ELEC 350, fall 2018
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.
In this section, we describe how the communications signal may be written as the sum of two complex-valued waveforms resulting in a real-valued waveform.

Radio waves are used to carry a message over a distance determined by the link budget. The radio waves may travel through free space (wireless) or via transmission lines (e.g. waveguide, fiber, microstrip). The radio wave (called a carrier wave) with a defined frequency 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 modified carrier wave is called a communications signal $s(t)$ and has 3 main attributes: frequency, amplitude and phase.

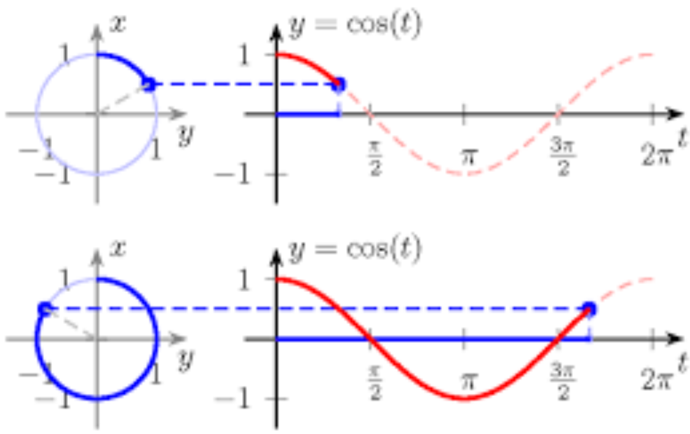
We wish to find a mathematical way to describe $s(t)$.

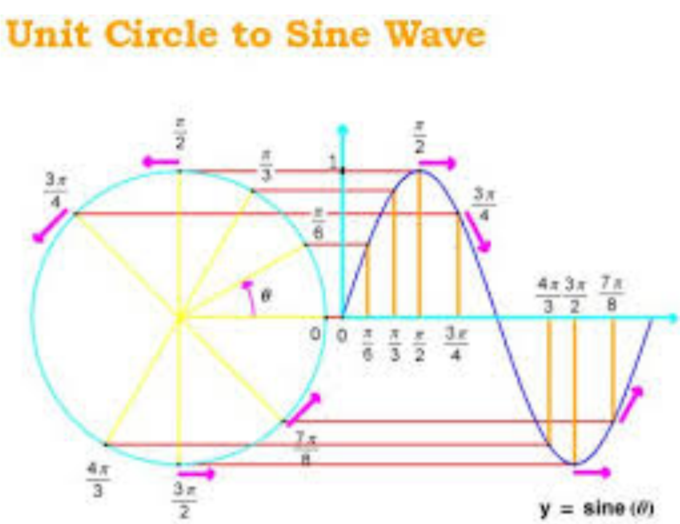
1.1.1 Carrier wave

We begin by considering the carrier wave, which is a “pure” sinewave, such as would be emitted by a laser. It contains one and only one frequency and has constant amplitude and phase that do not change with time. We will call this pure sine wave $c(t)$

We can describe the carrier wave mathematically as $c(t) = \cos 2\pi f_c t$ or $c(t) = \sin 2\pi f_c t$
The carrier wave is a real (not complex) waveform that can be represented as a time-varying analog voltage and plotted on an oscilloscope and connected to a transmission line (e.g. coaxial cable, fiber, waveguide, microstrip) to carry it from one point to another, or connected to an antenna to radiate it into space.

A cosine or sine wave is obtained from the projection of a point moving to trace out a unit circle.





