\psi(t) cos 2\pi f_c t - a(t)\sin\psi(t)\sin 2\pi f_c t$$
$$s(t) = v_i(t) cos 2\pi f_c t - v_q(t)\sin 2\pi f_c t$$

where $a(t)$ is the amplitude and $\psi(t)$ is the phase. The coefficients of the carrier waves in the transmitter are referred to as $v_i(t)$ and $v_q(t)$ respectively, where

$$v_i(t) = a(t)\cos\psi(t)$$
$$v_q(t) = a(t)\sin\psi(t)$$

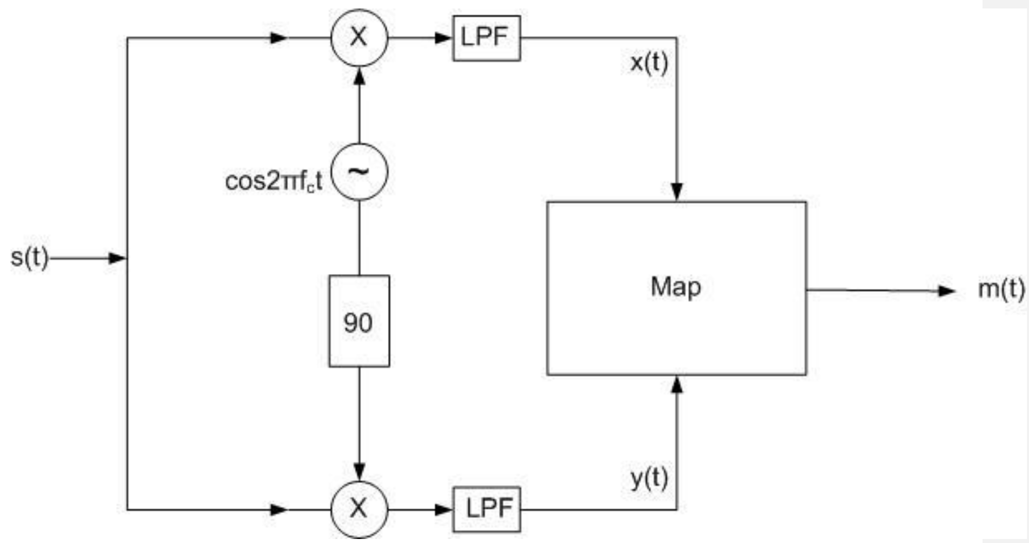
These signals are formed by the map and used to create $s(t)$ at the IQ Transmitter.

$$s(t) = v_i(t)\cos(2\pi f_c t) - v_q(t)\sin(2\pi f_c t)$$



General IQ Transmitter

The IQ Receiver multiplies the incoming signal $s(t)$ by two versions of the carrier wave functions: $\cos(2\pi f_c t)$ and the 90 degree phase-shifted version $\sin(2\pi f_c t)$.



General IQ Receiver

1.2.5 Messages

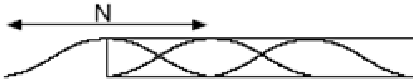
When we study modulation, we often choose a simple analog message $m(t) = A_m \cos(2\pi f_m t)$, where f_m is the modulation/message frequency and is usually on the order of Hz or kHz. This message is called a single frequency or single tone message. In practice, we wish to transmit a more interesting message that contains more than a single tone at a single frequency.

A general analog message (voice or music) would be represented by a sum of cosine waves with different amplitudes and phases A_i, ψ_i at each frequency f_i

$$m(t) = \sum_i A_i(t) \cos(2\pi f_i t + \psi_i(t))$$

We often assume that $m(t)$ is divided into frames of length T in the range 5-20 milliseconds. The frames may simply follow one after the other without overlap.

Frames may also be windowed and overlapped. In the figure, we show each frame of length T as being N samples long, where N is typically a power of 2 in the range 256 to 2048 for audio signals sampled at an audio rate of say $N/T = 48{,}000$ Hz. The figure also shows the window shape (raised cosine) and the overlap (N/2 samples). *Exercise*: what is the length in milliseconds of an N=256 sample frame?

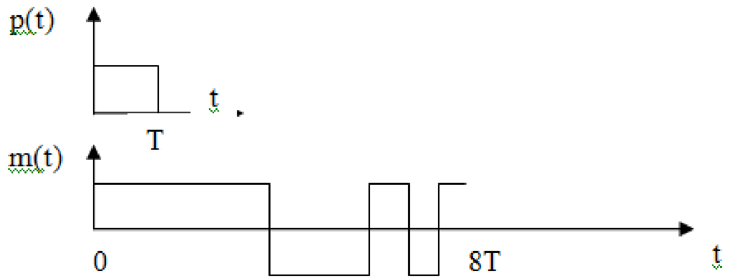


During each frame k at time $t = kT$, we assume $A_i(t) = A_{i,k}, \psi_i(t) = \psi_{i,k}$ are constant, so we can write $m(t = kT) = \sum_i A_{i,k} \cos(2\pi f_i t + \psi_{i,k})$. At the discrete times $t = kT$, the message is the sum of cosine waves with frequencies f_i amplitudes $A_{i,k}$ and phases $\psi_{i,k}$ that are different for each frame k. For each frame k, the amplitudes $A_{i,k}$ and phases $\psi_{i,k}$ at the frequencies f_i can be obtained from the short-time Fourier transform of $m(t)$

An analog message can also represent a digital symbol sequence a_k by writing

$$m(t) = \sum_k a_k p(t - kT)$$

where $p(t)$ is a pulse that spans a finite time period, a_k may be binary symbol ± 1 (to represent binary 1 or 0) or multilevel (e.g. $\pm 1, \pm 3$ to represent 00, 01, 10, 11) and T is the symbol time.



We can have two such sequences

$$m_I(t) = \sum_k a_k p(t - kT) = I(t)$$
$$m_Q(t) = \sum_k b_k p(t - kT) = Q(t)$$

Thus we can write the complex baseband signal as
$I(t) + jQ(t) = \sum_k (a_k + jb_k) p(t - kT),$ where

$c_k = a_k + jb_k = r_k e^{j\phi_k}$
is a complex data symbol. If both a_k, b_k are binary, then c_k represents 2 bits of information at each time $t = kT$ If both a_k, b_k are or multilevel $\pm 1, \pm 3$, then c_k represents 4 bits of information at each time $t = kT$

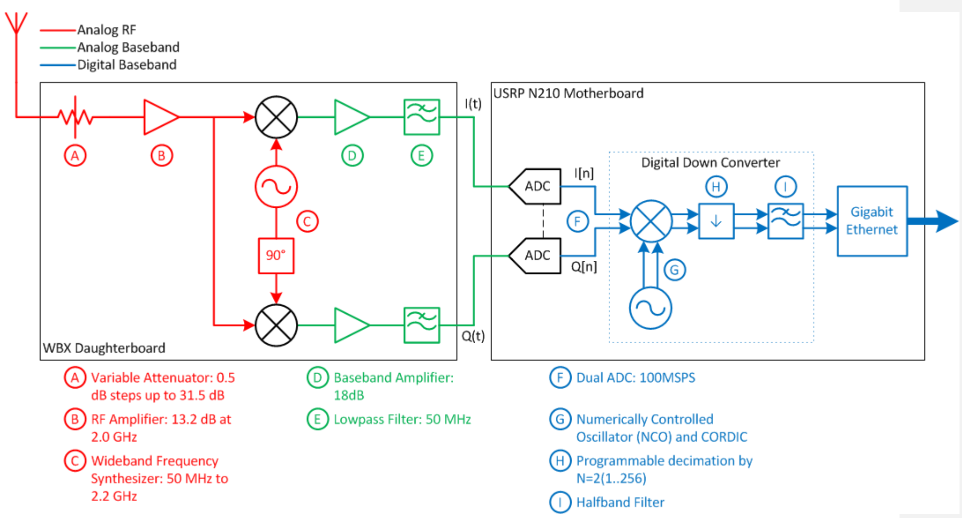
The pulse shape $p(t)$ can have shapes other than rectangular, and can be longer than T. Other pulse shapes are described in a later section.

1.3 Analog I-Q receiver.

1.3.1 Overview and block diagram

In this section, we apply the I-Q signal theory above to a particular example:

> Ettus USRP N210 receiver www.ettus.com used in the ELEC 350 labs (red and green portion of figure below). For this receiver, the daughterboard receives signals with carrier frequency in the range 50 to 2200 MHz and generates analog I-Q outputs that are in the frequency range below 50 MHz and are sampled at 100 MHz.
>
> The red section represents real passband signals, the green section represents the real and imaginary parts of complex baseband signals. The blue section represents digital processing in the software defined part of the radio receiver that will be described later.

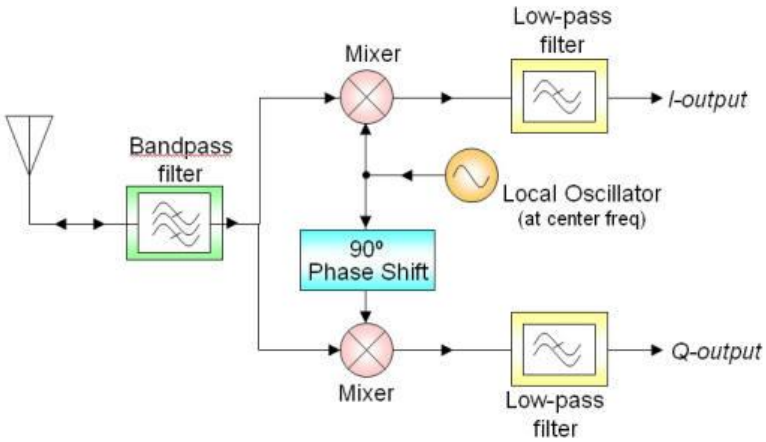


Many software defined radios that operate at frequencies above about 100 MHz have an analog I-Q stage similar to this example ahead of the analog-to-digital converter.

1.3.2 USRP daughterboard receiver operation

The USRP daughterboard is designed to receive radio frequency (RF) signals at frequencies f_c in the range $f_{LO} \pm f_s / 2$ MHz, where f_{LO} is the local oscillator (LO) frequency. In this example, we set f_{LO} to about 100 MHz and f_s is the sampling rate of USRP main board ADCs, set to 100 Msps.

The USRP daughterboard operates by generating two local oscillator (LO) signals at f_{LO} and mixing (multiplying) it with a desired radio frequency (RF) signal at f_c (picked up by the antenna or fed in by a signal generator) to yield a signal at the difference frequency $f_b = f_c - f_{LO}$



The USRP daughterboard has two local oscillators operating 90 degrees out of phase, $\cos 2\pi f_{LO} t$ and $-\sin 2\pi f_{LO} t$ and two mixers. Thus there are two receiver outputs that we call $I(t)$ and $Q(t)$ (see figure above)

The desired RF signal that we wish to receive is written $r(t) = a(t)\cos[2\pi f_c t + \phi(t)]$ We assume $a(t) = 1, \phi(t) = 0$ for the moment, so the desired RF signal is simply an unmodulated carrier wave $r(t) = \cos 2\pi f_c t$.

1.3.3 Receiver operation in real notation

In what follows, given this RF signal input we will calculate the two receiver outputs $I(t)$ and $Q(t)$

To do this, we will use some trigonometric identities

$\cos\alpha\cos\beta = [\cos(\alpha-\beta)+\cos(\alpha+\beta)]/2$

$\sin\alpha\cos\beta = [\sin(\alpha-\beta)+\sin(\alpha+\beta)]/2$

$\sin\alpha\sin\beta = [\cos(\alpha-\beta)-\cos(\alpha+\beta)]/2$

One of the local oscillator signals is written $\cos 2\pi f_{LO} t$.
The *cos* mixer multiplies this LO signal with the RF input signal to obtain

$$\cos 2\pi f_c t \cdot \cos 2\pi f_{LO} t \ .$$

Using one of the trigonometric identities , we can write

$$\cos 2\pi f_{LO} t \cdot \cos 2\pi f_c t = 0.5\cos 2\pi(f_c + f_{LO})t + 0.5\cos 2\pi(f_c - f_{LO})t$$

Thus multiplying two sine (or cosine) waves at frequencies f_{LO} and f_c results in two new sine waves: one at the sum frequency $f_c + f_{LO}$ and one at the difference frequency $f_c - f_{LO}$.

The signal at the sum frequency $f_c + f_{LO}$ is filtered out by analog low pass filters in the USRP daughterboard.

The signal at the difference frequency $f_c - f_{LO}$ is written

$I(t) = 0.5\cos 2\pi(f_c - f_{LO})t = 0.5\cos 2\pi f_b t$.

$I(t)$ can be sampled by the ADC, provided that f_c is close enough to f_{LO}, i.e. the difference is less than half the sampling rate, $|f_c - f_{LO}| < f_s/2$ or
$f_{LO} - f_s/2 < f_c < f_{LO} + f_s/2$

The difference frequency $f_b = f_c - f_{LO}$, where $|f_b| < f_s/2$

The USRP daughterboard has two local oscillators operating 90 degrees out of phase, $\cos 2\pi f_{LO} t$ and $-\sin 2\pi f_{LO} t$ and two mixers. Thus there are two outputs that we call $I(t)$ and $Q(t)$ (see figure above) where $I(t) = 0.5\cos 2\pi(f_c - f_{LO})t$ was calculated above.

The second local oscillator signal is written $-\sin 2\pi f_{LO} t$
The *sin* mixer multiplies this LO signal with the RF input signal to obtain

$\cos 2\pi f_c t \cdot (-\sin 2\pi f_{LO} t)$

We apply a trigonometric identity with $\alpha = 2\pi f_{LO}t, \beta = 2\pi f_c t$ and write

$$-\sin 2\pi f_{LO}t \cdot \cos 2\pi f_c t = -0.5 \sin 2\pi (f_c + f_{LO})t + 0.5 \sin 2\pi (f_c - f_{LO})t$$

The signal at the difference frequency is written

$$Q(t) = 0.5 \sin 2\pi (f_c - f_{LO})t = 0.5 \sin 2\pi f_b t$$

If the receiver outputs $I(t) = \cos 2\pi f_b t$ and $Q(t) = sin 2\pi f_b t$ are displayed on a *x-y* scope, then a circle is displayed. If $f_b < 5$ Hz or so, then the dot on the scope can be seen tracing out the circle. We can write these two receiver output signals $I(t)$ and $Q(t)$ as one complex signal $\tilde{r}(t) = I(t) + jQ(t) = a(t)e^{j\phi(t)}$ that has time varying amplitude and phase $a(t), \phi(t)$, where in this case

$$a(t) = \sqrt{I^2(t) + Q^2(t)} = 1$$

$$\phi(t) = \arctan \frac{Q(t)}{I(t)} = \arctan \frac{\sin 2\pi f_b t}{\cos 2\pi f_b t} = \arctan(\tan 2\pi f_b t) = 2\pi f_b t$$

$$\tilde{r}(t) = a(t)e^{j\phi(t)} = e^{j2\pi f_b t}$$

1.3.4 Receiver operation in complex notation

We can calculate the USRP daughterboard receiver output using complex signals as follows.

Recall that the desired RF signal that we wish to receive is written $a(t)\cos[2\pi f_c t + \phi(t)]$ We assume $a(t) = 1, \phi(t) = 0$ for the moment, so the desired RF signal is simply an unmodulated carrier wave $\cos 2\pi f_c t$. In complex notation we write

$$r(t) = \text{Re}\{r_+(t)\} = \text{Re}\{e^{j2\pi f_c t}\} = cos 2\pi f_c t$$

In complex notation, the unmodulated RF signal $r_+(t) = e^{j2\pi f_c t}$ is multiplied by the complex local oscillator

$$e^{-j2\pi f_{LO}t|} = \cos 2\pi f_{LO}t - j sin 2\pi f_{LO}t$$

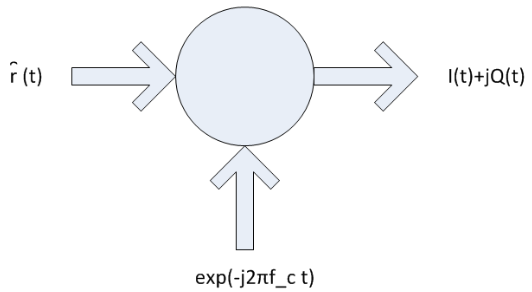
to yield

$$r_+(t)e^{-j2\pi f_{LO}t} = e^{j2\pi f_c t}e^{-j2\pi f_{LO}t} = e^{j2\pi f_b t} = I(t) + jQ(t) = \tilde{r}(t)$$

where the received complex baseband signal is

$\tilde{r}(t) = I(t) + jQ(t) = e^{j2\pi f_b t}$

The diagram below using complex signals performs the same function as the previous diagram above using real signals. Note that the complex signal diagram does not use the low pass filters.



USRP daughterboard analog IQ receiver in complex notation, in this diagram f_c = f_{LO}

In summary,

$r_+(t) = e^{j2\pi f_c t}$
$I(t) = \cos 2\pi f_b t$, $Q(t) = \sin 2\pi f_b t$
$\tilde{r}(t) = I(t) + jQ(t) = \cos 2\pi f_b t + j \sin 2\pi f_b t = e^{j2\pi f_b t}$

the same result as above, apart from a factor 0.5 arising from the complex notation.

If $I(t)$ and $Q(t)$ are displayed on a *x-y* scope, a circle is displayed. If $f_b < 5$ Hz or so, then the dot on the scope can be seen tracing out the circle.

Exercise: repeat the above analysis in both real and complex notation when the RF input signal is a general signal $a(t)\cos[2\pi f_c t + \phi(t)]$ to find $I(t)$ and $Q(t)$ as a function of $f_b, a(t), \phi(t)$.

Exercise: repeat the complex notation analysis for a real input signal $r(t) = \cos 2\pi f_c t$ Hint: write $r(t)$ as the sum of two complex exponentials. Some filters may be needed.

In the next section, we consider IQ signals in the frequency domain.

1.4 Spectrum Analyzers, analog and digital

While an oscilloscope is used to display electrical signals in an amplitude versus time format, a spectrum analyzer provides an amplitude squared (power) versus frequency format. Figure 1 shows the two representations for the same signal.

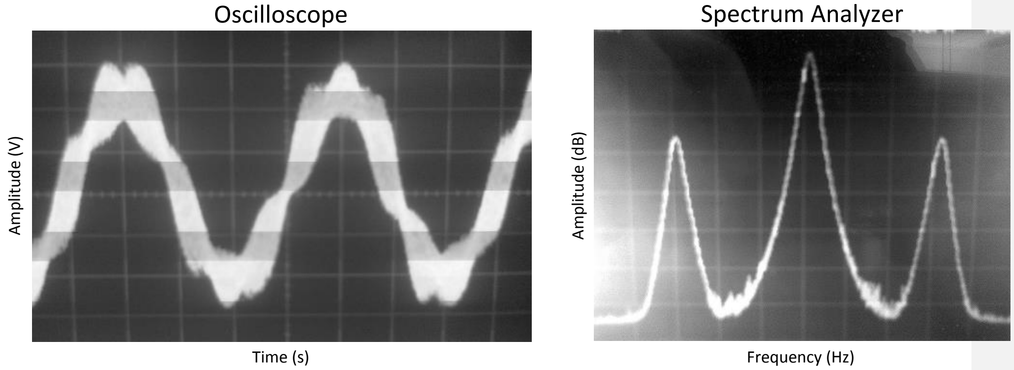


Figure 1 Oscilloscope and Spectrum Analyzer Displays

Fourier analysis shows that any complex waveform can be resolved into sinusoidal waveforms of a fundamental frequency and a number of harmonic frequencies. The spectrum analyzer effectively performs the Fourier integral:

$$S(f) = \int_{t=-\infty}^{\infty} s(t) e^{-j2\pi ft} dt = \int_{t=-\infty}^{\infty} s(t)\cos(2\pi ft)dt - j\int_{t=-\infty}^{\infty} s(t)\sin(2\pi ft)dt \qquad (1.1)$$

The integral finds the frequency components in *s(t)* by correlating *s(t)* with cosine and sine waves at each frequency *f*. For a particular frequency $f = f_c$, $s(t) = \cos(2\pi f_c t)$ and

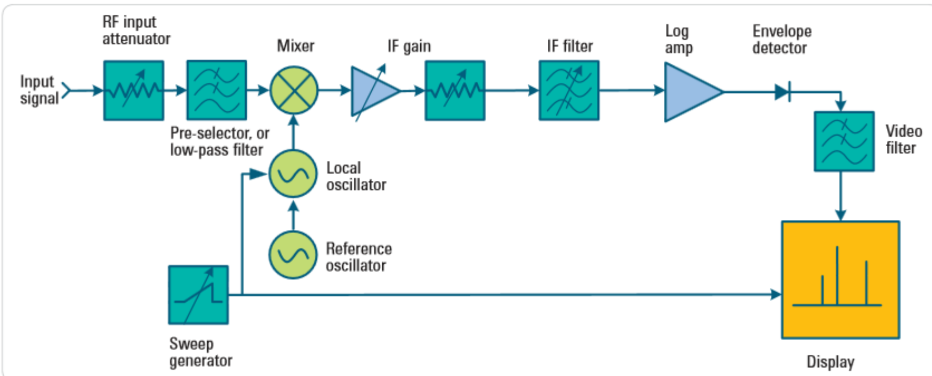
$$S(f_c) = \int_{-\infty}^{\infty} \cos(2\pi f_c t)\cos(2\pi f_c t)dt = \int_{-\infty}^{\infty} 0.5[1+\cos(4\pi f_c t)]dt$$

The cosine waves are in phase for all time, the product of the two cosine waves contains DC, thus the integral integrates DC over all time, resulting in infinity (delta functions) at $f = f_c$. For all other frequencies $f \neq f_c$ the cosine waves drift in and out of phase over time, there is no DC component, and the integral is zero. Thus for $s(t) = \cos(2\pi f_c t)$ we can write:

$$S(f) = \int_{-\infty}^{\infty} \cos(2\pi f_c t)\cos(2\pi ft)dt = [\delta(f - f_c) + \delta(f + f_c)]/2 \qquad (1.2)$$

For a digital spectrum analyzer where the signal is digitized (sampled) by an analog-to-digital converter (ADC), the Fourier transform is done using the Fast Fourier Transform (FFT) algorithm, as will be discussed in more detail in the next section.

There is a practical upper limit to the sampling rate of an ADC. Thus analog spectrum analyzers are used for high frequencies. For an analog spectrum analyzer, frequency scanning is accomplished by electronically tuning a bandpass filter network across the desired frequency range. In practice, a sweep generator along with a fixed frequency bandpass filter, rather than tuning the bandpass filter, as shown in the figure taken from an Agilent app note. http://cp.literature.agilent.com/litweb/pdf/5952-0292.pdf



The amplitudes of all the signals in the bandpass area (at each point during the scan) are measured to provide the amplitude versus frequency display.

It is important to note that in order to detect two signals which are narrowly spaced, the frequency span of the bandpass filter must be set appropriately. This concept is shown in the figure below: the frequencies f_1, f_2 and f_3 in the input spectrum are summed into a single peak on the displayed waveform due to the excessive frequency span. Note that the DC reference shown in this figure is a characteristic of the spectrum analyzer. It is always present, regardless of whether there is a DC bias on the input signal.

Unlike an oscilloscope, the amplitude scale on a spectrum analyzer can be set to a logarithmic scale (for power measurements) as well as a linear scale (for voltage measurements). The logarithmic scale is represented in dBm, or decibels with reference to a milliwatt.

$$\text{Power}_{\text{(dBm)}} = 10 \log \left(\frac{input \quad power}{1mW} \right) \tag{1.3}$$

It can then be shown that

$$\text{Power}_{\text{(dBm)}} = 10 \log \left(\frac{V^2}{Z} \cdot \frac{1}{1mW} \right) \tag{1.4}$$

If the load is a standard 50Ω input

$$\text{Power}_{\text{(dBm)}} = 20 \log \text{V} + 10 \log \frac{1}{50 \times 10^{-3}} = 20 \log \text{V} + 13 \tag{1.5}$$

So the relationship between input power and input voltage is

$$P_{(\text{dBm})} = 20 \log V + 13 \tag{1.6}$$

This is a handy formula worth remembering.

On the logarithmic dB scale, the power of signals is always positive. It is the magnitude of the signals above (in reference to) the noise level that is of interest.
The analog spectrum analyzer shows only the square magnitude (or power) of each Fourier component.

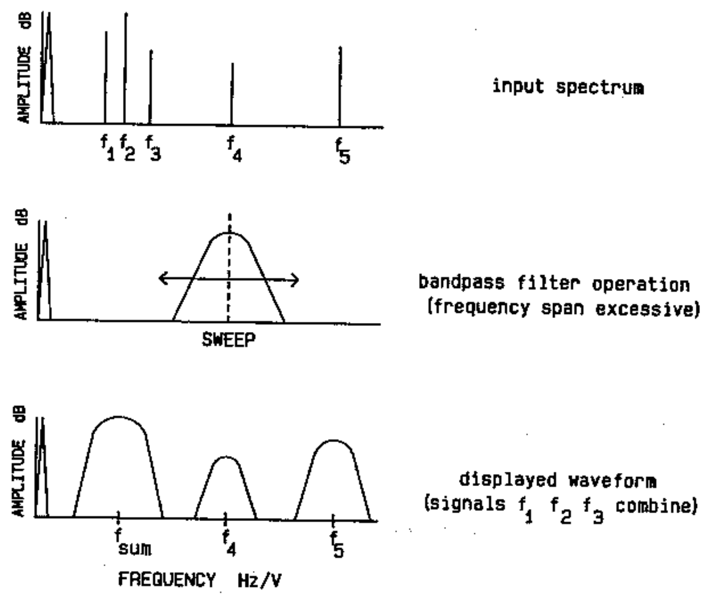


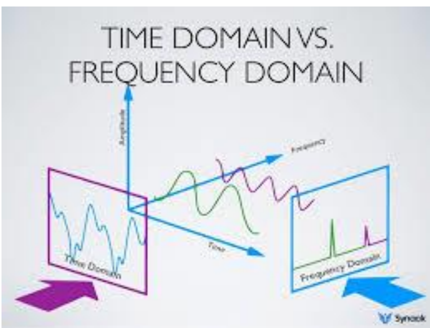
Figure 2 Effect of Span on Spectrum Analyzer Display

Most analog spectrum analyzers have a peak search function to identify the frequency and power (in dBm) of a peak.

A digital spectrum analyzer can show phase as well as amplitude. The block diagram of the digital spectrum analyzer is the same as the general I-Q receiver, followed by an ADC (analog-to-digital converter), FFT processor and display.

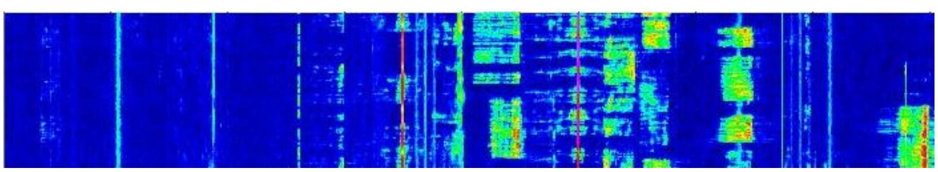
1.5 Spectrogram or spectral waterfall

Radio signals (and other types of one dimensional signals) can be displayed with a *spectrogram* that represents the signal simultaneously in the time and frequency domains. A spectrogram has 3 dimensions: time, frequency and amplitude.



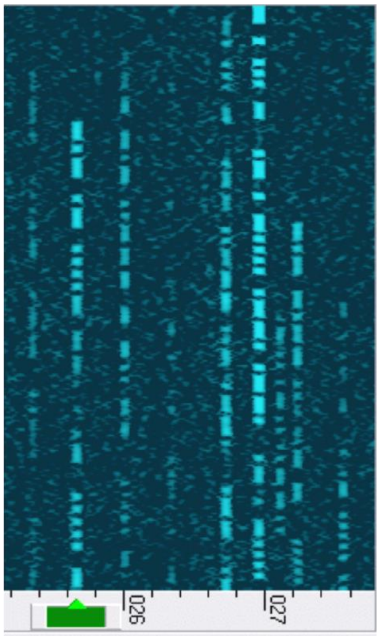
Often the amplitude is represented by colour, just like a topographical map. A *waterfall* display is a spectrogram where the time axis is moving, showing the most recent signal first.

The figure below shows an example snapshot in time of a spectrogram or moving spectral waterfall display. The *x* axis is frequency, the *y* axis is time (most recent at the top, time increasing downwards).
On the left of the display, the blue line represents a sine wave at a fixed frequency. The green blocks represent radio signal modulated with information, and thus occupy a finite non-zero bandwidth. Note that the radio signals on a given frequency start and stop at different times.



Waterfall spectrogram with frequency on *x* axis, time on *y* axis, amplitude color coded to display *z* axis.

Another example of a waterfall spectrogram is shown below. In this example, all the signals are morse code, i.e. sine waves switched on and off. The amplitude is grey scale coded to display z axis, a tunable filter shown in green to select desired signal.

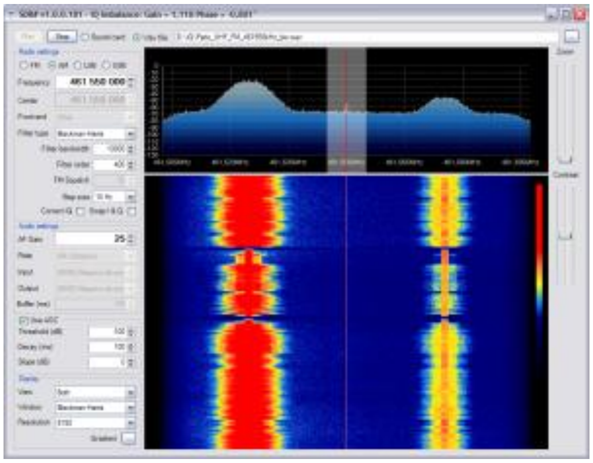


An example of a combination spectrum and spectral waterfall is shown below. The spectrum is the output of a digital spectrum analyzer and is valid in real time. The waterfall shows the history of the spectrum with frequency on the horizontal axis, time on the vertical axis and amplitude represented by color.

The waterfall is drawn as follows:

1. Segment the time domain signal $\tilde{s}(t)$ sampled at a given rate f_s samples per second into frames of a fixed number of samples *N*, where *N* is typically a power of two such as 1024. Each frame is of length $T_0 = N / f_s$ seconds.
2. Take the short-time Fourier transform of a frame $\tilde{S}(f) = FFT\{\tilde{s}(t)\}$ to yield a frame of *N* complex samples representing the amplitude and phase versus frequency. The frequency resolution will be $f_0 = f_s / N$ Hz.
3. Draw one horizontal line at the top of the spectrogram with the color of each sample chosen to represent the amplitude. For example, blue for low amplitude and red for high amplitude and all colors in between.
4. Push all other lines in the spectrogram down by one time unit (one frame time).
5. Repeat from Step 2 for the next segment (frame).

In the example figure, there are two signals present that are present continuously.



The spectrum for a particular frame is found by computing $\tilde{S}(f) = FFT\{\tilde{s}(t)\}$ for the samples of $\tilde{s}(t)$ in that frame. The spectrum extends from $-f_s/2$ to $f_s/2$. Since $\tilde{s}(t)$ is complex, the spectrum $\tilde{S}(f)$ is not symmetrical around zero, thus the spectrum will have a different shape for positive and negative frequencies.

The frequency axis zero point is often labeled f_{LO} in the waterfall plot instead of zero, since an RF signal at exactly $f_c = f_{LO}$ will be mixed by the IQ receiver down to $f_c - f_{LO} = f_b = 0$ Hz exactly. In this way, the frequency axis displays signals at their actual RF frequencies within a bandwidth f_s from $f_{LO} - f_s/2$ to $f_{LO} + f_s/2$. Thus the maximum bandwidth of a complex signal sampled at f_s is f_s,

For a real signal sampled at f_s the bandwidth is $f_s/2$ and the spectrum is symmetrical about zero frequency. The spectrum shape at negative frequencies from $-f_s/2$ to 0 will be the mirror image of the spectrum shape from 0 to $f_s/2$.

Live waterfall displays with tunable filter and audio output can be found at http://www.websdr.org/

Exercise: write the algorithm for creating a spectrogram waterfall showing the signal samples of $\tilde{s}(t)$ and $\tilde{S}(f) = FFT\{\tilde{s}(t)\}$ in discrete-time notation where $\tilde{s}_n = \tilde{s}(t = nT_s),\ T_s = 1/f_s,\ \tilde{S}_k = \tilde{S}(f = kf_0)$. Recall the discrete Fourier transform is defined as a Fourier transform operating on a sampled periodic signal

$$\tilde{S}_k = \int_{t=0}^{T_0 = nT_s} \tilde{s}(t) e^{-j2\pi ft} dt \,\Big|_{t=nT_s,\, f=kf_0} = \sum_{n=0}^{N-1} \tilde{s}_n e^{-j2\pi k f_0 n T_s} = \sum_{n=0}^{N-1} \tilde{s}_n e^{-j2\pi nk/N} \quad \text{where } T_s f_0 = 1/N$$

1.6 IQ imbalance

In practice, the two local oscillators of the USRP daughterboard analog IQ receiver are such that $Q(t)$ is not exactly the same amplitude as $I(t)$ and not exactly 90 degrees out of phase with $I(t)$. This is called IQ imbalance. The practical result is that in the spectrum analyer or waterfall view, the signals at positive frequencies will have weak images at the corresponding negative frequencies, and vica versa. These images are not really there, they are an artifact of the IQ imbalance.

If there is IQ imbalance present, then the amplitude and phase of $I(t)$ is shifted relative to what it should be by a complex factor $(1+\alpha)e^{j\theta}$ for small values of α,θ. We write the complex output $\tilde{r}'(t)$ of the imperfect receiver with IQ imbalance as

$\tilde{r}'(t) = I'(t) + jQ'(t) = (1+\alpha)e^{j\theta}I(t) + jQ(t)$
$I'(t) = (1+\alpha)cos\theta I(t)$
$Q'(t) = Q(t) + (1+\alpha)\sin\theta I(t)$

For small theta, this matches the approximation

- Re{out} = Re{in} * (1 + Magnitude)
- Im{out} = Im{in} + Re{in} * Phase

Where we define

$\mathrm{Re}\{out\} = I'(t)$

$\mathrm{Re}\{in\} = I(t)$

$\mathrm{Im}\{out\} = Q'(t)$

$\mathrm{Im}\{in\} = Q(t)$

$magnitude = \alpha$

$phase = \theta$

In the frequency domain, IQ imbalance appears in the form of an image, i.e. every signal at frequency f_b will have a mirror image at $-f_b$. For a signal at f_b, the desired complex baseband signal is $\tilde{r}(t) = \tilde{s}(t) = e^{j2\pi f_b t}$, but with IQ imbalance, the actual observed complex baseband signal is

$$\tilde{r}'(t) = \mu e^{j2\pi f_b t} + \nu e^{-j2\pi f_b t}$$

with complex constants μ,ν, thus explicitly showing the desired signal at f_b and the mirror image at $-f_b$. The complex constants μ,ν are functions of α,θ. When $\alpha = \theta = 0, \nu = 0$

In general, we can write the received complex signal obtained at the output of an imperfect IQ receiver as

$$\tilde{r}'(t) = \mu\tilde{r}(t) + \nu\tilde{r}^*(t)$$

where $\tilde{r}'(t)$ is the receiver output, $\tilde{r}(t) = \tilde{s}(t)$ is the desired complex signal, and $\tilde{r}^*(t)$ is the image The image rejection is $20\log|\nu/\mu|$ dB.

The mirror image can be nulled out by multiplying $Q'(t)$ by the inverse of this complex factor $(1+\alpha)e^{j\theta}$, i.e. $\frac{1}{1+\alpha}e^{-j\theta} = \beta e^{j\psi}$. The values of the magnitude β and phase ψ can be found manually (with hardware or software controls) or by an adaptive algorithm.

Exercises

1. For an complex baseband signal $\tilde{s}(t) = I(t) + jQ(t) = a(t)e^{j\phi(t)}$, find an expression for $a(t)$ as a function of $I(t), Q(t)$
2. Given a USRP daughterboard receiver with $f_{LO} = 142$ MHz and an incoming RF signal $f_c = 142.17$ MHz, find an expression for
   a. the two real output signals from the receiver $I(t), Q(t)$, and
   b. the complex output signal $I(t) + jQ(t)$ written in polar form $a(t)e^{j\phi(t)}$
3. Verify the expressions for the imperfect receiver outputs $I'(t)$ and $Q'(t)$ given in section 4 above.
4. Verify the approximate expressions for IQ imbalance correction given in Section 4 above.

1.7 Radio tuning, selecting a particular signal

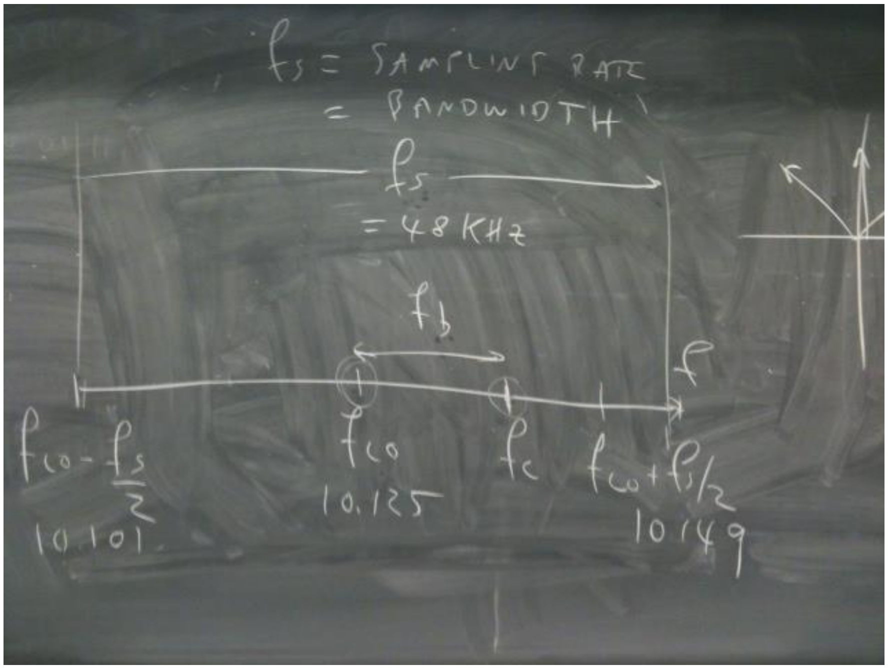
The USRP receiver is designed to receive radio frequency (RF) signals at any frequency f_c in the range $f_{LO} \pm f_s/2$ MHz, where f_{LO} is the local oscillator (LO) frequency. In this example, consider the complex signal at the output of the USRP (that is connected to the computer via Ethernet and processed in software such as GNURadio). The GNURadio block diagram software includes a so-called "USRP source block" that has a complex signal output and has the LO frequency and sampling rate as parameters. This complex signal can be live from the USRP source block output or a IQ file in 2 channel wav format that was recorded previously.
Note that this LO frequency in the GNURadio software is not the same as the LO frequency in the USRP daughterboard mentioned in an earlier section.

In this example, we choose f_{LO} = 10.125 MHz and $f_s = 48$ KHz, so that the frequency range on the spectrum or waterfall is 10.101 – 10.149 MHz. Selecting a signal at a particular frequency f_c is called "tuning" the radio.

The USRP source block operates by generating two local oscillator signals at f_{LO} and mixing (multiplying) it with a desired radio frequency (RF) carrier wave $r_+(t) = e^{j2\pi f_c t}$ at f_c to yield a complex baseband signal $\tilde{r}(t) = I(t) + jQ(t)$ at the difference frequency $f_b = f_c - f_{LO}$, where we write

$$r_+(t)e^{-j2\pi f_{LO}t} = e^{j2\pi f_c t}e^{-j2\pi f_{LO}t} = e^{j2\pi f_b t} = \cos 2\pi f_b t + jsin2\pi f_b t = I(t) + jQ(t)$$

$I(t)$ and $Q(t)$ is processed correctly, provided that f_c is close enough to f_{LO}, i.e. the difference is less than half the sampling rate, $|f_c - f_{LO}| < f_s / 2$ or $f_{LO} - f_s / 2 < f_c < f_{LO} + f_s / 2$. The difference frequency $f_b = f_c - f_{LO}$, where $|f_b| < f_s / 2$



The USRP source block function is to shift a 48 KHz wide slice of spectrum from 10.101-10.149 MHz centered at $f_{LO} = 10.125$ MHz down to -0.024 to +0.024 MHz (positive and negative frequencies centered around zero Hz). The complex baseband signal $\tilde{r}_1(t) = e^{j2\pi f_b t}$ can represent positive and negative frequencies, since f_b can be positive or negative and $| f_b |< 24$ KHz.. The 48 KHz slice of spectrum may contain many different signals at various frequencies within the 48 KHz range (recall waterfall plots mentioned above).

If the RF carrier wave is turned on and off to transmit information in e.g. Morse code, then we wish to listen to (or digitally decode) the complex baseband signal $\tilde{r}_1(t) = \cos 2\pi f_b t + \text{j} \sin 2\pi f_b t$ and no other signals.

If f_b is outside the audio range we want to listen to, or if f_b is not the frequency expected at the digital decoder input, then we multiply $\tilde{r}_1(t)$ by a complex exponential $e^{-j2\pi f_d t}$ at frequency f_d to obtain another complex baseband signal

$$\tilde{r}_2(t) = e^{j2\pi f_b t} e^{-2\pi f_d t} = e^{j2\pi(f_b - f_d)t} = e^{2\pi f_E t}$$

at frequency $f_E = f_b - f_d$, where f_E is chosen to be the frequency for listening (or for the decoder).

In effect, we have shifted the spectrum twice. We first shifted f_c by f_{LO} to obtain f_b and then shifted f_b by f_d to get the exact frequency f_E we want to listen to (for Morse code) or for a digital decoder.
This idea of two successive spectrum shifts will be seen again later when we discuss the Weaver demodulator.

Table of frequencies used above

f_s sampling rate of USRP source block
f_{LO} local oscillator frequency
f_c carrier frequency of desired signal at radio frequency near 10 MHz (passband)
$(f_{LO} - f_s / 2) < f_c < (f_{LO} + f_s / 2)$ passband frequency range
$f_b = f_c - f_{LO}$ desired signal obtained by converting to baseband
$(-f_s / 2) < f_b < (f_s / 2)$ baseband frequency range (centered at 0 Hz)
Conversion (spectrum shifting) is done by complex multiply
$\tilde{r}_1(t) = r_+(t) e^{-j2\pi f_{LO} t} = e^{j2\pi f_c t} e^{-j2\pi f_{LO} t} = e^{j2\pi f_b t} = \cos 2\pi f_b t + jsin 2\pi f_b t = I(t) + jQ(t)$
f_E desired frequency for listening or for decoder (could be zero or not zero)
$f_d = f_b - f_E$ digital local oscillator in baseband used for tuning into desired signal
f_d is controlled by tuning knob or mouse
Spectrum shift is done by complex multiply
$\tilde{r}_2(t) = e^{j2\pi f_b t} e^{-2\pi f_d t} = e^{j2\pi(f_b - f_d)t} = e^{2\pi f_E t}$

More generally, if the RF signal contains information encoded in its amplitude and phase, then the RF signal $r_+(t) = a(t)e^{j\phi(t)}e^{j2\pi f_c t}$ is multiplied by the complex local oscillator $e^{-j2\pi f_{LO}t|} = \cos 2\pi f_{LO}t - jsin2\pi f_{LO}t$
to yield

$$\tilde{r}_1(t) = r_+(t)e^{-j2\pi f_{LO}t} = [a(t)e^{j\phi(t)}e^{j2\pi f_c t}]e^{-j2\pi f_{LO}t} = a(t)e^{j\phi(t)}e^{j2\pi f_b t} = I(t) + jQ(t)$$

where the received complex baseband signal is

$$\tilde{r}_1(t) = I(t) + jQ(t) = a(t)\cos\phi(t)cos2\pi f_b t + \mathrm{j}\; a(t)\sin\phi(t)sin2\pi f_b t$$

If we want to receive the information contained in $a(t), \phi(t)$ then we multiply $\tilde{r}_1(t)$ by a complex exponential $e^{-j2\pi f_b t}$ at exactly $-f_b$ to obtain

$\tilde{r}_2(t) = \tilde{r}_1(t)e^{-j2\pi f_b t} = a(t)e^{j\phi(t)}e^{-j2\pi f_b t} = a(t)e^{j\phi(t)}$ centered at $f_E = 0$, followed by a low pass filter to filter out any other signals. We have shifted the spectrum twice, once by f_{LO} using analog circuits and a second time by f_b using software to receive the desired signal. In this case, the decoder uses $f_E = 0$

Exercise:

Explain the operation of a transmitter that has a selectable carrier frequency. Sketch a block diagram and the signal spectrum. The transmitter uses the same frequency shifting operations in reverse order.

2 Amplitude modulation (AM) transmitters

So how do we generate a radio signal $s(t)$ using our desired carrier $c(t)$ and the message $m(t)$?

We wrote a general radio signal $s(t) = a(t)\cos(2\pi f_c t + \phi(t))$. We can use $a(t)$ and/or $\phi(t)$ or some combination thereof to represent the message $m(t)$.

In this chapter, we introduce several types of modulation that use $a(t)$ only.

2.1 Amplitude modulation with full carrier

One method to represent the message is to use the message $m(t)$ to modulate the amplitude $a(t)$ and leave the phase $\phi(t)$ constant. This method is called *amplitude modulation (AM).*

AM waves are usually carried at frequencies on the order of MHz, whereas the message frequency is typically on the order of KHz.

Amplitude modulation is used for broadcasting (AM radio), using carrier frequencies both in the medium wave band (540-1700 KHz), the long wave band 153-279 KHz (in Europe only) and the short wave bands (3-30 MHz). AM is also used for aircraft navigational aids (190-535 KHz and 108-118 MHz) and aircraft voice communications (118-137 MHz).

For AM, we set the amplitude

$$a(t) = A_c[1 + k_a m(t)]$$

with $|k_a m(t)| < 1$ or $[1 + k_a m(t)] > 0$ and $f_c >> w$
where w is the bandwidth of *m(t).*

We have chosen the modulation index k_a such that $a(t) > 0$

For AM, we assume $\varphi(t)$ is a constant or zero.

Thus the transmitted AM signal is of the form:

$$s(t) = a(t)\cos(2\pi f_c t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t); \qquad \varphi(t) = 0$$

For the special case where there is no message to send, $m(t) = 0$ and $s(t) = A_c \cos(2\pi f_c t)$ which is simply the carrier wave.

In practice, $m(t)$ could an analog or digital message as described in an earlier section.

For an analog message, $m(t) = A_m \cos(2\pi f_m t)$, the resulting AM wave is given by the following:

$$s(t) = A_c[1 + k_a A_m \cos(2\pi f_m t)]\cos(2\pi f_c t)$$

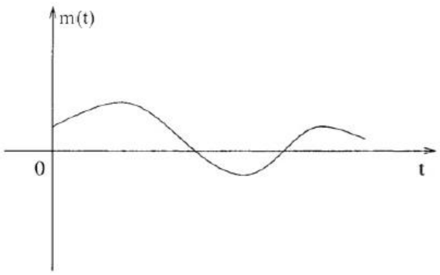
Recall that an analog message (voice or music) is time varying in the practical case and will normally change from one frame to the next. (Recall that frames for audio are typically 5-23 msec). The AM signal above with $m(t) = A_m \cos(2\pi f_m t)$ is valid for one frame in which the message contains only one frequency f_m during that frame. In the next frame, the message could be the same or could be different.

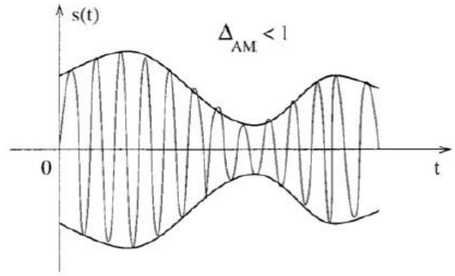
2.1.1 AM waveform with single tone message

For our purposes here, we will assume the message is of constant frequency for all frames. The AM wave appears in the time domain to be the carrier wave with an envelope that replicates the form of the modulating tone.

$$m(t) = A_m \cos 2\pi f_m t \qquad f_m = \text{message frequency}$$
$$a(t) = A_c(1 + k_a m(t)], \phi(t) = 0$$
$$s(t) = A_c(1 + k_a m(t)]\cos 2\pi f_c t$$
$$s(t) = A_c(1 + k_a A_m \cos 2\pi f_m t]\cos 2\pi f_c t$$
$$s(t) = A_c(1 + \mu \cos 2\pi f_m t]\cos 2\pi f_c t$$
$$s(t) = A_c \cos 2\pi f_c t + A_c \mu \cos 2\pi f_m t \cos 2\pi f_c t$$
$$\mu = k_a A_m < 1 \text{ is called modulation index}$$

We use the character $\mu = A_m k_a$ for simplification and to serve as the modulation index. The modulation index μ can be given as a percentage where 1 = 100%.





Using the complex I-Q notation

$$s(t) = a(t)\cos(2\pi f_c t + \phi(t)) = \mathrm{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\}$$

Thus for AM

$$s(t) = \mathrm{Re}\{A_c[1 + k_a m(t)]e^{j2\pi f_c t}\} = \mathrm{Re}\{A_c[1 + \mu cos 2\pi f_m t]e^{j2\pi f_c t}\}$$
$$a(t) = A_c[1 + \mu cos 2\pi f_m t]$$
$$\phi(t) = 0$$

2.1.2 AM spectrum

What does the AM wave look like in the frequency domain?
We first consider the special case where the message $m(t) = A_m \cos(2\pi f_m t)$, and take the Fourier transform

$$m(t) \xrightarrow{Fourier} M(f)$$
$$A_m \cos(2\pi f_m t) \xrightarrow{Fourier} \frac{1}{2} A_m[\partial(f - f_m) + \partial(f + f_m)]$$

This result is obtained from the Fourier transform properties for a cos wave.

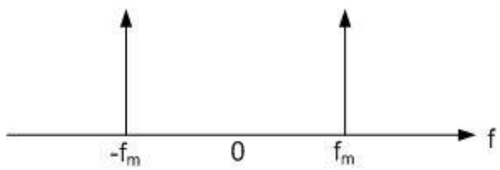
$$m(t) \leftrightarrow M(f)$$
$$m(t) = \cos(2\pi f_m t) = (e^{j2\pi f_m t} + e^{-j2\pi f_m t})/2$$
$e^{j2\pi f_m t} \leftrightarrow \delta(f - f_m)$ positive frequencies

$e^{-j2\pi f_m t} \leftrightarrow \delta(f + f_m)$ negative frequencies
$$M(f) = \frac{1}{2}[\delta(f - f_m) + \delta(f + f_m)]$$

Note that the cos wave contains both positive and negative frequencies arising from the exponentials rotating clockwise and counterclockwise respectively as time increases.

The spectrum for any real signal is symmetrical about zero frequency.

The spectrum of the message $M(f)$ can be drawn as shown in the figure, where both positive and negative frequencies are shown:



Frequency Spectrum of a Single Tone

Note that the delta functions are drawn as vertical lines with arrows that suggest they go to infinity. The delta function $\delta(f - f_m)$ is zero for all frequencies $f \neq f_m$. For the frequency $f = f_m$, the delta function $\delta(f - f_m)$ has infinite height, zero width and area 1, so that we can write the area under the curve $\int_{-\infty}^{\infty} \delta(f - f_m) = 1$

The spectrum of the carrier wave is written

$c(t) \leftrightarrow C(f)$

$c(t) = \cos(2\pi f_c t) = (e^{j2\pi f_c t} + e^{-j2\pi f_c t})/2$

$e^{j2\pi f_c t} \leftrightarrow \delta(f - f_c)$ positive frequencies

$e^{-j2\pi f_c t} \leftrightarrow \delta(f + f_c)$ negative frequencies

$C(f) = \frac{1}{2}[\delta(f - f_c) + \delta(f + f_c)]$

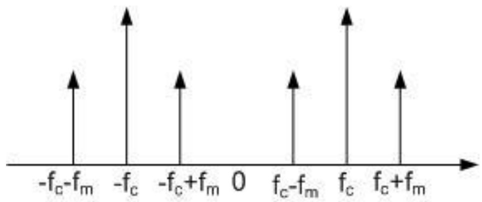
$C(f)$ appears as two delta function spikes, one at f_c and one at $-f_c$

To find the spectrum of the AM signal

$$\begin{aligned} s(t) &= A_c(1 + \mu \cos 2\pi f_m t]\cos 2\pi f_c t \\ &= A_c \cos 2\pi f_c t + A_c \mu \cos 2\pi f_m t \cos 2\pi f_c t \end{aligned}$$

we take the Fourier transform

$s(t) \xrightarrow{Fourier} S(f)$ and obtain the result shown in the figure below.



Frequency Spectrum of AM Wave

Observe that the spectrum of an AM signal contains both positive and negative frequencies. The AM signal spectrum is the sum of two components: the spectrum of the carrier wave at $\pm f_c$ plus the spectrum of the message shifted both up and down by f_c

There is very important principle at work here. Notice that in the expression

$$s(t) = A_c \cos 2\pi f_c t + A_c \mu \cos 2\pi f_m t \cos 2\pi f_c t$$

the message signal is multiplied by the carrier wave. Whenever a message signal is multiplied by a carrier wave, the spectrum of the message signal is shifted both up and down by the carrier frequency.

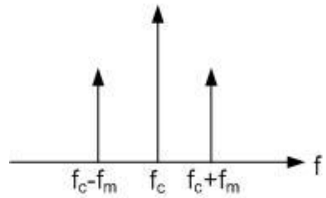
This result for the AM signal spectrum can be demonstrated in two different ways.

One method is done algebraically by simplifying $s(t)$ in the time domain and then converting the resulting collection of individual sinusoidal terms to their frequency domain representation.

Note: The trigonometric identity $\cos\alpha\cos\beta = \frac{1}{2}[\cos(\alpha+\beta)+\cos(\alpha-\beta)]$ is used to produce line three from line two below:

$$s(t) = A_c[1+\mu\cos(2\pi f_m t)]\cos(2\pi f_c t)$$
$$s(t) = A_c\cos(2\pi f_c t) + A_c\mu\cos(2\pi f_m t)\cos(2\pi f_c t)$$
$$s(t) = A_c\cos(2\pi f_c t) + (A_c\mu/2)\cos(2\pi[f_c + f_m]t) + (A_c\mu/2)\cos(2\pi[f_c - f_m]t)$$

The second and third cosine terms represent the sidebands as seen in the frequency domain. The frequency of the sidebands relative to the carrier frequency holds the useful information describing the message.



Spectral Representation of AM signal showing positive frequencies only

$$S(f) = \frac{A_c}{2}[\delta(f - f_c) + \delta(f + f_c)]$$
$$+\frac{A_c\mu}{4}[\delta(f - f_c - f_m) + \delta(f + f_c + f_m)]$$
$$+\frac{A_c\mu}{4}(\delta(f - f_c + f_m) + \delta(f + f_c - f_m))]$$

Exercise: show that

$$\mu = \frac{A_{cMAX} - A_{cMIN}}{A_{cMAX} + A_{cMIN}} = A_m k_a, \text{ where}$$

A_{cMAX} and A_{cMIN} are respectively the highest and lowest positive amplitude obtained by $s(t)$.

Hint: $A_{cMAX} = A_c(1+\mu)$; $A_{cMIN} = A_c(1-\mu)$, $\frac{A_{c\max}}{A_{c\min}} = \frac{A_c(1+\mu)}{A_c(1-\mu)}$

Observe that even though we seem to only modify the amplitude of the carrier wave at f_c sidebands are observed at $f_c + f_m$ and $f_c - f_m$ in the frequency domain representation. How could there by frequencies at $f_c + f_m$ and $f_c - f_m$ when all we are doing is varying the amplitude of the carrier at f_c without changing its frequency?

The reason is that to see the sidebands at $f_c + f_m$ and $f_c - f_m$ we must observe the signal over a sufficiently long time (at least one cycle of the message *m(t)*). If we observe the signal for a very short time, i.e. over a small fraction of a cycle of *m(t)*, then we see only the carrier wave with amplitude $a(t) = A_c[1+\mu cos 2\pi f_m t]$ and we don't see any sidebands.

The second method of showing the result is done using Fourier transform properties (*exercise*). Note that a Fourier transform is taken over a long time including at least one cycle of the message $m(t)$, so that we expect to see the sidebands.

The result can also be shown using a Fourier series, where the signal is assumed to be periodic with a period equal to one cycle of the message.

2.1.3 Power in carrier wave and sidebands

Power is a quantity of interest and one may wish to calculate the power of the carrier or sidebands specifically. The power is proportional to the square of the cosine coefficients.

The power versus frequency curve is displayed on a spectrum analyzer, typically in dBm, or dB relative to one milliwatt at 50 ohms. For this case, P(dBm) = 20 log V + 13 where V is rms voltage.

In what follows we assume 1 ohm rather than 50 ohms for convenience.
The power into one ohm $[s(t)]^2 = |S(f)|^2$

Consider an AM signal with single tone modulation $m(t) = A_m \cos 2\pi f_m t$

$$s(t) = A_c[1+\mu\cos(2\pi f_m t)]\cos(2\pi f_c t)$$

which, in the frequency domain is:

$$s(t) = A_c \cos(2\pi f_c t) + (A_c\mu/2)\cos(2\pi[f_c + f_m t)]t) + (A_c\mu/2)\cos(2\pi[f_c - f_m]t)$$

Power at carrier at f_c: $A_c^2/2$
Power of sideband at $f_c + f_m$: $(1/2)(A_c\mu/2)^2 = (1/8)A_c^2\mu^2$
Power of sideband at $f_c - f_m$: $(1/2)(A_c\mu/2)^2 = (1/8)A_c^2\mu^2$

carrier power $2\left(\frac{A_c}{2}\right)^2 = \frac{A_c^2}{2}$
factor of 2 because positive and negative frequencies
upper sideband power $2\left(\frac{A_c\mu}{4}\right)^2 = \frac{A_c^2\mu^2}{8}$
lower sideband power $2\left(\frac{A_c\mu}{4}\right)^2 = \frac{A_c^2\mu^2}{8}$

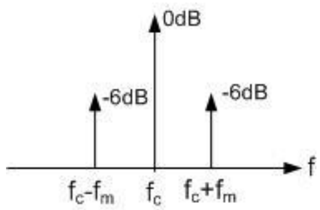
$$\frac{USB + LSB\,Power}{Total\,Power} = \frac{\frac{A_c^2\mu^2}{8} + \frac{A_c^2\mu^2}{8}}{\frac{A_c^2}{2} + \frac{A_c^2\mu^2}{8} + \frac{A_c^2\mu^2}{8}} = \frac{\mu^2}{2+\mu^2}$$

Question: If the power at $f_c = 0$ dBm, what is the power at $f_c + f_m$?

Answer: Given: $A_c^2/2 = 1mW$; $(1/8)A_c^2\mu^2 = (1mW)\mu^2/4$

If $\mu = 1$ then the power at $f_c + f_m$ is $(1mW)/4 = 0.25mW$

A drop in power by half is -3dB, thus one quarter of the power is -6dB

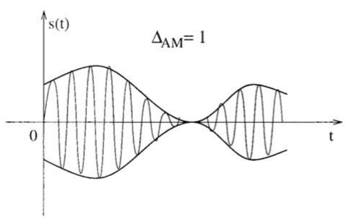


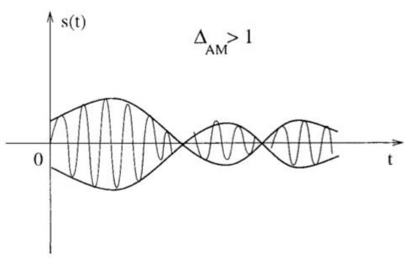
AM Spectrum: Sidebands with modulation index 1

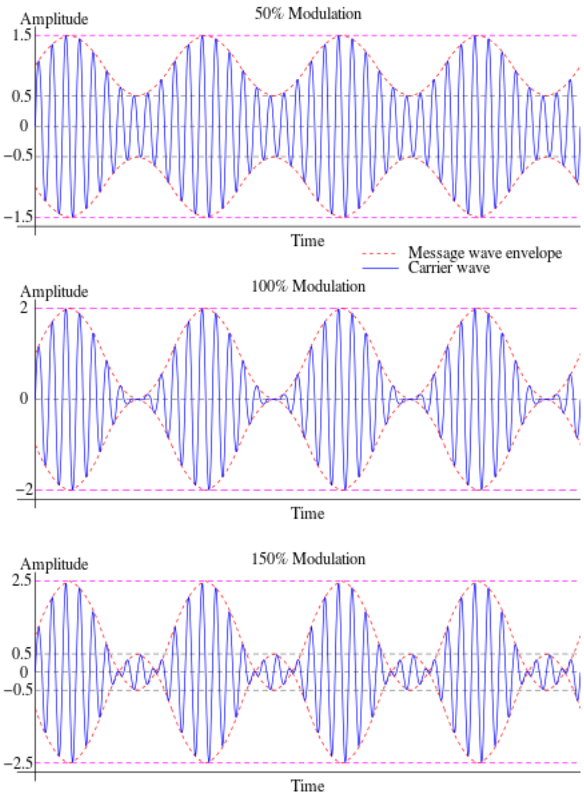
In the special case where $\mu = 1$, we have 100% modulation (minimum envelope value is zero).

2.1.4 Overmodulation

At modulation $\mu > 1$ (greater than 100%) the envelope, $[1 + \mu\cos(2\pi f_m t)]$ becomes less than zero and no longer looks like the message being sent. This case is called "overmodulation". The phase of the carrier wave is shifted by 180 degrees when $[1 + \mu\cos(2\pi f_m t)]$ is less than zero.







2.1.5 AM waveform with a general message

We can write the AM signal in time and frequency domain for a general message *m(t)*

$$s(t) = A_c[1 + k_a m(t)]\cos 2\pi f_c t$$
$$s(t) = A_c[1 + k_a m(t)]\frac{(e^{j2\pi f_c t} + e^{-j2\pi f_c t})}{2}$$
$$s(t) = \frac{A_c}{2}e^{j2\pi f_c t} + \frac{A_c}{2}e^{-j2\pi f_c t} + \frac{A_c k_a}{2}m(t)e^{j2\pi f_c t} + \frac{A_c k_a}{2}m(t)e^{-j2\pi f_c t}$$

To find the frequency domain expression *S(f)* use the Fourier transform properties

$$
\begin{aligned}
&m(t) \leftrightarrow M(f) \\
&e^{j2\pi f_c t} \leftrightarrow \delta(f - f_c) \\
&e^{-j2\pi f_c t} \leftrightarrow \delta(f + f_c) \\
&\exp(j2\pi f_c t) m(t) \leftrightarrow M(f - f_c) \\
&\exp(-j2\pi f_c t) m(t) \leftrightarrow M(f + f_c)
\end{aligned}
$$

The result is

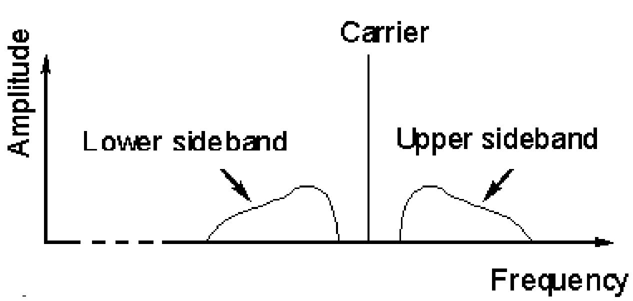
$$S(f) = \frac{A_c}{2}[\delta(f - f_c) + \delta(f + f_c)] + \frac{A_c k_a}{2}[M(f - f_c) + M(f + f_c)]$$

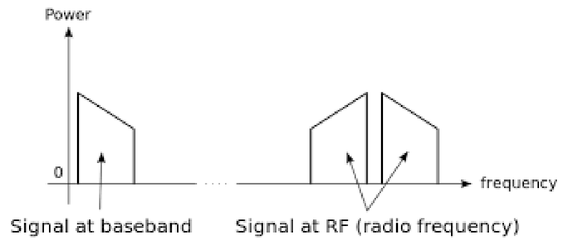
This expression shows both positive and negative frequencies. We often draw a picture of the positive frequencies only, and we can write

$$S(f > 0) = \frac{A_c}{2}\delta(f - f_c) + \frac{A_c k_a}{2} M(f - f_c)$$

The modulation spectrum $M(f)$ is shifted up so it is centered around f_c

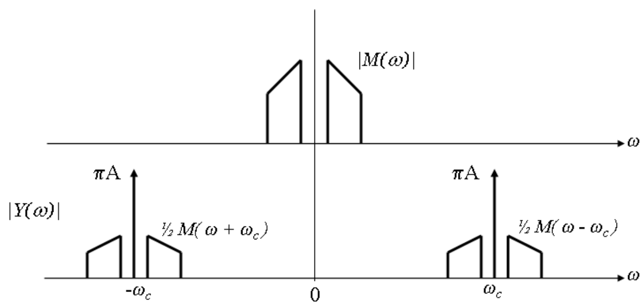
.





The above figures show only positive frequencies.

In the figure below, both positive and negative frequencies are shown. In this figure we write $\omega = 2\pi f$, $w_c = 2\pi f_c$



Observe that the AM signal spectrum is the sum of two components: the spectrum of the carrier wave at $\pm f_c$ plus the spectrum of the message shifted both up and down by f_c

Observe that because the message spectrum contains both positive and negative frequencies, the AM signal spectrum contains components both above and below the carrier frequency.

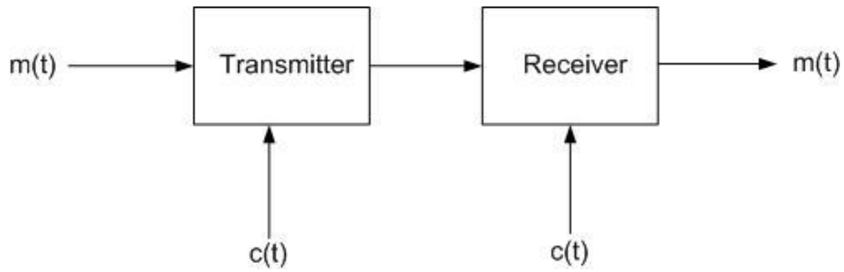
In general, $M(f)$ will change shape with each frame. For a particular frame, and for illustration purposes, we draw it as it would be for a voice signal: an asymmetrical shape that is zero for $f < 300$Hz , peaks near f=1000 Hz, and is zero for $f > 2{,}700$ Hz. (Think of the green line in sndpeek). The negative frequencies will be the mirror image of this shape.

2.1.6 Digital Messages Transmitted Using AM

Digital messages can be sent by using a binary modulating signal; the amplitude of the carrier is multiplied by a high voltage (logic one) and a low voltage (logic zero). A modulated wave with the capacity to represent a binary bit stream is the result. This method is known as amplitude shift keying (ASK). The most common case is where logic 1 is a positive voltage, say +1 volts and a logic 0 is a negative voltage say -1 volts. In general we can write $m(t) = \pm A_m$ Thus we can write the digital AM signal $s(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t) = A_c[1 \pm k_a A_m]cos 2\pi f_c t$ If we choose A_m such that $k_a A_m = 1$, then $s(t) = 2A_c cos 2\pi f_c t$ or 0 depending on whether a logic 1 or a logic 0 was sent. In this case the amplitude is shifted from $2A_c$ to 0, and it is called on-off keying (OOK) and can be used for Morse code transmission.

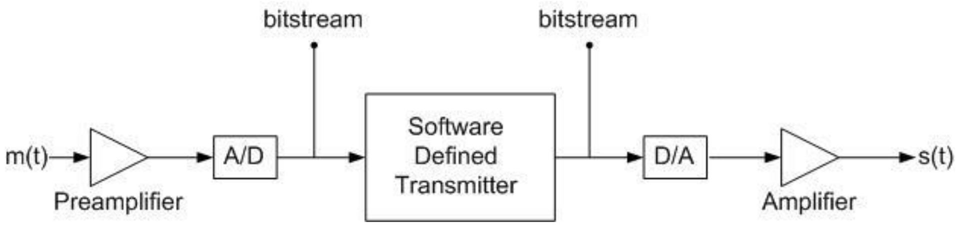
2.1.7 Building an AM transmitter digitally in software

Next we look at how to build transmitters and receivers. We have a message that we wish to send and we need a way to send and receive it.



Conceptual Modulation and Demodulation of the Signal m(t)

One approach is to use digital means, for example by programming the USRP using GNURadio. The USRP includes the DAC and analog IQ mixer, filters and amplifiers connected to an antenna.



Digital Transmitter Configuration

The transmitted AM signal is of the form:

$$s(t) = a(t)\cos(2\pi f_c t + \phi) = A_c[1 + k_a m(t)]\cos(2\pi f_c t + \phi)$$

where ϕ is a constant phase.

We can write the AM signal $s(t)$ as the real part of a complex signal:

$$s(t) = \mathrm{Re}[a(t)e^{j\phi}e^{j2\pi f_c t}] = \mathrm{Re}[\tilde{s}(t)e^{j2\pi f_c t}],$$

where the complex envelope:

$$\tilde{s}(t) = a(t)e^{j\phi} = a(t)\cos\phi + ja(t)\sin\phi$$
$$= i(t) + jq(t)$$

Usually we choose $\phi = 0$, so that $i(t) = a(t) = 1 + k_a m(t),\ q(t) = 0$

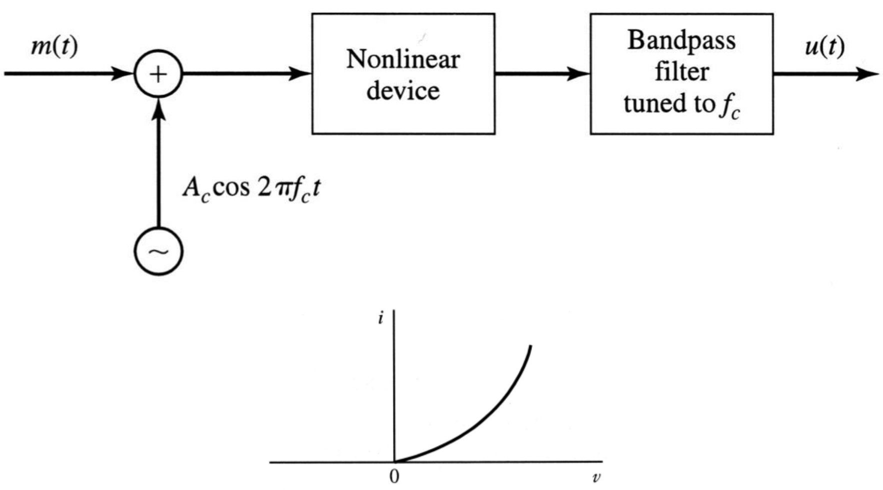
Given the message $m(t)$, the digital AM transmitter simply implements the equations $i(t) = 1 + k_a m(t)$

2.1.8 Analog-only method to create AM wave

An AM wave can be created using a power law modulator

One could implement a purely analog method instead using a nonlinear circuit with output $v_2(t) = \sum_i a_i [v_i(t)]^i = a_1 v(t) + a_2 v^2(t) + \ldots$

The nonlinearity could be a diode operating with a small input voltage. The nonlinearity includes a square term where the input is multiplied by itself.



Fundamental Analog Transmission Method

We want to multiply the message signal m(t) with the carrier wave c(t), in particular: $s(t) = [1 + k_a m(t)]c(t)$. Any non-linearity $v_2(t) = \sum_i a_i [v_i(t)]^i$ will cause multiplication where $i \geq 2$

As shown by the identity, $\cos\alpha\cos\beta = \frac{1}{2}[\cos(\alpha+\beta)+\cos(\alpha-\beta)]$, the result of multiplying two signals is sinusoidal components located at the sum and the difference of carrier and message frequencies.

We model the nonlinearity as the first two terms of the Taylor series expansion, $v_2(t) = \sum_{i=1,2} a_i[v_i(t)]^i$, where the input signal $v_1(t) = m(t) + c(t)$ is the sum (not the product) of the message and the carrier.

Thus we write

$$
\begin{aligned}
&v_2(t) = a_1 v_1(t) + a_2[v_2(t)]^2 \\
&v_1(t) = m(t) + c(t) \\
&v_i(t) = m(t) + A_c \cos(2\pi f_c t) \\
&\qquad\therefore \\
&v_2(t) = a_1[m(t) + A_c \cos(2\pi f_c t)] + a_2[m(t) + A_c \cos(2\pi f_c t)]^2 \\
&v_2(t) = a_1 m(t) + a_1 A_c \cos(2\pi f_c t) + a_2 m^2(t) + 2a_2 m(t) A_c \cos(2\pi f_c t) + a_2 A_c^2 \cos^2(2\pi f_c t)
\end{aligned}
$$

Using the trigonometric identity, $\cos^2(\alpha) = (1/2)[1+\cos(2\alpha)]$, we arrive at

$$a_2 A_c^2 \cos^2(2\pi f_c t) = (1/2) a_2 A_c^2 [1 + \cos(4\pi f_c t)]$$

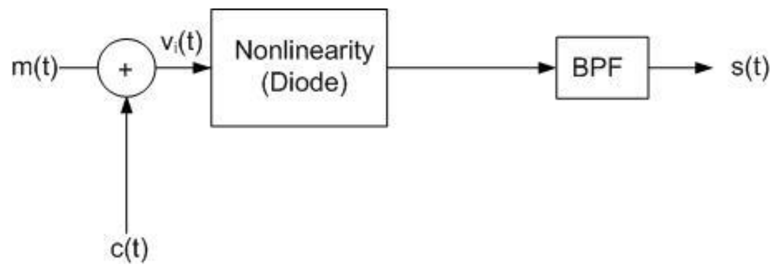
Sorting by frequency we get the following simplified expression:

$$
\begin{aligned}
&v_2(t) = a_1 m(t) + a_2 m^2(t) + (1/2) a_2 A_c^2 \\
&+\cos(2\pi f_c t)[a_1 A_c + 2a_2 A_c m(t)] \\
&+\cos(4\pi f_c t)[(1/2) a_2 A_c^2]^2
\end{aligned}
$$

We now apply a bandpass filter centred at f_c ,to get the modulator output

$$s(t) = a_1 A_c \cos 2\pi f_c t[1 + (2a_2 / a_2) m(t)]$$

Thus we have created an AM wave $s(t)$ by adding $m(t) + c(t)$ and then using a nonlinearity and band pass filter. For this AM wave $\mu = 2a_2 / a_1$



Adding Message and Carrier Prior to Nonlinearity

Exercise, write a frequency domain expression for $v_2(t)$.

$$
\begin{aligned}
&V_2(f) = a_1 \frac{A_c}{2}[\delta(f - f_c) + \delta(f + f_c)] \\
&+2a_2A_c[M(f - f_c) + M(f + f_c)] \\
&+a_1M(f) + a_2FT[m^2(t)] \\
&+\frac{a_2A_c^2}{2}[\delta(f) + \delta(f - 2f_c) + \delta(f + 2f_c)]
\end{aligned}
$$

Exercise: Another type of nonlinearity is diode switching. $v_2(t) = \begin{matrix} v_1(t) & c(t) > 0 \\ 0 & c(t) < 0 \end{matrix}$

In this case, the message waveform plus the carrier wave signal is half-wave rectified and the result is bandpass filtered at the carrier wave frequency. Show that the filter output contains an AM wave. Hint:

$v_1(t) = c(t) + m(t)$

$v_2(t) = [c(t) + m(t)]g_p(t)$

$g_p(t) =$ unipolar square wave at f_c

Find the Fourier series for the square wave

Result

$v_2(t) = [c(t) + m(t)][\frac{1}{2} + \frac{2}{\pi}\cos 2\pi f_c t + ...]$

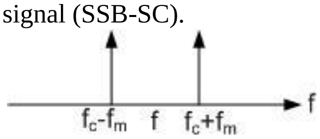
$= \frac{A_c}{2}[1 + \frac{4}{\pi A_c}m(t)]\cos 2\pi f_c t + ...$

$=$ AM wave with $k_a = \frac{4}{\pi A_c} +$ components at frequencies removed from f_c

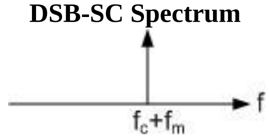
2.2 Double Sideband Modulation

Carrier waves require power to transmit but they do not in themselves contain any information about the message.

Sometimes we want to transmit an AM wave without its carrier wave (this lessens the power requirements for transmission and thus may be a more economical course of action for certain applications). Filtering out the carrier frequency band leads to a double sideband - suppressed carrier signal (DSB-SC) and further filtering out one of either the upper sideband (USB) or lower sideband (LSB) will produce what is called a single sideband – suppressed carrier signal (SSB-SC).



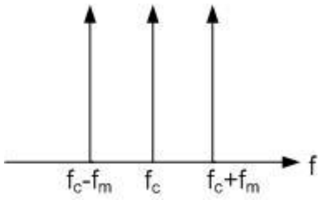
DSB-SC Spectrum



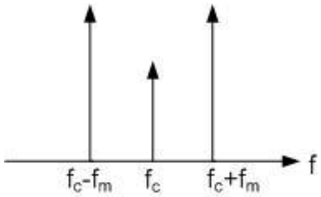
SSB-SC Spectrum

Yet another type is vestigial sideband modulation (VSB) which can be used practically for transmission of television signals. In essence VSB = SSB + Carrier.

Modulation Type	Advantages	Disadvantages
AM	Easily demodulated	High power requirements due to carrier
DSB-SC	Lower power requirements	Carrier location required. Redundant SB information
SSB-SC	Less bandwidth and power requirements	Don't know where carrier is
VSB	Less bandwidth and power requirements. Lack of redundancy	



100% Modulation



Overmodulation

2.2.1 Double sideband suppressed carrier waveform

For a double sideband signal we know that the only difference from the AM case is the lack of carrier term. Thus with a message $m(t) = A_m \cos 2\pi f_m t$ and a carrier wave $c(t) = A_c cos 2\pi f_c t$, the DSB-SC signal is represented by

$$s(t) = m(t)c(t) = A_c A_m \cos(2\pi f_m t)\cos(2\pi f_c t)$$
$$= \frac{A_c A_m}{2}[\cos(2\pi[f_c - f_m]t) + \cos(2\pi[f_c + f_m]t)]$$

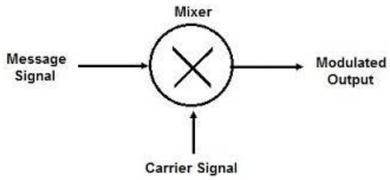
where we let the constant $A_c A_m$ represent any scaling that has taken place by way of amplification or any hardware effects.

Here we have used the trigonometric identity

$$\cos\alpha\cos\beta = \frac{1}{2}[cos(\alpha-\beta) + cos(\alpha+\beta)]$$

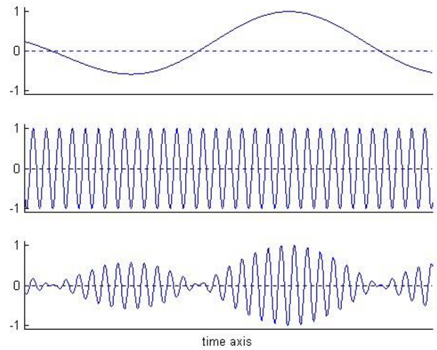
For DSB-SC with message $m(t) = A_m \cos 2\pi f_m t$, $I(t) = A_c \cos(2\pi f_m t)$ and $Q(t) = 0$.

A DSB-SC modulator is simply a multiplier that multiplies the message signal $m(t)$ with the carrier $c(t)$



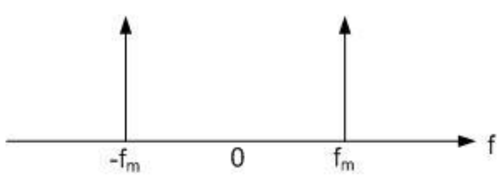
The figure below illustrates the waveforms $m(t), c(t), s(t)$ respectively.

Note the phase reversal of the carrier wave when the message signal is less than zero.

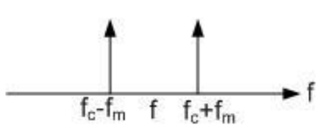


2.2.2 DSB spectrum

The spectrum of the single tone message $m(t) = A_m \cos 2\pi f_m t$ is



The spectrum of a DSB signal with carrier $c(t) = A_c \cos 2\pi f_c t$ and message $m(t) = A_m \cos 2\pi f_m t$ is two peaks at $f_c - f_m$ and $f_c + f_m$ and no power at f_c



The DSB signal spectrum will also include negative frequencies with peaks at $-(f_c - f_m)$ and $-(f_c + f_m)$ (not shown in the figure above).

Notice the general principle at work here: multiplying two cos waves together yields the sum and difference frequencies, as can be seen from the trigonometric identity

$$\cos\alpha\cos\beta = \frac{1}{2}[cos(\alpha-\beta)+cos(\alpha+\beta)]$$

There is another related principle at work here. Notice that in the expression

$$s(t) = m(t)c(t) = A_c A_m \cos 2\pi f_m t \cos 2\pi f_c t$$

the message signal is multiplied by the carrier wave. Whenever a message signal is multiplied by a carrier wave, the spectrum of the message signal is shifted both up and down by the carrier frequency.

Observe that because the message spectrum contains both positive and negative frequencies, the signal spectrum contains components both above and below the (suppressed) carrier frequency.

Exercise: for a DSB signal, find the power in each sideband.

Exercise: for a DSB-SC signal, find the complex envelope
Exercise: write the spectrum of a DSB signal with a single tone message using delta functions.
Exercise: write the spectrum of a DSB signal with a general message $m(t) \leftrightarrow M(f)$

2.3 Single sideband suppressed carrier (SSB-SC) and Hilbert transform

SSB type modulation may be viewed as double sideband with one of the sidebands removed.

Ironically, we will see that to *remove* the sideband, we *add* an additional term to the expression for DSB-SC.

2.3.1 Derivation of SSB-SC using Hilbert transform and analytic signals

We can start the derivation of SSB-SC signal mathematics by considering the message spectrum $M(f)$. Note that $M(f)$ has two sidebands on either side of DC or zero frequency. Thus $M(f)$ contains both positive and negative frequencies. For any $m(t)$ that is real-valued with Fourier transform $M(f)$, we find that $M(-f) = M^*(f)$ where * denotes complex conjugate (same amplitude, minus the phase).

We can create an SSB-SC waveform in the same way as we created DSB-SC, except that we add one new step:

- first, eliminate one of the sidebands (say the negative frequency sideband) in the message spectrum to create a new message wtih asymmetrical spectrum,
- multiply the new message by the carrier wave as was done for DSB-SC.

The negative frequency sideband is eliminated by multiplying $M(f)$ by the Heaviside step function $2u(f) = 1 + \mathrm{sgn}(f) = 2$ for $f > 0$ and 0 for $f < 0$

The new message can be written

$$M_+(f) = 2u(f)M(f) = [1 + \mathrm{sgn}(f)]M(f) = M(f) + \mathrm{sgn}(f)M(f)$$

The new message in the time domain

$$\begin{aligned}
m_+(t) &= \text{ inverse FT } \{M_a(f)\} \\
&= \text{ inverse FT } \{M(f)\} + \text{ inverse FT } \{\mathrm{sgn}\, f \;\; M(f)\} \\
&= m(t) + \text{inverse FT } \{\mathrm{sgn}(f)\} \otimes \text{ inverse FT } \{M(f)\} \\
&= m(t) + \frac{j}{\pi t} \otimes m(t) \\
&= m(t) + j\hat{m}(t)
\end{aligned}$$

where $\otimes$ is the convolution operator and we have used the Fourier transform pair $\frac{j}{\pi t} \leftrightarrow \mathrm{sgn}(f)$. $\hat{m}(t)$ is the so-called Hilbert transform of $m(t)$ defined below. Recall that multiplication in the frequency domain corresponds to convolution in the time domain.

The new message $m_+(t) = m(t) + j\hat{m}(t)$ with only positive frequencies is called an *analytic signal* or *pre-envelope.*

2.3.2 Hilbert transform definition and properties

$\hat{m}(t)$ is the Hilbert Transform of $m(t)$. The Hilbert transform is defined as

$$\hat{m}(t) = m(t) \otimes (1/\pi t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{m(\tau)}{t-\tau} d\tau$$

The Hilbert transform is a special kind of non-causal filter with impulse response $1/\pi t$.

It turns out that the Hilbert transform shifts each sinusoidal component of $m(t)$ by 90 degrees.

For $m(t) = \cos 2\pi f_m t$, $\hat{m}(t) = \cos(2\pi f_m t - \pi/2) = \sin 2\pi f_m t$ for any and all f_m

The proof is in showing that $\cos 2\pi f_m t \otimes (1/\pi t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\cos 2\pi f_m \tau}{t-\tau} d\tau = \sin 2\pi f_m t$

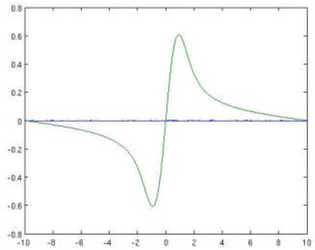
The proof is easier in the frequency domain

$$m(t) \leftrightarrow M(f)$$
$$\hat{m}(t) \leftrightarrow -j \,\mathrm{sgn}\, f \; M(f)$$
$$m_+(t) = m(t) + j\hat{m}(t) \leftrightarrow M(f) + \mathrm{sgn}\, f \; M(f) = M_+(f)$$

Thus the Hilbert transformer is an LTI system with transfer function $H(f) = -j \,\mathrm{sgn}\, f$ and impulse response $h(t) = \frac{1}{\pi t}$. The transfer function shows that every positive frequency is multiplied by $-j$ (and every negative frequency is multiplied by $+j$), thus shifiting the phase by 90 degrees. For example, for a single tone message

$$m(t) = \cos 2\pi f_m t = \frac{1}{2}[e^{j2\pi f_m t} + e^{-j2\pi f_m t}]$$
$$\text{recall } e^{j2\pi f_m t} \leftrightarrow \delta(f - f_m)$$
$$M(f) = \frac{1}{2}[\delta(f - f_m) + \delta(f + f_m)]$$
$$\hat{M}(f) = -j \,\mathrm{sgn}\, f \; M(f) = \frac{1}{2}[-j\delta(f - f_m) + j\delta(f + f_m)]$$
$$\hat{M}(f) = \frac{1}{2j}[\delta(f - f_m) - \delta(f + f_m)]$$
$$\text{thus } \hat{m}(t) = \frac{1}{2j}[e^{j2\pi f_m t} - e^{-j2\pi f_m t}] = \sin 2\pi f_m t$$

The figure below shows an approximation of the impulse response. In theory $h(t)$ flips from minus infinity to infinity at $t = 0$



An analytic signal (or pre-envelope) is defined as $m_+(t) = m(t) + j\hat{m}(t)$, where $\hat{m}(t) = m(t) \otimes (1/\pi t)$ is the Hilbert transform of $m(t)$. The pre-envelope contains only positive frequencies. We will use this pre-envelope to create an SSB-SC signal with only the upper sideband, also called simply an *upper sideband (USB)* signal.

The reasoning above can be repeated to define a pre-envelope with only negative frequencies $m_-(t) = m(t) - j\hat{m}(t)$ which is used to create a *lower sideband* (LSB) signal.

2.3.3 SSB-SC signal derived from pre-envelope

An SSB-SC signal is created by taking the analytic signal or pre-envelope as the message signal, shifting it up to the carrier frequency, and taking the real part.

Recall from the Fourier transform properties that multiplying any signal by a complex exponential at f_c shifts the spectrum of that signal by f_c.
For a baseband message signal $M(f)$ centered at $f = 0$, the spectrum is shifted so that it is centered at f_c

$$\begin{aligned} &m(t) \leftrightarrow M(f) \\ &m(t)e^{j2\pi f_c t} \leftrightarrow M(f - f_c) \end{aligned}$$

For DSB-SC, $\quad s(t) = m(t)c(t)$

For SSB-SC with upper sideband only

$$\begin{aligned} s(t) &= \mathrm{Re}\{[m(t) + j\hat{m}(t)]e^{j2\pi f_c t}\} \\ &= \mathrm{Re}\{[m(t) + j\hat{m}(t)][\cos 2\pi f_c t + j\sin 2\pi f_c t]\} \\ &= m(t)\cos 2\pi f_c t - \hat{m}(t) sin 2\pi f_c t \end{aligned}$$
,

In general, an SSB-SC signal with a general message can be written

$s(t) = m(t)\cos 2\pi f_c t \mp \hat{m}(t) sin 2\pi f_c t$

where the minus sign is for upper sideband and the plus sign for lower sideband

If $m(t) = \cos(2\pi ft)$, then $\hat{m}(t) = \sin(2\pi ft)$ where we denote the Hilbert transform by the hat ^ symbol, and

$s(t) = \cos 2\pi f_m t \cos 2\pi f_c t \mp \sin 2\pi f_m t \sin 2\pi f_c t = \cos 2\pi (f_c \pm f_m) t$

which is the single sideband at $f_c \pm f_m$ (plus for USB, minus for LSB).

Thus for USB, $I(t) = m(t)$ and $Q(t) = \hat{m}(t)$ For a single tone message
$m(t) = A_m cos 2\pi f_m t$, $I(t) = \cos(2\pi f_m t)$ and $Q(t) = \sin(2\pi f_m t)$.
The constant A_c accounts for the signal power.

A single sideband USB signal $s(t) = m(t)\cos 2\pi f_c t - \hat{m}(t) sin 2\pi f_c t$ can be written as

$s(t) = \mathrm{Re}\{a(t) e^{j\phi} e^{j2\pi f_c t}\} = \mathrm{Re}\{[m(t) + j\hat{m}(t)] e^{j2\pi f_c t}\}$

Thus the complex envelope of an SSB-SC signal $a(t)e^{j\phi} = m(t) + j\hat{m}(t)$ is an analytic signal created from the (real) message $m(t)$.

2.3.4 Single tone modulation for SSB

$$
\begin{aligned}
m(t) &= A_m \cos 2\pi f_m t \\
\hat{m}(t) &= A_m \sin 2\pi f_m t
\end{aligned}
$$

$$
\begin{aligned}
s(t) &= \frac{A_c}{2}[m(t)\cos 2\pi f_c t \pm \hat{m}(t) \sin 2\pi f_c t] \\
&= \frac{A_c A_m}{2}[\cos 2\pi f_m t \cos 2\pi f_c t \mp \sin 2\pi f_m t \sin 2\pi f_c t] \\
&= \frac{A_c A_m}{2}\cos[2\pi (f_c \pm f_m) t]
\end{aligned}
$$

using

$cos(\alpha \pm \beta) = \cos\alpha \cos\beta \mp \sin\alpha \sin\beta$

$s(t)$ is a single tone at frequency $f_c \pm f_m$

2.3.5 SSB-SC in time and frequency domain

$$\text{message } m(t) \Leftrightarrow M(f)$$
$$\text{SSB-SC } s(t) \Leftrightarrow S(f) = M(f - f_c) \text{ for } f > f_c$$
$$\text{for positive frequencies}$$

proof

$$s(t) = \frac{A_c}{2}[m(t)\cos 2\pi f_c t \pm \hat{m}(t)\sin 2\pi f_c t]$$
$$\hat{M}(f) = M(f)H(f) = -j\,\text{sgn}\, f\, M(f)$$
$$S(f) = \frac{A_c}{2}[M(f) \otimes \frac{1}{2}\{\delta(f - f_c) + \delta(f + f_c)\backslash$$
$$\pm -j\,\text{sgn}\, f\, M(f) \otimes \frac{1}{2j}\{\delta(f - f_c) - \delta(f + f_c)\}]$$
$$= \frac{A_c}{4}[M(f) \otimes \delta(f - f_c) + M(f) \otimes \delta(f + f_c)$$
$$\pm \text{sgn}\, f\, M(f) \otimes \delta(f - f_c)$$
$$\mp sgn\, f\, M(f) \otimes \delta(f + f_c)]$$
$$= \frac{A_c}{4}[M(f - f_c) + M(f + f_c)$$
$$\pm \text{sgn}(f - f_c) M(f - f_c) \mp sgn(f + f_c) M(f + f_c)$$
$$= \frac{A_c}{2} M(f - f_c),\, f - f_c > 0,\, f > f_c$$

Similarly

$$S(f) \;=\; \frac{A_c}{2} M(f + f_c),\, f + f_c > 0,\, f < -f_c$$

2.3.6 SSB-SC transmitter 1 using Hilbert transform

To create an SSB-SC signal at f_c using the SSB-SC equations above, we can:

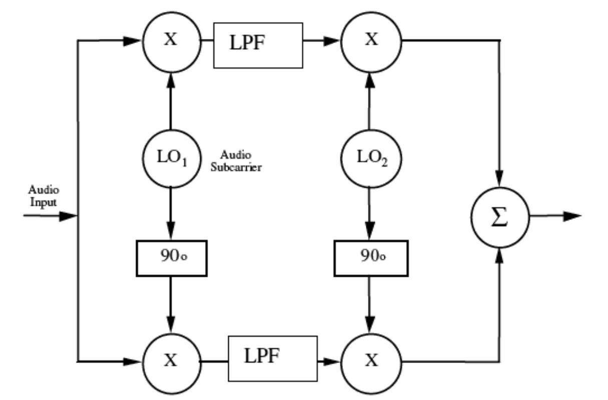
1. Take the Hilbert transform $\hat{m}(t) = m(t) \otimes (1/\pi t)$ of the message.
2. Create the analytic signal $\tilde{m}(t) = m(t) + j\hat{m}(t)$ for USB or $\tilde{m}(t) = m(t) - j\hat{m}(t)$ for LSB

3. Upconvert it to the desired carrier frequency by multiplying by $e^{j2\pi f_c t}$
4. Take the real part.

The upconversion to a radio frequency (RF) wave at f_c (steps 3 and 4) is the function of the *USRP Sink* block. Thus an SSB signal is generated by a USRP sink block (standard IQ transmitter) with inputs $i(t) = m(t)$ and $q(t) = \hat{m}(t)$ for USB.

2.3.7 SSB-SC transmitter 2, avoids Hilbert transform

The Hibert transform can be inconvenient to compute.
An SSB-SC signal can be created without using the Hilbert transform with the so-called Weaver modulator, shown below in real form



and in complex form

$$
\begin{array}{l}
m(t) \rightarrow \otimes \Rightarrow LPF(f_1) \Rightarrow \otimes \Rightarrow \rightarrow s(t) \\
\qquad\quad \Uparrow \qquad\qquad\qquad\qquad \Uparrow \\
\quad e^{-j2\pi f_1 t} \qquad\qquad\qquad e^{j2\pi(f_c+f_1)t}
\end{array}
$$

The signs of the exponentials are chosen for USB in this example.

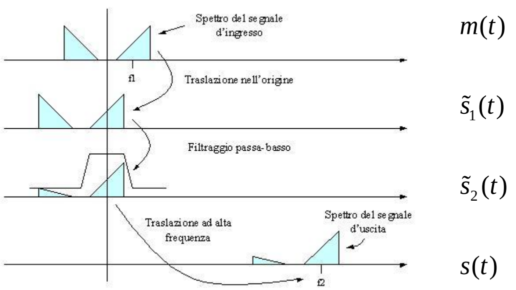
LO1 operating at f_1 is chosen to be in the centre of the band of the message $m(t)$, e.g. $f_1 = 1500$ Hz for a voice bandwidth message 300-2700 Hz.

LO2 operates at $f_c + f_1$.

In the figure below, we observe how the multiplying the messsage by the complex exponetials shift the spectrum of the message.

The operations to create the USB signal are as follows:

- The real message spectrum $m(t)$ is first shifted down by f_1 thus making it a complex signal $\tilde{s}_1(t)$ with an asymmetrical spectrum.
- $\tilde{s}_1(t)$ is low pass filtered to remove the lower sideband resulting in $\tilde{s}_2(t)$ which is now a complex signal
- $\tilde{s}_2(t)$ is shifted up by $f_c + f_1$ resulting in another complex signal related to $s(t)$ with positive frequencies only.
- The real part of this signal is selected, or the real and and imaginary parts of this signal are added together, resulting in a real signal $s(t)$ with positive and negative frequency components.
- The net result of the two frequency conversions is that the zero frequency component of $m(t) \leftrightarrow M(f)$ has been shifted up to the suppressed carrier frequency f_c and the components of $m(t) \leftrightarrow M(f)$ above zero frequency (i.e. the message content are shifted to be above f_c, i.e. $M(f) \to M(f - f_c)$



$$m(t) \to \otimes \Rightarrow LPF(f_1) \Rightarrow \otimes \Rightarrow \to s(t)$$

$$\Uparrow \quad \tilde{s}_1(t) \qquad \tilde{s}_2(t) \quad \Uparrow$$

$$e^{\pm j2\pi f_1 t} \qquad\qquad e^{j2\pi(f_c \mp f_1)t}$$

The operations to create the USB signal are written mathematically as follows:

input $m(t)$
after complex multiply with $e^{-j2\pi f_1 t}$
$\tilde{s}_1(t) = m(t)e^{-j2\pi f_1 t} = m(t)\cos 2\pi f_1 t - jm(t)sin2\pi f_1 t$
after LPF the result $\tilde{s}_2(t)$ is the analytic signal shifted by f_1
This does not arise from algebra, but is clear from the frequency diagram
$\tilde{s}_2(t) = [m(t) + j\hat{m}(t)]e^{-j2\pi f_1 t}$
$= m(t)\cos 2\pi f_1 t - jm(t)sin2\pi f_1 t + j\hat{m}(t)cos2\pi f_1 t + \hat{m}(t)sin2\pi f_1 t$
after mixer
$\tilde{s}_3(t) = \tilde{s}_2(t)e^{j2\pi(f_c+f_1)t} = [m(t) + j\hat{m}(t)]e^{-j2\pi f_1 t}e^{j2\pi(f_c+f_1)t} = [m(t) + j\hat{m}(t)]e^{j2\pi f_c t}$
$= m(t)\cos 2\pi f_c t - \hat{m}(t)sin2\pi f_c t + j[\hat{m}(t)\cos 2\pi f_c t + m(t)sin2\pi f_c t]$
$s(t)$ SSB-SC output USB, can choose i or q

Exercise: repeat the figure and mathematics for LSB

It is also possible to make a real USB signal by adding the real and imaginary branch outputs together, adding $i+q$ not $i+jq$

if add $i+q$

$$s(t) = m(t)\cos 2\pi f_c t - \hat{m}(t)sin2\pi f_c t + \hat{m}(t)\cos 2\pi f_c t + m(t)sin2\pi f_c t$$
$$= m(t)[\cos 2\pi f_c t + sin2\pi f_c t] + \hat{m}(t)[\cos 2\pi f_c t - \sin 2\pi f_c t]$$
$$= m(t)\cos(2\pi f_c t - \pi/4) + \hat{m}(t)\cos(2\pi f_c t + \pi/4)$$
$$= m(t)\cos(2\pi f_c t - \pi/4) + \hat{m}(t)\cos(2\pi f_c t - \pi/4 + \pi/2)$$
$$= m(t)\cos(2\pi f_c t - \pi/4) - \hat{m}(t)\sin(2\pi f_c t - \pi/4)$$

thus

$$s(t) = \mathrm{Re}\{\tilde{s}_3(t)e^{-j\pi/4}\} = \mathrm{Re}\{[m(t) + j\hat{m}(t)]e^{j(2\pi f_c t - \pi/4)}\}$$

using

$$\cos(\alpha + \pi/4) = \cos\alpha\cos\pi/4 - \sin\alpha\sin\pi/4 = \sqrt{(2)}[\cos\alpha - \sin\alpha]$$
$$\cos(\alpha - \pi/4) = \cos\alpha\cos\pi/4 + \sin\alpha\sin\pi/4 = \sqrt{(2)}[\cos\alpha + \sin\alpha]$$
$$\cos(\alpha + \pi/2) = -\sin\alpha$$

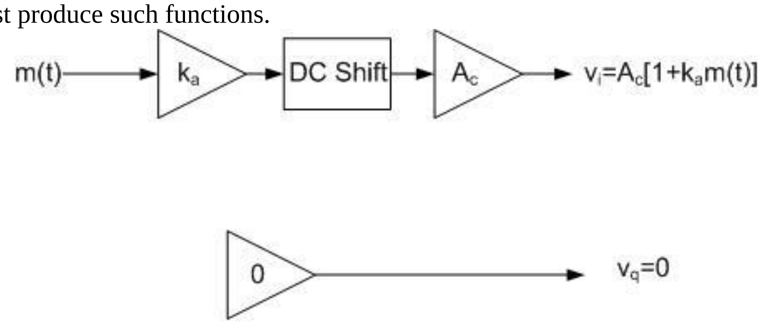
This completes the discussion of the operations in an SSB-SC transmitter that creates an SSB-SC signal $s(t) = m(t)\cos 2\pi f_c t \mp \hat{m}(t)sin2\pi f_c t$ with message $m(t)$, where the upper sign is for USB and the lower sign is for LSB.

2.4 Review of AM, DSB, SSB transmitters in IQ format

In preparation for discussing receivers, we review the AM, DSB, SSB signals in IQ format. Each transmission method (AM, DSB-SC, SSB, etc) requires its own unique map that serves to form $v_i(t)$ and $v_q(t)$ from the message $m(t)$.

2.4.1 Amplitude Modulation:

We already know that $s(t) = A_c[1+k_a m(t)]\cos(2\pi f_c t)$, and for the special case where $m(t) = \cos 2\pi f_m t$ then $s(t) = A_c[1+k_a A_m \cos(2\pi f_m t)]\cos(2\pi f_c t)$ Since in general $s(t) = v_i(t)\cos(2\pi f_c t) - v_q(t)\sin(2\pi f_c t)$, to form $v_i(t)$ and $v_q(t)$ in the IQ Transmitter/Receiver we see that $v_i(t) = A_c[1+k_a A_m \cos(2\pi f_m t)]$ and $v_q(t) = 0$. Thus our map must produce such functions.



Map to transmit AM

Recall that $v_i(t) = a(t)\cos\psi(t)$ and $v_q(t) = a(t)\sin\psi(t)$, thus

$a(t) = A_c[1+k_a m(t)] = A_c[1+k_a A_m \cos(2\pi f_m t)]$ and $\psi(t) = 0$

We could also choose $\psi(t) \neq 0$ and write

$v_i(t) = a(t)\cos\psi(t) = A_c[1+k_a m(t)]\cos\psi(t)$

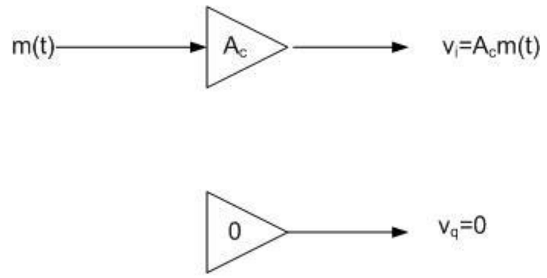
$v_q(t) = a(t)\sin\psi(t) = A_c[1+k_a m(t)]\sin\psi(t)$

The value of the phase $\psi(t)$ does not affect the message contained within $a(t)$

2.4.2 DSB-SC

For a double sideband signal we know that the only difference from the AM case is the lack of carrier term. Thus with a message $m(t) = A_m \cos 2\pi f_m t$ the DSB-SC signal is represented by

$s(t) = A_c A_m \cos(2\pi f_m t)\cos(2\pi f_c t) = \dfrac{A_c A_m}{2}[\cos(2\pi[f_c - f_m]t) + \cos(2\pi[f_c + f_m]t)]$ where we let the constant $A_c A_m$ represent any scaling that has taken place by way of amplification or any hardware effects. For DSB-SC, $v_i(t) = A_c \cos(2\pi f_m t)$ and $v_q(t) = 0$.



Map to transmit DSB-SC

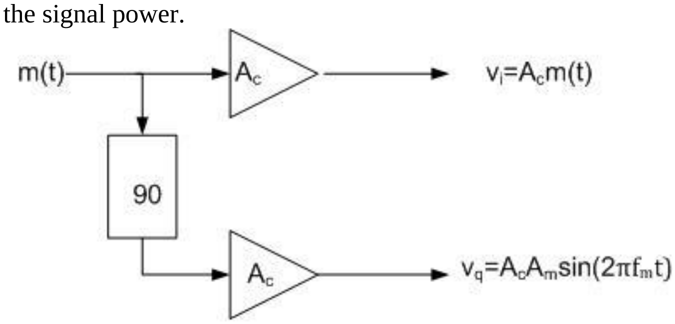
2.4.3 SSB-SC

SSB type modulation might be understood as double sideband with one of the sidebands removed. It would follow that adding something to the double sideband map would be adequate to generate single sideband signal. We require a transform which when taken of a certain signal, would negate that signal if the two were added together. We may use, for this purpose, the Hilbert Transform that shifts all frequencies by 90 degrees.

In general $s(t) = m(t)\cos 2\pi f_c t \mp \hat{m}(t) sin 2\pi f_c t$ where the minus sign is for upper sideband and the plus sign for lower sideband

If $m(t) = \cos(2\pi ft)$, then $\hat{m}(t) = \sin(2\pi ft)$ where we denote the Hilbert transform by the hat ^ symbol, and $s(t) = \cos 2\pi f_m t \cos 2\pi f_c t \mp \sin 2\pi f_m t \sin 2\pi f_c t = \cos 2\pi(f_c \pm f_m)t$

Thus the map for SSB is $v_i(t) = m(t)$ and $v_q(t) = \hat{m}(t)$ For a single tone message $m(t) = A_m cos 2\pi f_m t$, $v_i(t) = \cos(2\pi f_m t)$ and $v_q(t) = \sin(2\pi f_m t)$. The constant A_c accounts for the signal power.



Map to transmit SSB-SC

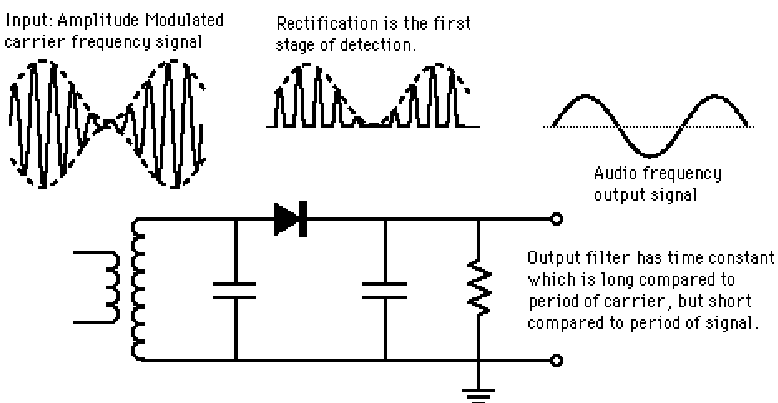
3 AM Receivers

The purpose of a receiver is to extract or recover the message $m(t)$ from the modulated wave $s(t)$. There exist many different kinds of receiver structures and designs. Digital and analog methods are used, each having a specific structure that is unique to the method of modulation used. (We shall focus here on AM, DSB, and SSB).

3.1 Analog AM receiver

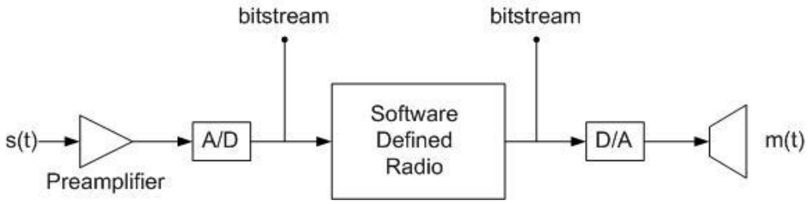
Analog demodulation of an AM wave (see below) can be achieved via an envelope detector. The envelope represents the modulating message wave and is extracted through the use of a rectifying diode and a lowpass filter (high frequency carrier term is removed while the message with lower frequency f_m remains.

$$s(t) = a(t)\cos(2\pi f_c t + \phi) = A_c[1 + k_a m(t)]\cos(2\pi f_c t + \phi)$$



3.2 Digital software receivers

The figure below is a representation of the main principles involved in demodulating a transmission using a digital architecture.



3.2.1 Digital software AM receiver in complex notation

The transmitted AM signal is of the form:

$$s(t) = a(t)\cos(2\pi f_c t + \phi) = A_c[1 + k_a m(t)]\cos(2\pi f_c t + \phi)$$

where ϕ is a constant phase. The AM receiver recovers $m(t)$ from $s(t)$. One method is to recover $a(t) = 1 + k_a m(t)$ and subtract the DC component to obtain $m(t)$.

To show how this is done in software with the USRP and GNURadio Companion (GRC), recall that the USRP source block has a complex output with real and imaginary components $i(t)$ and $q(t)$.

We can write the AM signal $s(t)$ as the real part of a complex signal:

$$s(t) = \mathrm{Re}[a(t)e^{j\phi}e^{j2\pi f_c t}] = \mathrm{Re}[\tilde{s}(t)e^{j2\pi f_c t}],$$

where the complex envelope:

$$\tilde{s}(t) = a(t)e^{j\phi} = a(t)\cos\phi + ja(t)\sin\phi$$
$$= i(t) + jq(t)$$

Thus the USRP source block with frequency set to f_c will have outputs:

$$i(t) = a(t)\cos\phi, \; q(t) = a(t)\sin\phi.$$

To obtain $a(t)$, we take the magnitude of the complex envelope $\tilde{s}(t)$, thus we can write:

$$
\begin{aligned}
|\tilde{s}(t)| &= |i(t) + jq(t)| \\
&= |a(t)\cos\phi + ja(t)\sin\phi| \\
&= a(t)\,|\cos\phi + j\sin\phi| \\
&= a(t)\sqrt{\cos^2\phi + \sin^2\phi} = a(t)
\end{aligned}
$$

This shows that we can recover $a(t) = 1 + k_a m(t)$ regardless of the value of ϕ.

The GRC *Complex to Magnitude* block allows us to obtain the magnitude of the complex envelope by performing the function $a(t) = |i(t) + jq(t)|$.

If there is frequency offset, then $\phi = 2\pi\Delta ft$, but as we have just seen, $|\tilde{s}(t)| = a(t)$ is not affected by the value of ϕ and thus not affected by any frequency offset Δf .

3.2.2 Tuning in (selecting) one particular AM signal with the USRP

The USRP multiplies the real valued radio frequency signal $s(t)$ by $e^{j2\pi f_c t}$ to generate $i(t) + jq(t)$. This process is called *complex downmixing* and is equivalent to the standard IQ receiver shown in Figure 3.

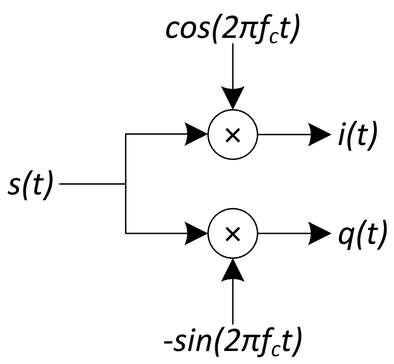


Figure 3 Complex Mixer

Recall:

$$
\begin{aligned}
&s(t) \leftrightarrow S(f) \\
&e^{-j2\pi f_c t} s(t) \leftrightarrow S(f + f_c)
\end{aligned}
$$

The spectrum $S(f)$ of the real radio frequency (RF) signal $s(t)$ will be symmetric about zero. After complex downmixing, the resulting signal is complex and the frequency spectrum $S(f + f_c)$ is no longer symmetric about zero.

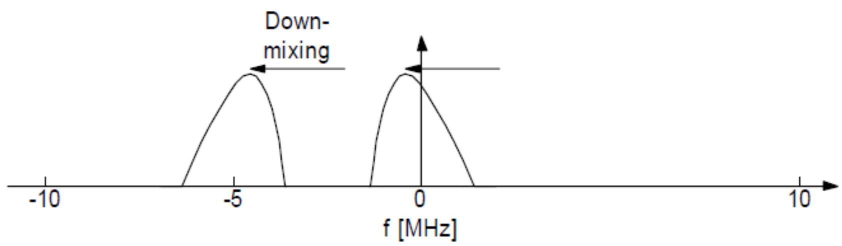


Figure 4 Downmixing

The complex signal output from the USRP source block $i(t)+jq(t)$ is bandlimited to the sampling rate of the USRP source block. The USRP source block output can be recorded to a file and used again at a later time. This file source will have the same sampling rate and bandwidth as the USRP sink block used to record it.

With a sampling rate of 256 kHz and complex samples, the bandwidth will be 256 kHz (because the complex signal spectrum is not symmetric and does not have redundant mirror-image positive and negative frequencies).

AM radio broadcast signals have a bandwidth of ± 5 KHz and the carrier frequencies are spaced apart by 10 KHz in North America (9 KHz elsewhere), e.g. there can be carriers at 710, 720, 730 KHz, etc.

With a file source sampled at 256 kHz, there can be as many as 24 different AM broadcast signals.

The AM broadcast signal with carrier frequency $f_c = f_d$ (f_d is set in the USRP source block) will appear at zero Hz after the downconversion (at the USRP source output). Other signals at carrier frequencies $f_c \pm nf_0$ kHz will appear at multiples of $f_0 = 10$ KHz away from zero Hz

We need to create a filter to select the one signal we want (the one with carrier frequency f_c that now appears at 0 Hz). A low pass filter with 5 KHz cutoff frequency will do the job, since all the other signals are centered at frequencies at least 10 KHz away from zero Hz.

To "tune in" (receive) one of the other signals, we can shift the spectrum of the USRP source output by nf_0 KHz by multiplying the complex signal $i(t)+jq(t)$ by $e^{-j2\pi nf_0 t} = \cos 2\pi nf_0 t - j\sin 2\pi nf_0 t$, so that the signal that first appeared at nf_0 Hz now appears at zero Hz.

3.3 DSB-SC receiver

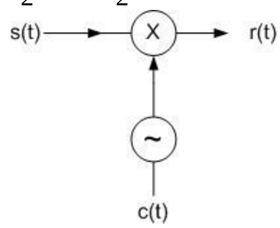
Recall this time that a DSB wave is comprised of only the two (redundant) message frequency bands (the carrier has been filtered out before transmission):

$$s(t) = m(t)c(t) = m(t)cos(2\pi f_c t)$$

The modulated double sideband signal has no carrier frequency and therefore exhibits a periodic phase reversal which makes the simple envelope detector described previously inadequate for demodulation (the rectifying diode will leave us with a rectified version of the envelope: received envelope = $|m(t)|$).

How to properly receive $m(t)$ from $s(t)$ for a DSB-SC signal ? One general method may be understood by observing the result of multiplying the modulated wave $s(t)$ by the carrier frequency, $c(t)$. Let $r(t)$ represent the received signal within the receiver structure.

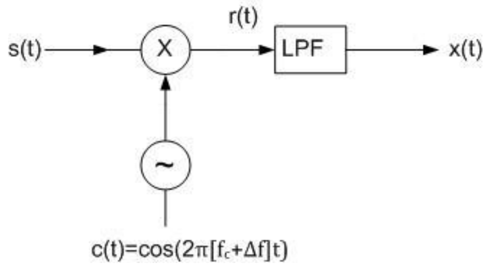
$$r(t) = s(t)c(t)$$
$$r(t) = m(t)c(t)c(t)$$
$$r(t) = m(t)\cos^2(2\pi f_c t)$$
$$r(t) = m(t)[\frac{1}{2}(1+\cos 4\pi f_c t)]$$
$$r(t) = \frac{1}{2}m(t) + \frac{1}{2}\cos 4\pi f_c t$$



DSB receiver first step

Thus we have created a signal comprised of a high frequency term (a DSB signal and an amplitude scaled version of our message. All that must be done to extract $m(t)$ from the signal is to apply a lowpass filter to the signal: $r_{LP}(t) = \frac{1}{2}m(t)$

This method, however effective, requires the party operating the receiver to know the exact carrier frequency, f_c and the phase of the carrier wave. Both instances of the carrier wave (that used here at the receiver and that which is used initially in the modulation process at the transmitter) must both have exactly the same phase and frequency.



Production of Signal x(t) with LPF

3.3.1 DSB receiver with frequency and phase offset

When the frequency and phase are not exactly the same, the error terms (frequency error Δf and phase shift φ) may be represented by $\psi = 2\pi\Delta f t + \varphi$. Using the appropriate trig identities, we now may write the received signal before the low pass filter

$$r(t) = m(t)\cos(2\pi f_c t)\cos[2\pi f_c t + 2\pi(\Delta f)t + \varphi]$$
$$r(t) = m(t)\cos(2\pi f_c t)\cos[2\pi f_c t + \psi]$$
$$r(t) = m(t)\cos(2\pi f_c t)[\cos(2\pi f_c t)\cos\psi - \sin(2\pi f_c t)\sin\psi]$$
$$r(t) = m(t)\cos^2(2\pi f_c t)\cos\psi - m(t)\cos(2\pi f_c t)\sin(2\pi f_c t)\sin\psi$$
$$r(t) = \frac{m(t)}{2}[1 + \cos(4\pi f_c t)]\cos\psi - \frac{m(t)}{2}\sin(4\pi f_c t)\sin\psi$$

Thus a more general resultant signal is attained when we take into account the possibility of frequency error and phase shift in the local oscillator that provides the receiver's version of $c(t)$. After lowpass filtering this signal, we achieve the following result:

$$r_{LP}(t) = \frac{m(t)}{2}cos(\psi)$$

So if $\psi = 0$ then we get the ideal case where $\cos\psi = 1$ and the output of the LPF is a scaled version of the message. If, however $\psi \neq 0$ then get above LPF output where the message is multiplied by the cosine of the error term.

Various values of ψ lead to different effects on the extracted message.

Recall: $\mathrm{r}_{\mathrm{LP}}(\mathrm{t}) = \frac{\mathrm{m(t)}}{2}\cos(\psi)$

Field Code Changed

Field Code Changed

Specific details of ψ	**Effect on** $r_{LP}(t)$
$\psi = 2\pi\Delta ft + \varphi$	$r_{LP}(t) = \frac{m(t)}{2}\cos(2\pi\Delta ft + \varphi)$ Message will be scaled and will be distorted
$\Delta f = 0$, $\psi = \phi =$ constant	$r(t) = k\frac{m(t)}{2}$ Where "k" is a scaling factor equal to $\cos\varphi$ (message can be recovered)
$\varphi = 0$, $\psi = 2\pi\Delta ft$	$r(t) = \frac{m(t)}{2}\cos(2\pi\Delta ft)$ Message will include distortion
$\psi = 0$	$r_{LP}(t) = \frac{m(t)}{2}$ Exact message is readily recoverable

The distorted signals can be interpreted by considering the frequency error cosine term $\cos 2\pi\Delta ft$ to be a sort of low frequency "carrier", where typically Δf =100 Hz
Assuming $m(t) = \cos 2\pi f_m t$, with typical $f_m = 1000$ Hz the receiver output

$$r(t) = m(t)\cos 2\pi\Delta ft = \cos 2\pi f_m t \cos 2\pi\Delta ft$$
$$= 0.5\cos 2\pi(f_m + \Delta f)t + 0.5\cos 2\pi(f_m - \Delta f)t$$

has outputs with frequencies at $f_m + \Delta f$ and $f_m - \Delta f$, so the message at frequency f_m now appears in two places, at $f_m - \Delta f$ and $\Delta f + f_m$. For our example, with frequency error 100 Hz, the message at 1000 Hz now appears at 900 and 1100 Hz.

We may extract the true message frequency only if we know the exact carrier frequency. Typical error ratios compared to carrier frequency ($\Delta f / f_c$) are as follows: a factor of 10^{-4} for cheap equipment, 10^{-6} for a good crystal oscillator, 10^{-11} for a Rubidium oscillator, and approaching 10^{-13} with GPS discipline (control).

Later we will see a method whereby a DSB-SC receiver can find the frequency error and eliminate it.

3.3.2 Review of I-Q receivers

Here we review the general I-Q receiver configuration that may be implemented for all types of signals The receiver for every type (for AM, DSB-SC, SSB) consists of a standard configuration coupled with a "map" that is specific to the transmission type.

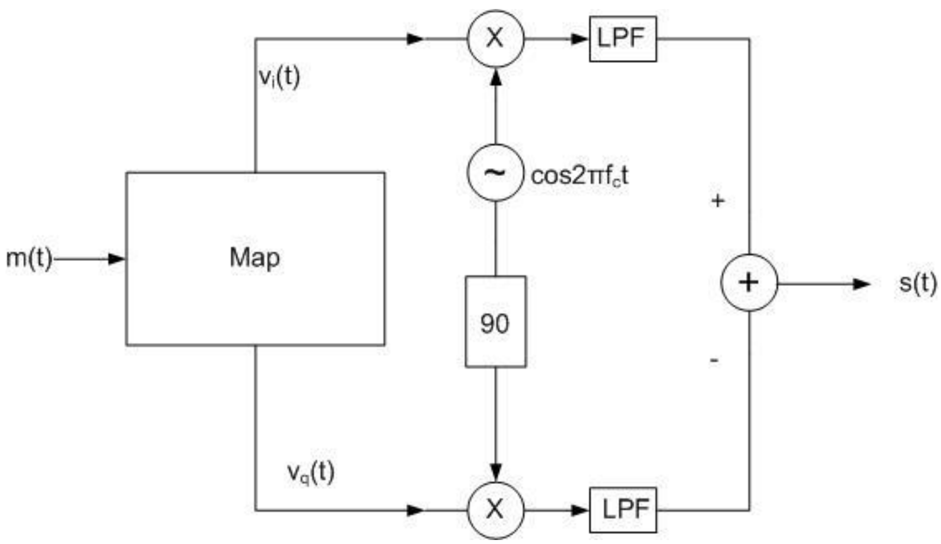
$$s(t) = a(t)\cos[2\pi f_c t + \psi(t)]$$
$$s(t) = a(t)[\cos(2\pi f_c t)\cos\psi(t) - \sin(2\pi f_c t)\sin\psi(t)]$$
$$s(t) = a(t)\cos\psi(t) cos 2\pi f_c t - a(t)\sin\psi(t)\sin 2\pi f_c t$$
$$s(t) = v_i(t) cos 2\pi f_c t - v_q(t)\sin 2\pi f_c t$$

where $a(t)$ is the amplitude and $\psi(t)$ is the phase. The coefficients of the carrier waves in the transmitter are referred to as $v_i(t)$ and $v_q(t)$ respectively, where

$$v_i(t) = a(t)\cos\psi(t)$$
$$v_q(t) = a(t)\sin\psi(t)$$

These signals are formed by the map and used to create $s(t)$ at the IQ Transmitter.

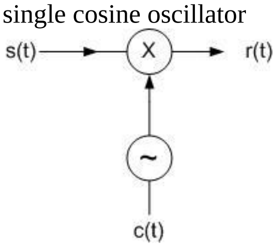
$$s(t) = v_i(t)\cos(2\pi f_c t) - v_q(t)\sin(2\pi f_c t)$$



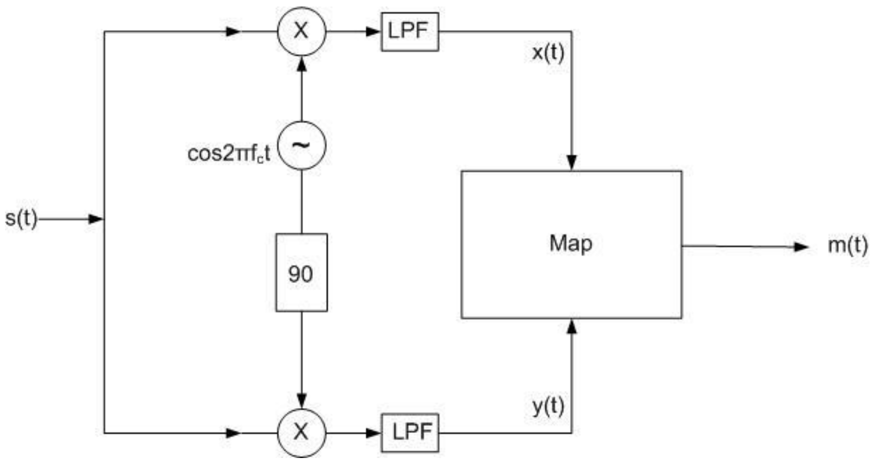
General IQ Transmitter

The IQ Receiver multiplies the incoming signal $s(t)$ by two versions of the carrier wave functions: $\cos(2\pi f_c t)$ and the 90 degree phase-shifted version $\sin(2\pi f_c t)$.

Recall the DSB-SC receiver uses only one version of the carrier wave and thus has a single cosine oscillator



The general I-Q receiver has two oscillators, cos and sin



General IQ Receiver

3.3.3 Review of AM, DSB, SSB receivers in IQ format

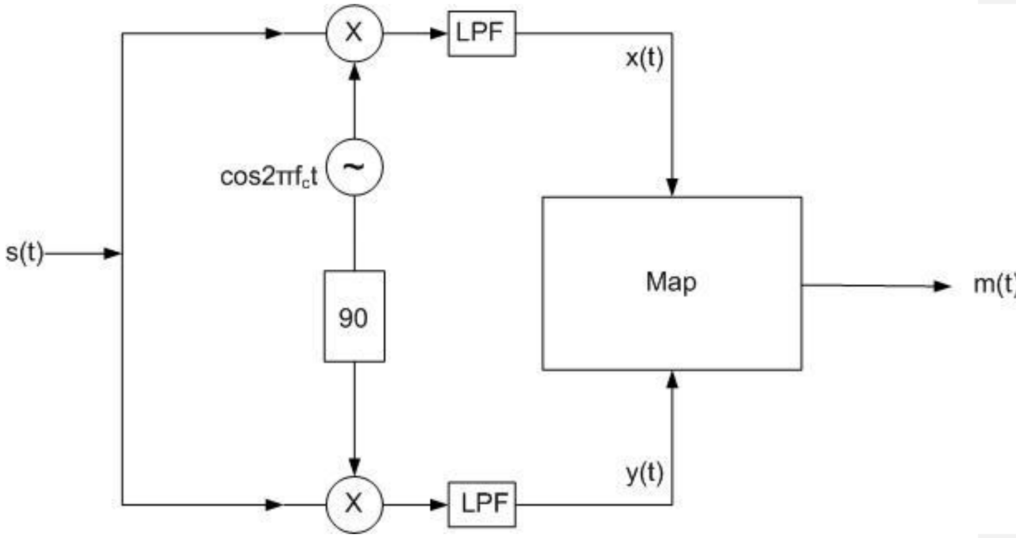
The receiver must determine a best guess of the information stored in the signals $v_i(t)$ and $v_q(t)$. The generic IQ receiver architecture produces $x(t)$ and $y(t)$ by multiplying the received signal $s(t)$ by local carriers $\cos 2\pi f_c$ and $\sin 2\pi f_c t$ The algebra describing the production of these signals assuming no frequency error appears below. We will use the identities

$$\cos\alpha\cos\beta = [\cos(\alpha-\beta)+\cos(\alpha+\beta)]/2$$

$$\cos^2\alpha = [1+\cos(2\alpha)]/2$$

$$\sin\alpha\sin\beta = [\cos(\alpha-\beta)-\cos(\alpha+\beta)]/2$$

$$\sin\alpha\cos\beta = [\sin(\alpha-\beta)+\sin(\alpha+\beta)]/2$$



$x(t)$: Before the LPF we have,

$$s(t)\cos 2\pi f_c t = [v_i(t)\cos(2\pi f_c t) - v_q(t)\sin(2\pi f_c t)]\cos(2\pi f_c t)$$

$$= v_i(t)\cos^2(2\pi f_c t) - v_q(t)\sin(2\pi f_c t)\cos(2\pi f_c t)$$

$$= \frac{v_i(t)}{2}[1+\cos(2\pi[2f_c]t)] - \frac{v_q(t)}{2}\sin(2\pi[2f_c]t)$$

$$= \frac{v_i(t)}{2} + \frac{v_i(t)}{2}\cos(2\pi[2f_c]t) - \frac{v_q(t)}{2}\sin(2\pi[2f_c]t)$$

Recall $v_i(t)$ and $v_q(t)$ are formed from the message frequency only, so of their frequency components are at or near f_m. The expression $v_i(t)\cos[2\pi(2f_c)t]$ results in two terms at the sum and difference frequencies ($2f_c \pm f_m$), both of which are near double the carrier frequency. The same can be said of $v_q(t)\sin[2\pi(2f_c)t]$. Thus after the LPF is applied to $s(t)\cos 2\pi f_c t$ the only remaining term will be $x(t) = \dfrac{v_i(t)}{2}$.

$y(t)$: Similarly, before the LPF we have,

$$s(t)\sin 2\pi f_c t = [v_i(t)\cos(2\pi f_c t) - v_q(t)\sin(2\pi f_c t)]\sin(2\pi f_c t)$$

$$= v_i(t)\sin(2\pi f_c t)\cos(2\pi f_c t) - v_q(t)\sin^2(2\pi f_c t)$$

$$= \frac{v_i(t)}{2}\sin(2\pi[2f_c]t) - \frac{v_q(t)}{2} + \frac{v_q(t)}{2}\cos(2\pi[2f_c]t)$$

And after the LPF we end up with $y(t) = \dfrac{v_q(t)}{2}$

We often write that the IQ receiver output $x(t) = v_i(t)$ and $y(t) = v_q(t)$, thus neglecting the scaling factor ½, since this scaling factor in real hardware will depend on the gains of the amplifiers.

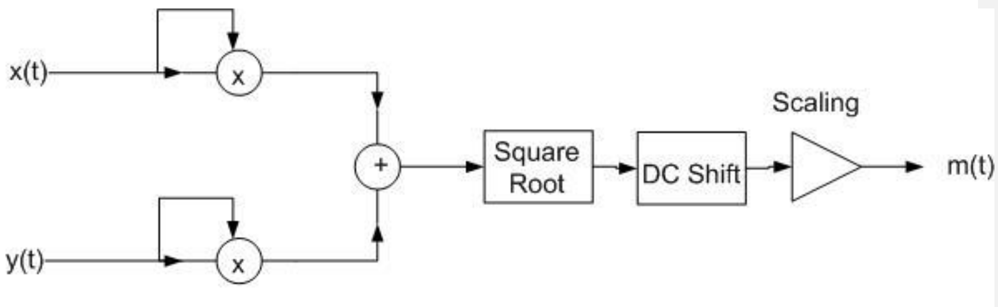
The receiver will still work in the case where $\psi = 2\pi\Delta ft + \varphi \neq 0$ as long as the oscillator that produces the signals $\cos 2\pi f_c t$ and $\sin 2\pi f_c t$ is corrected such that its frequency mirrors the error (a voltage controlled oscillator, VCO, outputs at frequency $2\pi(f_c + \Delta f)t + \varphi$ so that the apparent carrier frequency received is the same as that which is used in demodulation).
When we include the possibility of frequency error so to represent the most general case, the incoming signal can be represented by
$s(t) = v_i(t)\cos(2\pi[f_c + \Delta f]t + \phi) - v_q(t)\sin(2\pi[f_c + \Delta f]t + \phi)$.

3.4 Mapping from IQ receiver output to message signal

Each modulation method (AM, DSB-SC, SSB, etc) requires its own unique map that serves to extract the message $m(t)$ from $x(t)$ and $y(t)$.
The maps for receiving $m(t)$ from these signals for different transmission types are outlined below.

3.4.1 AM receiver.

Recall for AM
$x(t) = v_i(t) = a(t)\cos\psi(t) = A_c[1 + k_a m(t)]\cos\psi(t)$
$y(t) = v_q(t) = a(t)\sin\psi(t) = A_c[1 + k_a m(t)]\sin\psi(t)$
The AM receiver map extracts $m(t)$ by first obtaining $a(t) = A_c[1 + k_a m(t)]$, removing the DC component and scaling the signal by $1/k_a$. $a(t)$ is obtained from
$a(t) = \sqrt{x^2(t) + y^2(t)}$ since $\sqrt{x^2(t) + y^2(t)} = \sqrt{a^2(t)cos^2\psi(t) + a^2(t)sin^2\psi(t)} = a(t)$
regardless of the value of $\psi(t) = 2\pi\Delta ft + \varphi$.



Map to receive AM

Thus an AM receiver will function even with a frequency error Δf

3.4.2 DSB-SC receiver with frequency error correction

Recall for DSB-SC, $v_i(t) = m(t), v_q(t) = 0$ For the case where $m(t) = A_c cos 2\pi f_c t$
$v_i(t) = A_c \cos(2\pi f_m t)$ and $v_q(t) = 0$.
The demodulating carrier frequency must be controlled to correct for the possibility of frequency error Δf such that $\psi = 2\pi \Delta f t$. With a frequency error Δf, the incoming transmitted wave is of the following form:
$s(t) = v_i(t)\cos(2\pi[f_c + \Delta f]t) - v_q(t)\sin(2\pi[f_c + \Delta f]t)$ For DSB-SC with a single tone message $m(t) = \cos(2\pi f_m t)$ we have $v_i(t) = A_c \cos(2\pi f_m t)$ and $v_q(t) = 0$ (see DSB-SC transmitter above). Thus

$$s(t) = v_i(t)\cos(2\pi f_c t + 2\pi \Delta f t)$$
$$= A_c \cos 2\pi f_m t \cos 2\pi (f_c + \Delta f)t$$
$$= A_c \cos 2\pi[(f_c + \Delta f) + f_m]t + \cos 2\pi[(f_c + \Delta f) - f_m]t$$

To find $x(t)$: Before the LPF,

$$s(t)\cos 2\pi f_c t = A_c \cos(2\pi f_m t)\cos 2\pi(f_c + \Delta f)t \cos(2\pi f_c t)$$
$$= \frac{A_c}{2}[\cos 2\pi[(f_c + \Delta f) + f_m]t + \cos 2\pi[(f_c + \Delta f) - f_m]t]\cos 2\pi f_c t$$
$$= \frac{A_c}{4}[\cos(2\pi[(2f_c + \Delta f) + f_m]t) + \cos(2\pi[f_m + \Delta f]t) + \cos(2\pi[f_m - (\Delta f)]t) + \cos(2\pi[(2f_c + \Delta f) - f_m]t)]$$

After the LPF we have,

$$x(t) = \frac{A_c}{4}[\cos(2\pi[f_m + \Delta f]t) + \cos(2\pi[f_m - \Delta f]t)]$$

For a general message $m(t)$, we find $x(t) = \frac{1}{2} m(t)\cos 2\pi \Delta \mathrm{ft}$

Similarly for $y(t)$: Before the LPF,

$s(t)\sin 2\pi f_c t = A_c \cos 2\pi f_m t \cos 2\pi (f_c + \Delta f)t \sin 2\pi f_c t$

$= \frac{A_c}{2}[\cos(2\pi[f_m + (f_c + \Delta f)]t) + \cos(2\pi[f_m - (f_c + \Delta f)]t)]\sin(2\pi f_c t)]$

$= \frac{A_c}{4}[\sin(2\pi[f_m + (2f_c + \Delta f)]t) - \sin(2\pi[f_m + (\Delta f)]t) + (2\pi[f_m - (\Delta f)]t) - \sin(2\pi[f_m - (2f_c + \Delta f)]t)]$

After the LPF we have,

$y(t) = \frac{A_c}{4}[-\sin(2\pi[f_m + (\Delta f)]t) + \sin(2\pi[f_m - (\Delta f)]t)]$

For a general message $m(t)$, we find $y(t) = \frac{1}{2} m(t)\sin 2\pi \Delta \mathrm{ft}$

The received signals are distorted by the frequency offset. The DBS-SC receiver does not inherently know the frequency error but it manipulates the signals $x(t)$ and $y(t)$ to acquire it. We multiply $x(t)$ and $y(t)$ together and produce the following control signal for the VCO:

Using $\sin\alpha\cos\alpha = (1/2)\sin 2\alpha$

$e(t) = x(t)y(t)$

$= \frac{A_c^2}{16}[-\sin(2\pi[f_m + (\Delta f)]t) + \sin(2\pi[f_m - (\Delta f)]t)]\cdot$

$[\cos(2\pi[f_m + (\Delta f)]t) + \cos(2\pi[f_m - (\Delta f)]t)]$

$= \frac{A_c^2}{32}[-\sin(4\pi[f_m + \Delta f]t) - \sin(4\pi f_m t) - \sin(4\pi \Delta f t) +$

$\sin(4\pi f_m t) + \sin(-4\pi \Delta f t) + \sin(4\pi[f_m - \Delta f]t)$

$= \frac{A_c^2}{32}[-2\sin(4\pi \Delta f t)] = -\frac{A_c^2}{16}\sin(4\pi \Delta f t)$

For a general message $m(t)$

$e(t) = x(t)y(t) = \frac{1}{4} m^2(t)\cos 2\pi \Delta \mathrm{ft} \sin 2\pi \Delta \mathrm{ft}$

$= \frac{1}{8} m^2(t)\sin 4\pi \Delta \mathrm{ft}$

Typical values of f_m are in the range 300-3,000 Hz for voice, whereas a typical value for Δf is 100 Hz or less. $e(t)$ is low pass filtered with a cutoff frequency less than 300 Hz so that all frequency components f_m of $m(t)$ are filtered out. Thus after the low pass filter, $e_{LPF}(t) = \frac{1}{8} < m^2(t) > \sin 4\pi \Delta \mathrm{ft}$ where $< m^2(t) >$ is a DC value representing the average power in $m(t)$. Thus even without knowing the error frequency Δf, we can produced a sinusoid, directly dependant on Δf. This allows us to control the frequency for the local oscillator at the receiver to replicate the incoming carrier frequency. In the special case

$$m(t) = A_c \cos 2\pi f_m t$$

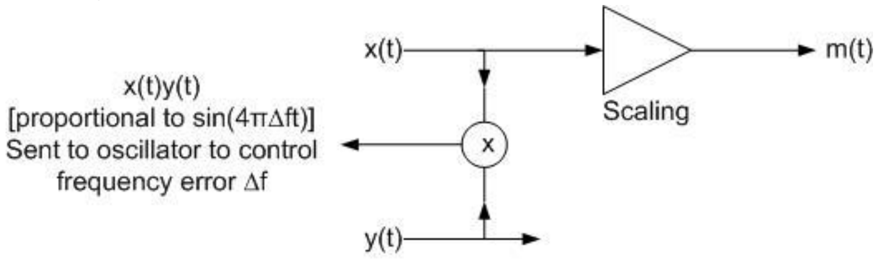
$$m^2(t) = \frac{A_c^2}{2} + \frac{A_c^2}{2} \cos 4\pi f_m t$$

where

$$< m^2(t) >= \frac{A_c^2}{2}$$

$$e_{LPF}(t) = \frac{1}{8} < m^2(t) > \sin 4\pi \Delta \text{ft} = \frac{A_c^2}{16} sin 4\pi \Delta \text{t}$$

The above analysis can be repeated using a general message $m(t)$ and a general offset $\psi = 2\pi \Delta f t + \varphi \neq 0$



Map to receive DSB-SC

3.4.3 SSB-SC receiver (analog)

We consider two methods of receiving an SSB-SC signal.

In Method 1, we multiply the real SSB signal by $\cos 2\pi f_c t$ and low pass filter to get $m(t)$ This method is used in analog receivers.

Exercise: show the mathematical steps to obtain $m(t)$ from $s(t)$ using this method.\

To demodulate SSB-SC, the receiver is required to know the exact carrier frequency for proper demodulation.

For an SSB-SC signal with $s(t) = m(t) \cos 2\pi f_c t \mp \hat{m}(t) sin 2\pi f_c t$ and considering $m(t) = A_c cos 2\pi f_m t$, we can write

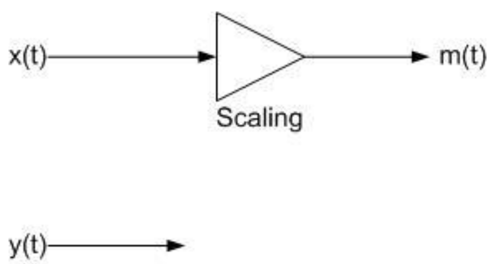
for lower sideband, LSB: $s(t) = A_c \cos(2\pi [f_c - f_m] t)$

for upper sideband, USB: $s(t) = A_c \cos(2\pi [f_c + f_m] t)$).

For USB, the IQ receiver output

$$x(t) = A_c \cos 2\pi(f_c + f_m)t \cos 2\pi f_c t$$
$$= \frac{A_c}{2}[\cos 2\pi(2f_c + f_m)t + \cos 2\pi f_m t]$$
$$y(t) = A_c \cos 2\pi(f_c + f_m t \sin 2\pi f_c t$$
$$= \frac{A_c}{2}[\sin 2\pi(2f_c + f_m)t - \sin 2\pi f_m t]$$

After LPF, $x(t) = \frac{A_c}{2}\cos 2\pi f_m t$ and $y(t) = \frac{A_c}{2}\sin 2\pi f_m t$. In both cases, an LPF leaves us with the message (cosine term) and a phase shifted version (sine term). We simply choose the signal in $x(t)$ and further demodulation only serves to scale the amplitude of the message.



Map to receive SSB-SC

3.4.4 SSB-SC receiver using Weaver Demodulator description

The second method of receiving SSB-SC uses the Weaver demodulator which is analogous to the Weaver modulator discussed earlier.

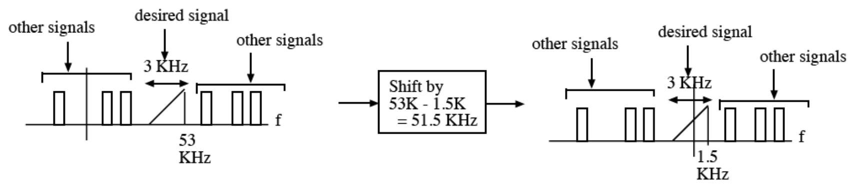
The Weaver demodulator operates in several steps, shown below in complex notation for the USB case. We assume the message occupies a bandwidth of 0-3000 Hz.

$$\begin{array}{ccccccc} s(t) \rightarrow & \otimes & \Rightarrow LPF(f_1) \Rightarrow & & \otimes & \Rightarrow\rightarrow m(t) \\ & \Uparrow \; \tilde{s}_1(t) & & \tilde{s}_2(t) & \Uparrow & \tilde{s}_3(t) \\ & e^{-j2\pi(f_0+f_1)t} & & & e^{j2\pi f_1 t} & \end{array}$$

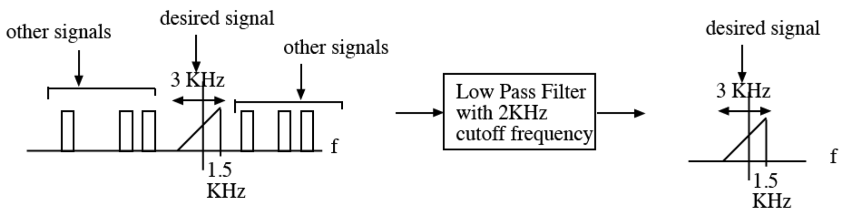
- downconvert the real USB signal $s(t) = m(t)\cos 2\pi f_c t - \hat{m}(t) sin 2\pi f_c t$ to $\tilde{s}_1(t)$ using a complex local oscillator $e^{j2\pi(f_c+f_1)t}$ with a frequency offset $f_1 = 1.5$ KHz relative to $f_0 = f_c = 53$ KHz as shown in the figure below. We choose $f_1 = B/2$

to be a frequency in the approximate middle of the message bandwidth, e.g. 1500 Hz for a 0-3000 Hz voice signal.

-



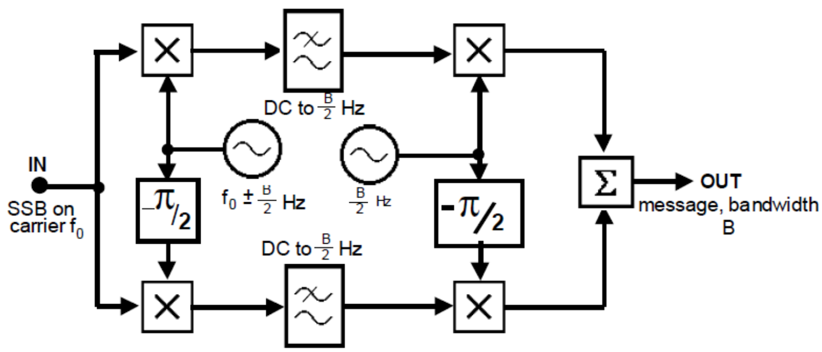
-
- Low pass filter the result to eliminate adjacent undesired signals. The resulting signal $\tilde{s}_2(t)$ is a complex signal with asymmetric spectrum in the frequency range $-B/2$ to $B/2$ or -1.5 KHz to +1.5 KHz.



-
- Do a frequency shift of $\tilde{s}_2(t)$ with a complex local oscillator $e^{j2\pi f_1 t}$ where $f_1 = B/2 = 1.5\,\text{KHz}$ as defined above, resulting in $\tilde{s}_3(t)$ in the frequency range 0 to B or 0 – 3000 Hz. $\tilde{s}_3(t)$ is an analytic signal with positive frequencies only.
- Take the real part of $\tilde{s}_3(t)$ to obtain the real signal $m(t)$ with both positive and negative frequencies that can be listened to or decoded correctly.
- The net result of the two frequency shifts and LPF in the Weaveer demodulator is that the suppressed carrier frequency in $s(t)$ is shifted to zero frequency, the message is recovered and all other unwanted signals are filtered out.

The Weaver demodulator also works for Morse code or other digital signals that are processed by ear or other decoder that operates on a real (not complex) signal. In this case, the bandwidth B may be set to an appropriate value (e.g. 50 Hz for Morse code) and the frequency offset f_1 is set to a tone (pitch) that is pleasant to the ear (e.g. 400 Hz).

The operations of a Weaver demodulator can also be shown in real notation as shown in the figure below.



USRP downmixer GNURadio mixer

The complex version is repeated below for convenient comparison with the real version

$$
\begin{array}{l}
s(t) \rightarrow \otimes \Rightarrow LPF(f_1) \Rightarrow \otimes \Rightarrow\rightarrow m(t) \\
\qquad\quad \Uparrow \ \tilde{s}_1(t) \qquad \tilde{s}_2(t) \ \ \Uparrow \ \ \tilde{s}_3(t) \\
\qquad e^{-j2\pi(f_0+f_1)t} \qquad\qquad e^{j2\pi f_1 t}
\end{array}
$$

3.4.5 SSB-SC Weaver demodulator operation in complex notation

We consider two cases for the input signal, real and complex

The mathematics below is for a real input signal that arises when the signal is taken directly from antenna and preamp and (perhaps) an analog real downconverter stage (cos oscillator only)

$$
\begin{array}{l}
s(t) = \operatorname{Re}\{\tilde{s}(t)e^{j2\pi f_c t}\} = m(t)\cos 2\pi f_c t - \hat{m}(t) sin 2\pi f_c t \\
\text{where } \tilde{s}(t) = m(t) + j\hat{m}(t) \\
\text{after first mixer} \\
\tilde{s}_1(t) = [m(t)\cos 2\pi f_c t - \hat{m}(t) sin 2\pi f_c t]e^{-j2\pi(f_c+f_1)t} \\
= 0.5[m(t)e^{j2\pi f_c t} + m(t)e^{-j2\pi f_c t} + j\hat{m}(t)e^{j2\pi f_c t} - j\hat{m}(t)e^{-j2\pi f_c t}]e^{-j2\pi(f_c+f_1)t} \\
= 0.5[m(t)e^{-2\pi f_1 t} + m(t)e^{-j4\pi(f_c+f_1)t} + j\hat{m}(t)e^{-j2\pi f_1 t} - j\hat{m}(t)e^{-j4\pi(f_c+f_1)t}] \\
\text{afterLPF must be analytic signal shifted} \\
\tilde{s}_2(t) = m(t)e^{-2\pi f_1 t} + j\hat{m}(t)e^{-j2\pi f_1 t} = [m(t) + j\hat{m}(t)]e^{-j2\pi f_1 t} \\
\text{after second mixer} \\
\tilde{s}_3(t) = \tilde{s}_2(t)e^{j2\pi f_1 t} = \tilde{s}(t) = m(t) + j\hat{m}(t) \\
m(t) = \operatorname{Re}\{\tilde{s}(t)\}
\end{array}
$$

The mathematics below is for a complex input that arises when a complex downconversion (e.g. cos and sin oscillators in an analog mixer with two outputs I and Q) has taken place ahead of the USB receiver complex input

$$\tilde{s}(t)e^{j2\pi f_c t} = m(t)\cos 2\pi f_c t - \hat{m}(t) sin 2\pi f_c t + j[\hat{m}(t)\cos 2\pi f_c t + m(t) sin 2\pi f_c t]$$

where $\tilde{s}(t) = m(t) + j\hat{m}(t)$

after first mixer

$$\tilde{s}_1(t) = \tilde{s}(t)e^{j2\pi f_c t}e^{-j2\pi(f_c+f_1)t} = \tilde{s}(t)e^{-j2\pi f_1 t}$$

after LPF, eliminate other unwanted signals

$$\tilde{s}_2(t) = \tilde{s}_1(t)$$

after second mixer

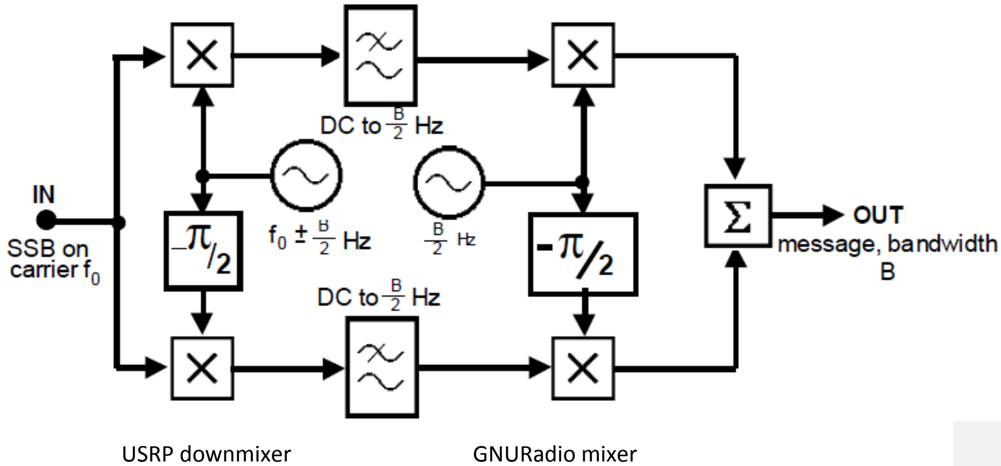
$$\tilde{s}_3(t) = \tilde{s}_2(t)e^{j2\pi f_1 t} = \tilde{s}(t)e^{-j2\pi f_1 t}e^{j2\pi f_1 t} = \tilde{s}(t) = m(t) + j\hat{m}(t)$$

$$m(t) = \mathrm{Re}\{\tilde{s}(t)\}$$

3.4.6 SSB-SC Weaver demodulator operation in real notation

The real version of the Weaver demodulator can be built entirely in analog, or entirely in digital. It may also be split, with the first cos/sin mixer in analog (as done in the USRP daugtherboard) and the second mixer digital (implemented in GNURadio software).

In the figure, for USB, the first oscillator is $f_0 + B/2$



The mathematics for the Weaver demodulator in real notation is done in the example below, assuming that the message is the sum of cos waves at 500 and 2000 Hz. This example should give an appreciation of the value of complex signals to simply both the concepts and the algebra.

$m(t) = \cos 2\pi f_2 t + \cos 2\pi f_3 t = \cos 2\pi 500t + \cos 2\pi 2000t$

Recall

$\cos\alpha \cos\beta = [\cos(\alpha-\beta) + \cos(\alpha+\beta)]/2$

$\sin\alpha \cos\beta = [\sin(\alpha-\beta) + \sin(\alpha+\beta)]/2$

$\sin\alpha \sin\beta = [\cos(\alpha-\beta) - \cos(\alpha+\beta)]/2$

$\cos\alpha \sin\beta = [-\sin(\alpha-\beta) + \sin(\alpha+\beta)]/2$

$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$

$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$

Message

$m(t) = \cos 2\pi f_2 t + \cos 2\pi f_3 t = \cos 2\pi 500t + \cos 2\pi 2000t$

Real SSB signal (upper sideband)

$s(t) = m(t)\cos 2\pi f_c t - \hat{m}(t) sin 2\pi f_c t$

$= [\cos 2\pi f_2 t + \cos 2\pi f_3 t]\cos 2\pi f_c t - [\sin 2\pi f_2 t + \sin 2\pi f_3 t]\sin 2\pi f_c t$

$= \cos 2\pi f_2 t \cos 2\pi f_c t - \sin 2\pi f_2 t \sin 2\pi f_c t + \cos 2\pi f_3 t \cos 2\pi f_c t - \sin 2\pi f_3 t \sin 2\pi f_c t$

$= \cos 2\pi (f_2 + f_c)t + \cos 2\pi (f_3 + f_c)t$

Writing $f_0 = f_c$

$s(t) = \cos 2\pi (f_2 + f_0)t + \cos 2\pi (f_3 + f_0)t$

First oscillator is at $f_0 + B_2$ for upper sideband

Upper branch after first mixer and low pass filter and before second mixer

Writing $B_2 = B/2 = f_1$ and recall $\cos\alpha\cos\beta = [\cos(\alpha-\beta) + \cos(\alpha+\beta)]/2$

$s(t)\cos 2\pi (f_0 + B_2)t = 0.5[\cos 2\pi (f_2 + f_0)t + \cos 2\pi (f_3 + f_0)t]\cos 2\pi (f_0 + B_2)t$

$= 0.5[\cos 2\pi (f_2 - B_2)t + \cos 2\pi (f_3 - B_2)t]$

$= 0.5[\cos 2\pi (500 - 1500)t + \cos 2\pi (2000 - 1500)t]$

$= 0.5[\cos 2\pi (-1000)t + \cos 2\pi (500)t]$

Upper branch after second mixer

$0.5[\cos 2\pi (f_2 - B_2)t + \cos 2\pi (f_3 - B_2)t]\cos 2\pi B_2 t$

$= 0.25[\cos 2\pi f_2 t + \cos 2\pi (f_2 - 2B_2)t + \cos 2\pi f_3 t + \cos 2\pi (f_3 - 2B_2)t]$

$= 0.25[\cos 2\pi (500)t + \cos 2\pi (500 - 3000)t + \cos 2\pi (2000)t + \cos 2\pi (2000 - 3000)t]$

$= 0.25[\cos 2\pi (500)t + \cos 2\pi (-2500)t + \cos 2\pi (2000)t + \cos 2\pi (-1000)t]$

Lower branch after first mixer and low pass filter and before second mixer

Recall $\cos\alpha\sin\beta = [-\sin(\alpha-\beta)+\sin(\alpha+\beta)]/2$

$$
\begin{aligned}
& s(t)[-\sin 2\pi(f_0+B_2)t] = -[\cos 2\pi(f_2+f_0)t + \cos 2\pi(f_3+f_0)t]\sin 2\pi(f_0+B_2)t \\
& = -0.5[\cos 2\pi(f_2+f_0)t\sin 2\pi(f_0+B_2)t - \cos 2\pi(f_3+f_0)t\sin 2\pi(f_0+B_2)t] \\
& = 0.5[\sin 2\pi(f_2-B_2)t + \sin 2\pi(f_3-B_2)t] \\
& = 0.5[\sin 2\pi(500-1500)t + \sin 2\pi(2000-1500)t] \\
& = 0.5[\sin 2\pi(-1000)t + \sin 2\pi(500)t]
\end{aligned}
$$

Lower branch after second mixer

Recall $\sin\alpha\sin\beta = [\cos(\alpha-\beta)-\cos(\alpha+\beta)]/2$

$$
\begin{aligned}
& 0.5[\sin 2\pi(f_2-B_2)t + \sin 2\pi(f_3-B_2)t][-\sin 2\pi B_2 t] \\
& = 0.25[-\cos 2\pi(f_2-2B_2)t + \cos 2\pi f_2 t - \cos 2\pi(f_3-2B_2)t + \cos 2\pi f_3 t] \\
& = 0.25[-\cos 2\pi(500-3000)t + \cos 2\pi(500)t - \cos 2\pi(2000-3000)t + \cos 2\pi(2000)t] \\
& = 0.25[-\cos 2\pi(-2500)t + \cos 2\pi(500)t - \cos 2\pi(-1000)t + \cos 2\pi(2000)t]
\end{aligned}
$$

Sum of second mixer outputs

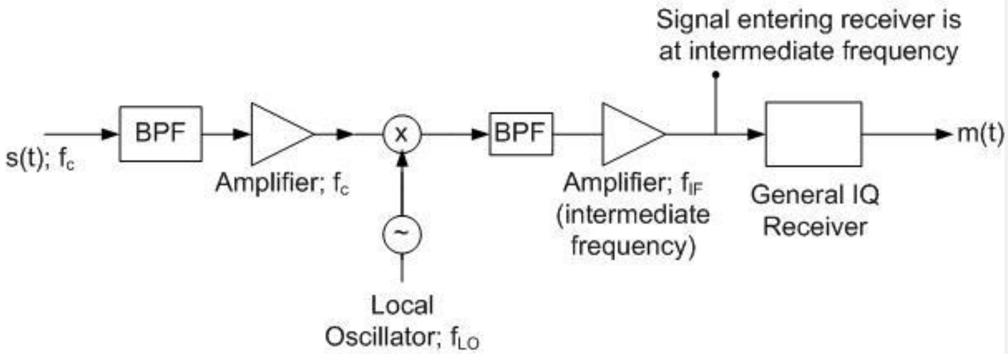
$$
0.5[\cos 2\pi f_2 t + \cos 2\pi f_3 t] = 0.5[\cos 2\pi(500)t + cos 2\pi(2000)t] = 0.5m(t)
$$

4 Super-heterodyne Receiver

Before a transmission signal is inputted into an IQ receiver, it is often advantageous to apply amplification as the typical amplitude range at this point is on the order of microvolts. A far more practical signal level to work with is one on the order of Volts and such a gain is usually achieved through a series of op amps. Such a large gain requirement around a single carrier frequency can raise practical issues such as the occurrence of feedback due to the high nature of the frequencies needing to be amplified.

A super-heterodyne (or colloquially: superhet) receiver offers a means of overcoming such issues by shifting the carrier (usually down in magnitude) and effectively splitting the gain requirements over multiple frequencies: the original carrier and what is referred to as the intermediate frequency (f_{IF}). This action helps to prevent feedback which may lead to inaccuracies in the demodulated message.

Another advantage of using the superhet design is that it allows the receiver to shift any carrier frequency to an industry standard. This allows the components of the standardized IQ receiver to be mass produced at a very low cost. Amplifiers that support higher frequencies tend to have a narrower bandwidth as well so by shifting the carrier we also remove this limitation. The architecture of the super-heterodyne appears below.



Super-heterodyne Receiver

The receiver is tuned to receive different carrier frequencies by changing the frequency of the so-called Local Oscillator (as distinct from the Remote oscillator in the transmitter some distance away from the receiver). The tuning knob on a radio receiver controls the frequency of the Local Oscillator. A common intermediate frequency used for AM receivers is 455kHz. Once f_{IF} is selected, we may decide how to tune the local oscillator so to produce such a carrier:

$f_{LO} = f_c \pm f_{IF}$;

With the local oscillator tuned to this frequency we get the following pair:
$f_c + (f_c \pm f_{IF}) = 2f_c \pm f_{IF}$ and $f_c - (f_c \pm f_{IF}) = \mp f_{IF}$

We can easily bandpass filter the high frequency term out meaning we have successfully shifted our carrier to a desired intermediate frequency.

An issue with the superhet which is important to note is the existence of an "image" frequency (f_{IM}). If a signal exists at this frequency and proper filtering is not implemented then it will be shifted to the intermediate frequency, along with the signal at the desired f_c, and will cause direct interference.

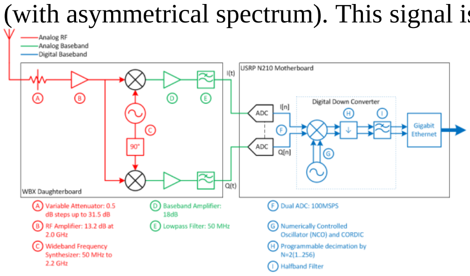
If $f_{LO} = f_c \pm f_{IF}$ then $f_{IM} = f_{LO} \pm f_{IF}$

It is often best to consider specific numbers rather than formulas with plus/minus signs. For example, consider f_RF = f_c = 1070 KHz and f_IF = 455 KHz (we often refer to the carrier frequency f_c as the radio (RF) frequency f_RF

Then we can choose f_LO = 1070-455=615 KHz, so that the incoming signal at 1070 KHz mixes with the LO at 615 KHz to yield an output at the 455 KHz IF. Thus by setting the LO to 615 KHz we have "tuned in" the signal at 1070 KHz.
The image frequency will be 615-455=160 KHz, since an incoming (image) signal at 160 KHz can also mix with the same LO at 615 KHz to yield an output at the 455 KHz IF.

We can also choose f_LO = 1070+455=1525 KHz, so that the incoming signal at 1070 KHz mixes with the LO at 1525 KHz to yield an output at the 455 KHz IF. . Thus by setting the LO to 1525 KHz we have "tuned in" the signal at 1070 KHz.
In this case, the image frequency will be 1525+455=1980 KHz, since an incoming (image) signal at 1980 KHz can also mix with the same LO at 1525 KHz to yield an output at the 455 KHz IF

The superheterodyne principle can also be applied to complex signals. For example, in the USRP, the daughterboard is a complex local oscillator that downconverts a slice of radio frequencies centered around f_LO to a zero frequency IF complex baseband signal (with asymmetrical spectrum). This signal is sampled by the main USRP board.



5 Frequency Modulation

5.1 Overview

In this chapter, we describe angle modulation, which includes both frequency and phase modulation. In both cases, amplitude is kept constant.

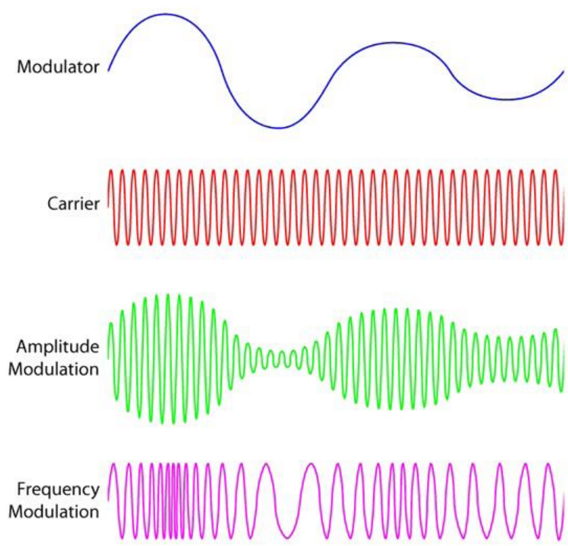
The phase angle of the carrier wave varies with the message signal.
For phase modulation, the phase angle is linearly related to the message
For frequency modulation (FM), the derivative of the phase angle is linearly related to the message.

FM and PM signals have a constant amplitude or *constant envelope*, thus enabling the use of power-efficient nonlinear power amplifiers.

FM signals are more resistant to noise than AM, but at the cost of larger bandwidth.

FM is widely used for radio broadcasting, analog TV audio and public safety two-way radios.

The figure below compares AM and FM



5.2 FM with General Message

To derive the equation for an FM wave, we have 3 starting points:

0. Recall that idea of AM is that the instantaneous amplitude $a(t)$ of a carrier wave is varied linearly with the baseband message signal $m(t)$ around a constant bias value A_c so that $a(t) = A_c + A_c k_a m(t)$

1. By analogy, the idea of frequency modulation is that the instantaneous frequency $f_i(t)$ of a carrier wave is varied linearly with the baseband message signal $m(t)$ around a constant value f_c Thus for a general message $m(t)$ the idea of FM is:

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

   where the constant k_f represents the frequency sensitivity of the modulator, expressed in hertz per volt.

2. Recall that a general signal is written:

   $$s(t) = a(t)\cos\theta(t) = a(t)cos[2\pi f_c t + \phi(t)]$$

   where $\theta(t) = 2\pi f_c t + \phi(t)$ has a linear variation at rate f_c and a time varying part $\phi(t)$

3. Recall the instantaneous frequency of a signal is:

   $$f_i(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt} = f_c + \frac{1}{2\pi}\frac{d\phi(t)}{dt}$$

Combining these 3 starting points, we can write:

$$\theta(t) = 2\pi\int_0^t f_i(\alpha)d\alpha = 2\pi\int_0^t [f_c + k_f m(\alpha)]d\alpha$$
$$= 2\pi f_c t + k_f \int_0^t m(\alpha)d\alpha$$
$$s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\alpha)d\alpha\right]$$

The FM signal $s(t)$ can be written in standard IQ format:

$$s(t) = i(t)\cos 2\pi f_c t - q(t) sin 2\pi f_c t$$
$$s(t) = \mathrm{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\} =$$
$$\mathrm{Re}\{[I(t) + jQ(t)][\cos 2\pi f_c t + j sin 2\pi f_c t]\}$$
$$= I(t)\cos 2\pi f_c t - Q(t) sin 2\pi f_c t$$
$$= a(t)\cos[2\pi f_c t + \phi(t)]$$

Thus for FM:

$$a(t) = A_c$$

$$\phi(t) = 2\pi k_f \int_0^t m(\alpha)d\alpha$$

$$I(t) = A_c \cos 2\pi k_f \int_0^t m(\alpha)d\alpha$$

$$Q(t) = A_c \sin 2\pi k_f \int_0^t m(\alpha)d\alpha$$

5.3 FM with Sinusoidal Message

Now consider a sinusoidal modulating wave defined by
$m(t) = A_m \cos(2\pi f_m t)$

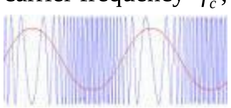
We re-derive the equation for the FM wave with the same 3 starting points.

1. The idea of FM is that the instantaneous frequency of the resulting FM wave equals:

   $$f_i(t) = f_c + k_f A_m \cos(2\pi f_m t)$$
   $$= f_c + \Delta f \cos(2\pi f_m t)$$

   where $\Delta f = k_f A_m$.

   Thus the message causes the instantaneous frequency to vary above and below the carrier frequency f_c, from $f_c - \Delta f$ to $f_c + \Delta f$.

   The quantity Δf is called the <u>frequency deviation</u>, since the instantaneous frequency deviates from the carrier by that amount.

2. A general signal:

   $$s(t) = a(t)\cos\theta(t)$$

3. The instantaneous frequency of a signal is:

   $$f_i(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt}$$

Combining these 3 starting points, we can write:

$$\theta(t) = 2\pi f_c t + \frac{\Delta f}{f_m}\sin(2\pi f_m t)$$
$$= 2\pi f_c t + \beta\sin(2\pi f_m t)$$

Where $\beta = \frac{k_f A_m}{f_m} = \frac{\Delta f}{f_m}$ is called the modulation index of the FM wave.

Exercise: fill in the missing steps in the above derivation.

Thus the FM wave itself is given In terms of β by:

$$s(t) = A_c\cos\left[2\pi f_c t + \beta\sin(2\pi f_m t)\right]$$

Thus for a sinusoidal modulating wave $m(t) = A_m\cos(2\pi f_m t)$ the FM wave can be written:

$$s(t) = I(t)\cos 2\pi f_c t - Q(t) sin 2\pi f_c t = a(t)\cos[2\pi f_c t + \phi(t)]$$

where:

$$a(t) = A_c$$
$$\phi(t) = \beta sin 2\pi f_m t$$
$$i(t) = A_c\cos\beta sin 2\pi f_m t$$
$$q(t) = A_c\sin\beta sin 2\pi f_m t$$

If β is small we have narrowband FM (NBFM) and if β is large (compared to one radian) we have wideband FM.

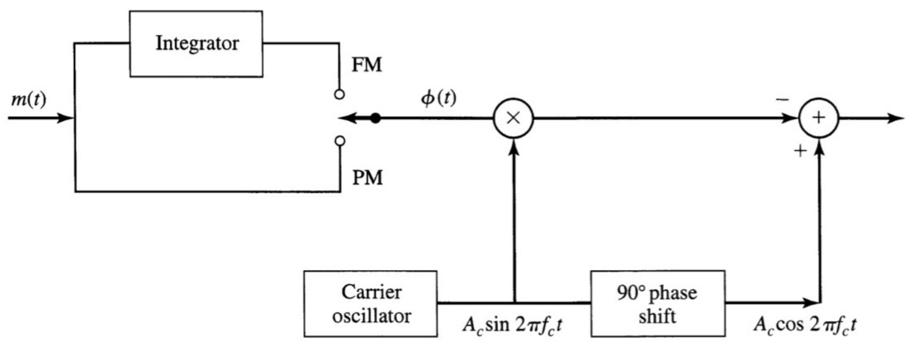
5.3.1 Narrowband FM with Sinusoidal Message

When β is small compared to one radian, i.e. $\beta << 1$, the FM wave may be approximated using $\cos x \approx 1, \sin x \approx x$ for $x << 1$ to obtain

$$s(t) \simeq A_c\cos(2\pi f_c t) - \beta A_c\sin(2\pi f_c t)\sin(2\pi f_m t)$$

Thus for $\beta << 1$

$$i(t) = A_c$$
$$q(t) = \beta A_c sin 2\pi f_m t$$



Narrow-band Frequency and Phase Modulator

Exercise: find the complex envelope $\tilde{s}(t)$ for a NBFM signal. Recall that in general

$$\tilde{s}(t) = a(t)e^{j\phi(t)}$$
$$s(t) = \mathrm{Re}\{\tilde{s}(t)e^{j2\pi f_c t}\}$$
$$a(t) = \sqrt{i^2(t) + q^2(t)}$$
$$\phi(t) = \arctan\frac{q(t)}{i(t)}$$

Compare the NBFM waveform equation with the AM equation, note the similarity and differences.

$$s(t) = A_c cos 2\pi f_c t + A_c \mu \cos 2\pi f_m t \cos 2\pi f_c t$$

Observe that NBFM signal wave requires essentially the same transmission bandwidth (i.e., $2f_m$) as an AM wave.

5.4 Power Spectrum of an FM Signal – Bessel functions

Consider the FM signal with a single tone message $m(t) = A_m \cos(2\pi f_m t)$ so that

$$s(t) = A_c \cos\left[2\pi f_c t + \beta \sin(2\pi f_m t)\right]$$
$$= \mathrm{Re}\{A_c e^{j(2\pi f_c t + \beta \sin(2\pi f_m t))}\} = \mathrm{Re}\{A_c e^{\beta \sin 2\pi f_m t} e^{j2\pi f_c t}\}$$

In what follows, we will evaluate the FM complex envelope and find that it is given in term of a special function called a Bessel function. Once we have defined this function, we continue with finding the FM power spectrum and bandwidth.

The FM complex envelope $a(t)e^{j\phi(t)} = A_c e^{j\beta \sin 2\pi f_m t}$ is periodic with period $1/f_m$ so we can write the complex envelope as a complex Fourier series with index *n*

$$A_c e^{j\beta \sin 2\pi f_m t} = A_c \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_m t}$$

The Fourier coefficients are given by the integral

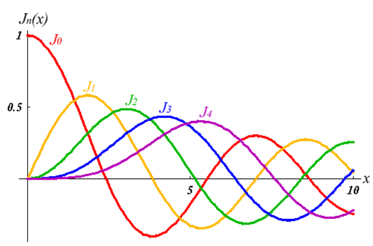
$$c_n = f_m \int_{t=0}^{1/f_m} e^{j\beta \sin 2\pi f_m t} e^{-j2\pi n f_m t} dt \triangleq J_n(\beta)$$

This integral depends on two parameters n and β, but cannot be evaluated in terms of elementary functions, so it is given a name $J_n(\beta)$ called the Bessel function of the first kind and order *n.*

By a change of variable $u = 2\pi f_m t$ we can write

$$J_n(\beta) = \frac{1}{2\pi} \int_{u=0}^{2\pi} e^{j(\beta \sin u - nu)} du$$

The function $J_n(\beta)$, the Bessel function of the first kind of order n, is plotted in the figure below for values of n from 0 to 4.



Bessel Functions

$J_n(\beta)$ looks like a damped cosine wave for $n = 0$ and damped sine waves for $n \geq 1$

The Bessel function has many properties and identities (just like cos and sin)

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

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

Also, for small values of β, we have:

$$J_0(\beta) \simeq 1, \; J_1(\beta) \simeq \frac{\beta}{2}$$

and:

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

Now that we have introduced the Bessel function, we continue to find the FM power spectrum and bandwidth.

The FM complex envelope can be written

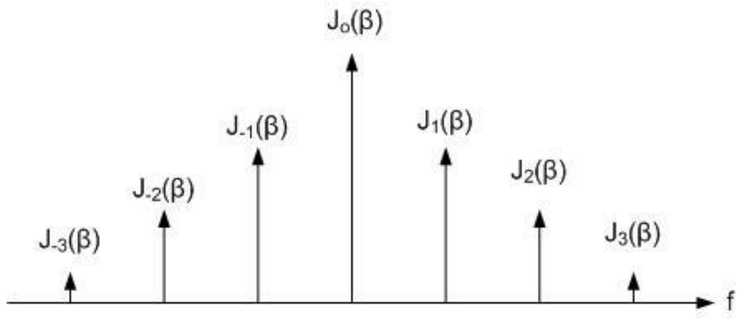
$$a(t)e^{j\phi(t)} = A_c e^{j\beta\sin 2\pi f_m t} = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) e^{j2\pi n f_m t}$$

Thus the FM signal is written

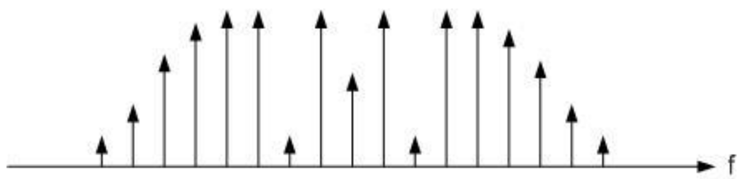
$$\begin{aligned} s(t) &= \mathrm{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\} = \mathrm{Re}\{A_c e^{\beta\sin 2\pi f_m t} e^{j2\pi f_c t}\} \\ &= \mathrm{Re}\{A_c[\sum_{n=-\infty}^{\infty} J_n(\beta) e^{j2\pi n f_m t}]\, e^{j2\pi f_c t}\} \\ &= A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \cos[2\pi(f_c + n f_m)t] \end{aligned}$$

The discrete spectrum of the FM wave is obtained as

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



Bessel Function value [$J_n(\beta)$] defining the FM spectrum depends upon β



Example of FM spectrum with relatively large β

As for the case of AM, we may ask why it is that for a sinusoidal message we see discrete sidebands at $f_c \pm n\, f_m$, since the instantaneous frequency given by $f_i(t) = f_c + \Delta f \cos(2\pi f_m t) = f_c + \beta f_m \cos(2\pi f_m t)$ may be any continuous value and is not restricted to a discrete set of values.

If we view the FM spectrum over an observation (time) interval much less than the message period $1/\,f_m$, the spectrum consists of one impulse or "spike" at $f_i(t)$ that moves back and forth in frequency from $f_c - \Delta f$ to $f_c + \Delta f$ in a continuous fashion as time progresses. If the observation interval is greater than $1/\,f_m$, then the spectrum is discrete with sidebands spaced at intervals of f_m. The derivation of the FM spectrum above uses the complex Fourier series and thus assumes an observation interval of one message period.

5.4.1 Observations about FM spectrum

We make the following observations about the FM spectrum as observed over an interval greater than the message period $1/\,f_m$

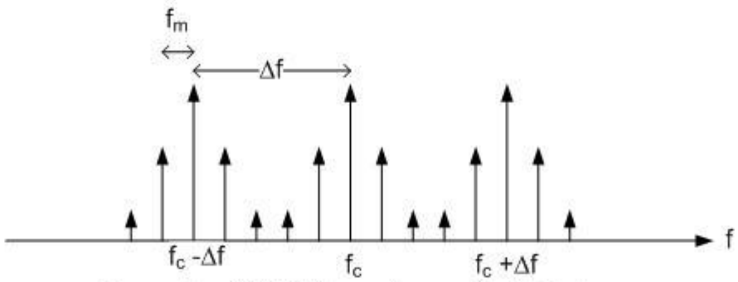
- the spectrum includes components at frequencies $f_c \pm n\, f_m$ that are integer multiples of the message frequency above and below the carrier frequency.
- The number of significant peaks increases as β increases.
- As β increases, more and more peaks become significant, but some peaks become smaller, all in accordance with the Bessel function curve.
- Thus the shape ('envelope') of the spectrum depends on β in a complicated way.
- It is not obvious from the spectrum plot, but for any value of β, and thus for any spectrum shape, the powers in the all of the peaks added together adds up to 1. This is because the average power of an FM wave developed across a 1 ohm resistor is given by:
  $P = \frac{A_c^2}{2} \cdot \sum_{n=-\infty}^{\infty} J_n^2(\beta) = \frac{A_c^2}{2}$ independent of the values of β, since all the Bessel function power spikes in the power spectrum add up to 1.

The FM spectrum depends upon the modulation index $\beta = \frac{\Delta f}{f_m}$ as shown above, this means it is a function of both Δf and f_m.

5.4.2 Wideband FM large modulation index

We consider the FM spectrum for $\beta >> 1$ or $\Delta f >> f_m$. For example

$$\beta = \frac{\Delta f}{f_m} = 10\ ; \Delta f = 10 f_m;$$
$$f_{max} = f_c + \Delta f; \qquad f_{min} = f_c - \Delta f$$

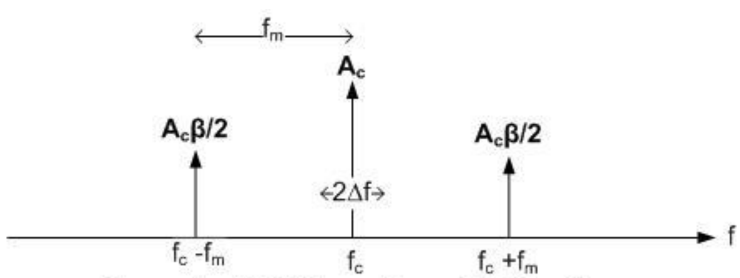


Example of WBFM spectrum with Δf >> f_m and therefore β >>1rad

5.4.3 Narrowband FM, small modulation index

We consider the FM spectrum for $\beta << 1$ or $\Delta f << f_m$. For example

$$\beta = \frac{\Delta f}{f_m} = 0.1; \Delta f = 0.1 f_m$$



Example of NBFM spectrum with Δf << f_m and therefore β <<1rad

5.4.4 Effective bandwidth of FM – Carson's rule

For practical purposes, the bandwidth of the FM wave corresponds to the bandwidth containing 98% of the signal power.

The effective bandwidth of the FM signal is approximately given by Carson's formula:

$$B = 2(1+\beta) f_m = 2\Delta f + 2 f_m = 2(\Delta f + f_m) = 2(k_f A_m + f_m)$$

Carson's formula is intuitively reasonable: the bandwidth must be at least twice the frequency deviation since the instantaneous frequency changes from $f_c - \Delta f$ to $f_c + \Delta f$. The frequency changes at a rate f_m , so this is added to the overall bandwidth (similar to AM).

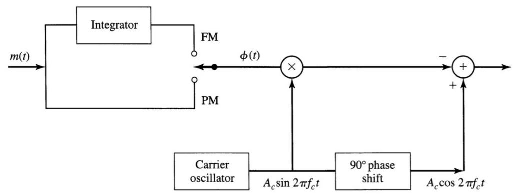
For a general message $m(t) = \sum_i A_i(t)\cos(2\pi f_i t + \psi_i(t))$ containing many cos waves, we can estimate the FM bandwidth by considering the highest frequency (i.e. the message bandwidth W) and highest amplitude A in the message. For this case, we replace $f_m \to W$ and $A_m \to A$ and find

$$B = 2k_f A + 2W = 2(1+D)W$$
$$DW = k_f A$$

Thus the deviation ratio $D = \dfrac{k_f A}{W}$ has the same role as the modulation index $\beta = \dfrac{k_f A_m}{f_m}$ in determining the bandwidth of an FM signal.

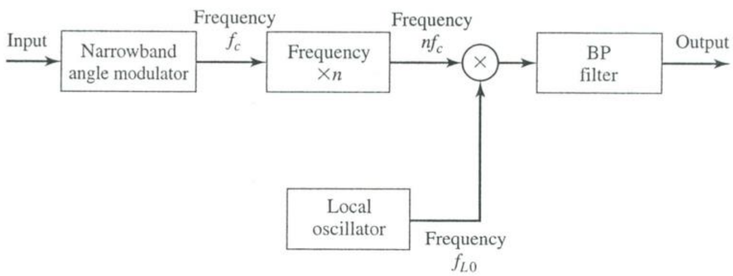
5.5 Frequency modulators (transmitters)

A common analog method is so-called indirect FM modulation.
The process starts with a narrowband modulator as shown above, and repeated below, that implements $s(t) \simeq A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_c t)\sin(2\pi f_m t)$



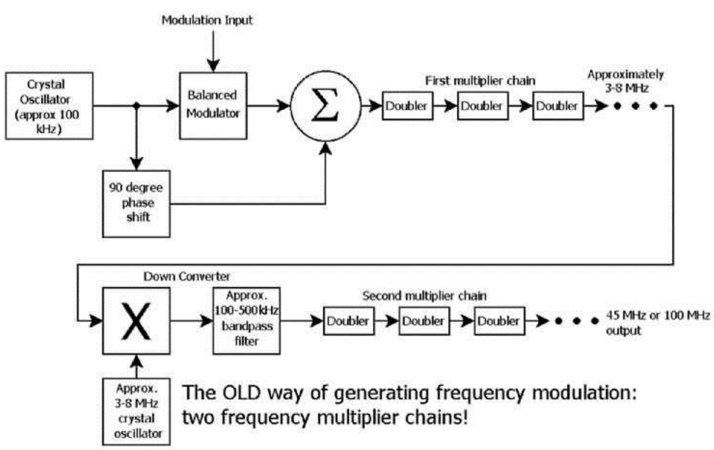
Narrowband angle modulator

The next step is to multiply the signal by itself several times, thus increasing both the carrier frequency and the modulation index, as shown in the figure and mathematics below



$$
\begin{aligned}
s(t) &= A_c \cos(2\pi f_c t + \beta sin 2\pi f_m t) \\
s^2(t) &= A_c^{\,2} \cos^2(2\pi f_c t + \beta sin 2\pi f_m t) \\
&= \mathrm{DC} + A_c^{\,2} \cos(4\pi f_c t + 2\beta sin 2\pi f_m t) \\
s^4(t) &= \mathrm{DC} + A_c^{\,4} \cos(8\pi f_c t + 4\beta sin 2\pi f_m t)
\end{aligned}
$$

We observe that both the deviation and modulation index increase when the signal is multiplied by itself. In practice the multiplication may be many times, as illustrated in the example below. A downconversion stage may be required if the desired carrier frequency is low and the desired deviation (β) is large.



In this figure, the message waveform is integrated and then phase modulated in accordance with the FM equation

$$s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\alpha)d\alpha\right]$$

A digital method of building an FM transmitter simply implements the FM equation in software to generate the I and Q signals that go to the complex multiplier at a local oscillator frequency, where

$$I(t) = A_c \cos 2\pi k_f \int_0^t m(\alpha)d\alpha$$

$$Q(t) = A_c \sin 2\pi k_f \int_0^t m(\alpha)d\alpha$$

I and Q go do a digital to analog converter and a complex upconverter at the desired carrier frequency.

Exercise: draw a diagram of the digital FM transmitter. Use the USRP receiver block diagram as a guide.

5.6 Frequency Demodulation

Frequency demodulation extracts the original message wave from the frequency-modulated wave. We describe two basic devices, the analog-like frequency discriminator and a digital FM demodulator.

How to extract the message from an FM signal? We consider four intuitive approaches.

5.6.1 Intuition 1: Differentiation

We know that from the idea of FM $f_i(t) = f_c + k_f m(t) = f_c + \frac{1}{2\pi}\frac{d\phi(t)}{dt}$ so that $m(t) = \frac{1}{2\pi k_f}\frac{d\phi(t)}{dt}$

Since $s(t) = a(t) cos[2\pi f_c t + \phi(t)]$ with $a(t) = A_c$ (constant) we see that if we differentiate $s(t)$ we will get a term in $\frac{ds(t)}{dt}$ that looks like $\frac{d\phi(t)}{dt}$

A demodulator based on this intuition is in the figure below.

$s(t) \rightarrow$ LIMITER $\rightarrow$ BPF $\rightarrow$ DIFFERENTIATOR $\rightarrow$
ENVELOPE DETECTOR $\rightarrow$ DC BLOCK $\rightarrow m(t)$

We differentiate $s(t)$ as a first step to extract the message $m(t)$. Assuming A_c is constant we have:

$$s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\alpha)d\alpha\right]$$

$$\frac{ds(t)}{dt} = -A_c\left[2\pi f_c + 2\pi k_f m(t)\right]\sin\left[2\pi f_c t + 2\pi k_f \int_0^t m(t)dt\right]$$

This expression is in the form of an AM signal

$$2\pi f_c A_c[1 + k_a m(t)]\cos[2\pi f_c t + \phi(t)] = a(t)\cos[2\pi f_c t + \phi(t)]$$

with envelope $2\pi A_c f_c \left[1 + \frac{k_j}{f_c} m(t)\right]$ with $k_a = k_f / f_c$ and phase $2\pi k_f \int_0^t m(\alpha)d\alpha - \pi/2$

The resulting AM signal can be demodulated by an envelope detector to obtain DC plus the message. The envelope detector ignores the phase.

Note that if the FM signal has $a(t) \neq A_c$ is not constant (e.g. due to channel fading or noise), then differentiation will not work, since for this case

$$\frac{ds(t)}{dt} = -a(t)\left[2\pi f_c + 2\pi k_f m(t)\right]\sin\left[2\pi f_c t + 2\pi k_f \int_0^t m(t)dt\right]$$
$$+\frac{da(t)}{dt}\cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\alpha)d\alpha\right]$$

To make $a(t) = A_c$ is constant, we can use a limiter (e.g. with back to back diodes) ahead of the discriminator. The limiter limits (clips) the input signal so that it is of constant amplitude. The output of the limiter looks like a square wave with changing frequency, and will contain harmonics at odd multiples of f_c. The limiter must be followed by a bandpass filter at f_c to restore the signal $s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\alpha)d\alpha\right]$

A block diagram of this FM demodulator is

$s(t) \rightarrow$ LIMITER $\rightarrow$ BPF $\rightarrow$ DISCRIMINATOR $\rightarrow$
ENVELOPE DETECTOR $\rightarrow$ DC BLOCK $\rightarrow m(t)$

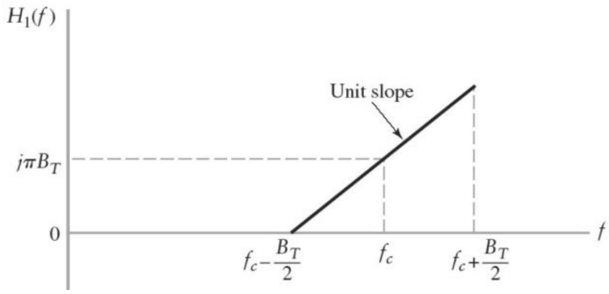
5.6.2 Intuition 2: Linear Amplitude vs. Frequency Characteristic

We know that from the idea of FM $f_i(t) = f_c + k_f m(t)$. If we have a circuit that has output amplitude that increases linearly with frequency, then the circuit output amplitude will vary in step with the message.

One such circuit is the ideal differentiator, with a transfer function given by $H(f) = j2\pi f$.

Proof: If $s(t) \rightarrow S(f)$ then $ds(t)/dt \rightarrow j2\pi fS(f)$

The transfer function acts as a frequency to voltage converter, and is illustrated below, centered at the carrier frequency



The slope is such that the transfer function changes by $j2\pi B_T$ over a bandwidth B_T centered at f_c

The action of an ideal differentiator (figure (a) below) can be approximated by any device whose magnitude transfer function is reasonably linear, within the range of frequencies of interest. In Figure 5(b) an RL circuit approximation to a differentiator is used followed by an envelope detector. A bandpass version of this circuit is shown in (c). These discriminators are known as <u>slope detectors</u>. A more linear response can be obtained by taking the difference between two bandpass magnitude responses, as is done by the balanced discriminator shown in Figure 5(d).

In all cases (a)-(d), a limiter and bandpass filter is required ahead of the discriminator

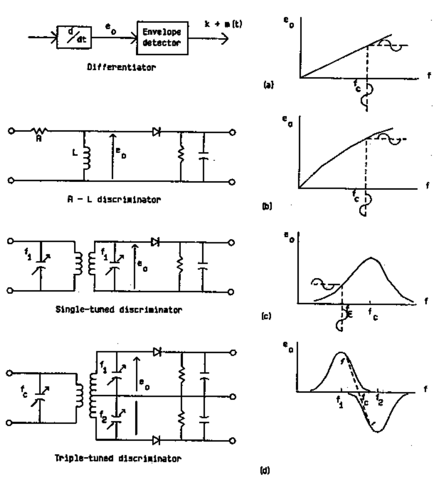
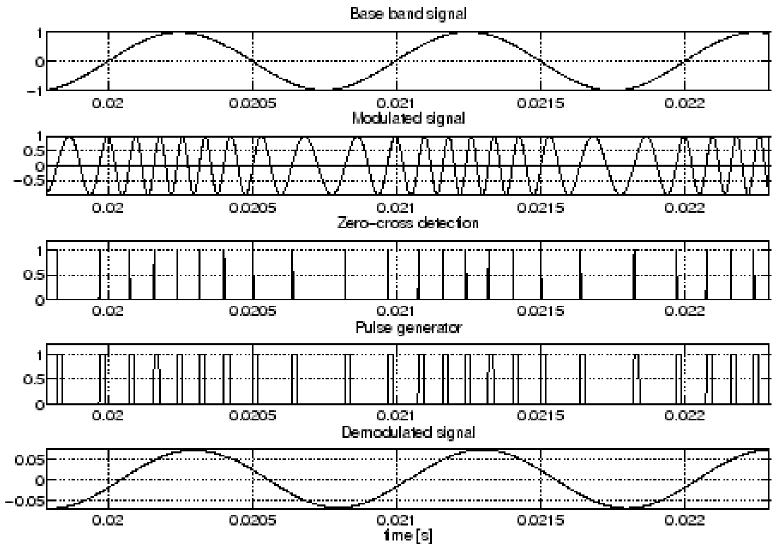


Figure 5 FM Detectors

5.6.3 Intuition 3: Zero Crossing Counter

The message information is contained in the time (location) of the zero crossings, and the amplitude can be ignored, as shown in the figure below.



5.6.4 Intuition 4: Phase Locked Loop

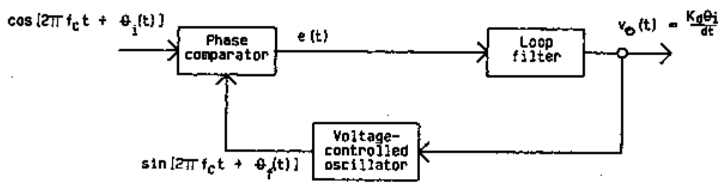


Figure 6 Phase-Locked Loop

This circuit uses ideas from the Costas loop receiver used for DSB. The circuit comprises a phase comparator (multiplier), low pass (loop) filter and voltage controlled oscillator (equivalent to an FM modulator).

With signal input $s(t) = \cos[2\pi f_c t + \theta_i(t)]$ and FM modulator (VCO) output $\sin[2\pi f_c t + \theta_f(t)]$, the phase comparator (multiplier) output signal $e(t) = k_c\left[\theta_i(t) - \theta_f(t)\right]$ plus double frequency terms. With high gain in the loop filter, $e(t) \approx 0$ and the VCO output frequency is the same as the input signal frequency. Thus the VCO input must be the same as the message. The loop filter output voltage is proportional to the instantaneous frequency of the input, and FM demodulation is achieved.

5.7 Digital FM Demodulator

A digital FM demodulator starts with the I and Q outputs of a general IQ receiver. Recall for an FM signal:

$$s(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_0^t m(\alpha)d\alpha\right]$$

$$I(t) = A_c \cos 2\pi k_f \int_0^t m(\alpha)d\alpha$$

$$Q(t) = A_c \sin 2\pi k_f \int_0^t m(\alpha)d\alpha$$

To extract $m(t)$ from $I(t), Q(t)$ we show two methods that can be implemented in software.

5.7.1 I(t), Q(t) as Real Signals

Formula: $$m(t) = \frac{d}{dt} arctan\left\{\frac{Q(t)}{I(t)}\right\}$$

Block diagram: $I(t), Q(t) \rightarrow DIVIDE \rightarrow ARCTAN \rightarrow d/dt \rightarrow m(t)$

Proof:

$$\frac{d}{dt}\arctan\left\{\frac{Q(t)}{I(t)}\right\} = \frac{d}{dt}\arctan\frac{A_c \sin 2\pi k_f \int_0^t m(\alpha)d\alpha}{A_c \cos 2\pi k_f \int_0^t m(\alpha)d\alpha}$$

$$= \frac{d}{dt}\arctan\{\tan 2\pi k_f \int_0^t m(\alpha)d\alpha\} =$$

$$= \frac{d}{dt} 2\pi k_f \int_0^t m(\alpha)d\alpha = 2\pi k_f m(t)$$

This method is not good in practice because when $I(t)$ is small, the division by a small number will cause numerical problems.

5.7.2 I(t) + jQ(t) as a Complex Signal

$$s(t) = \text{Re}\{a(t)e^{j\phi(t)}e^{j2\pi f_c t}\} = \text{Re}\{[I(t) + jQ(t)]e^{j2\pi f_c t}\} = \text{Re}\{\tilde{s}(t)e^{j2\pi f_c t}\}$$

Formula:
$m(t) = \arg[\tilde{s}(t-1)\tilde{s}^*(t)]$

where:

$(t-1) \rightarrow z^{-1}$
represents one sample delay

Block diagram:

$$\tilde{s}(t) \rightarrow \tilde{s}(t-1), \tilde{s}^{*}(t) \rightarrow MULTIPLY \rightarrow ARG$$

Proof:

$$\arg[s(t-1)s^{*}(t)] = \arg[a(t-1)e^{j\phi(t-1)}a(t)e^{-j\phi(t)}]$$

$$= \phi(t-1) - \phi(t) \approx \frac{d\phi}{dt} = 2\pi k_f m(t)$$

This method of FM demodulation is commonly used in software defined radios.