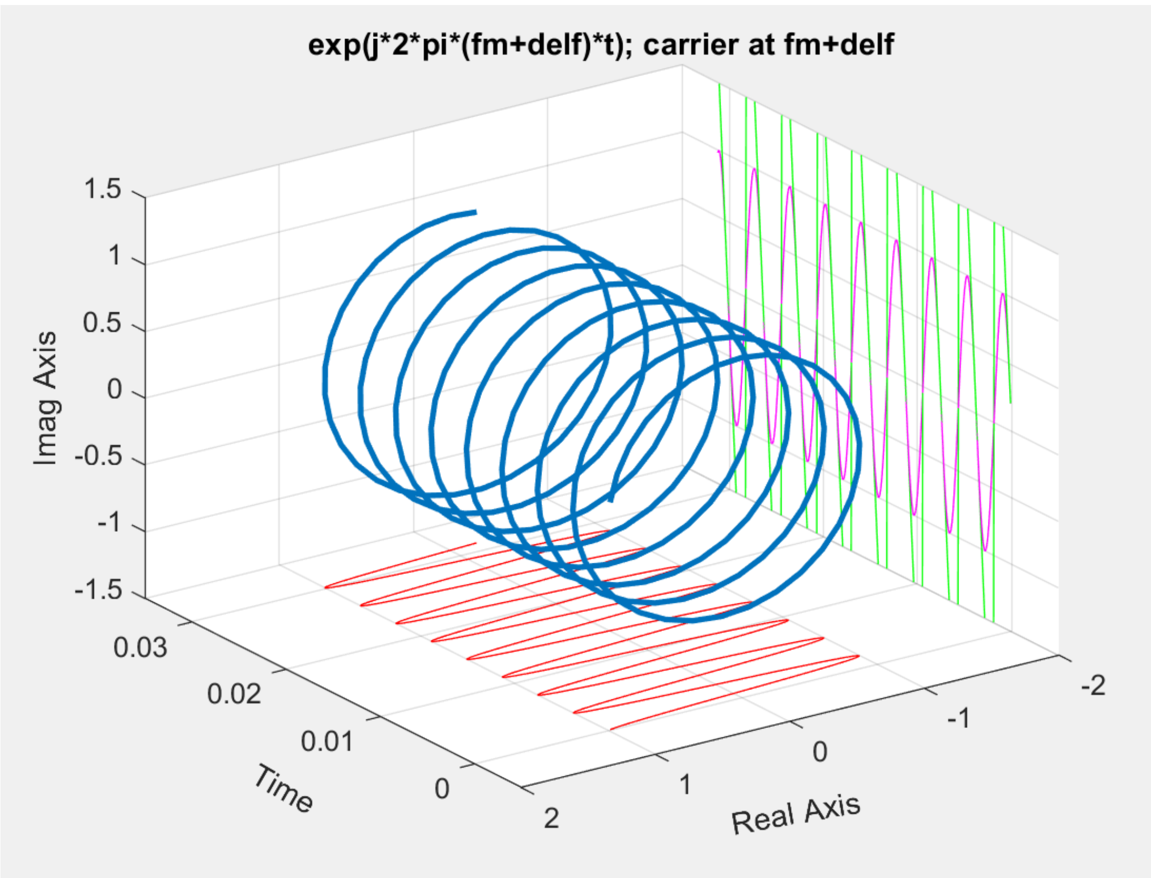
Note that we get the same cosine wave regardless of the direction of rotation, but we get the opposite polarity sine wave when we reverse the direction of rotation.

with theta = 2 pi f_c t

Using Euler's equation $e^{j\theta} = cos\theta + jsin\theta$ and also $\cos\theta = [e^{j\theta} + e^{-j\theta}]/2$ we can write $c(t) = e^{j2\pi f_c t} + e^{-j2\pi f_c t}$. Thus $c(t)$ can be represented mathematically as the sum of two complex exponentials, one positive and one negative. There is a scaling factor of 2 that is not important for the present discussion.

how plot a complex number versus time? 2D plot of a real number versus time, 3D plot of a complex number versus time

We can visualize the positive complex exponential $e^{j2\pi f_c t}$ in 3 dimensions: real axis (real), imaginary axis (imag) and time. This 3-D curve is called a *helix*. $e^{j2\pi f_c t}$ is a complex waveform that cannot be viewed on an oscilloscope as a voltage versus time, and cannot be carried over a transmission line or radiated into space via an antenna.



The end view of the helix (projection on the real-imag plane) shows the circle. The bottom (or top) view (projection on the real-time plane) of the helix shows the real part of the helix(a cosine wave) and the side view (projection on the imag-time plane) of the helix shows the imaginary part of the helix (a sine wave). Either the real or the imaginary parts are real waveforms that can be viewed on an oscilloscope. The end view is what would be seen on an x-y oscilloscope with the real and imaginary signals connected to the x and y inputs of the oscilloscope respectively. Since the cosine and sine wave are 90 degrees out of phase, the x-y plot is a circle.

The helix represents the carrier wave with constant amplitude and phase. In the next section we see what happens when the amplitude and phase are not constant.

In this figure, the helix arises from a counter-clockwise rotation as time increases. The positive complex exponential $e^{j2\pi f_c t}$ is defined to have a positive frequency.

We can visualize the negative complex exponential $e^{-j2\pi f_c t}$ as a helix in the opposite direction. The helix arises from a clockwise rotation as time increases. The negative complex exponential $e^{-j2\pi f_c t}$ is defined to have a negative frequency.

The sum of the two complex exponentials $c(t) = e^{j2\pi f_c t} + e^{-j2\pi f_c t}$ (helixes in opposite directions) is a cosine wave on the real-time plane, and zero in the imag-time plane.

We can also write $c(t) = c_+(t) + c_-(t)$ where $c_+(t) = e^{j2\pi f_c t}$ and $c_-(t) = e^{-j2\pi f_c t}$. This notation shows explicitly that the real carrier wave $c(t) = \cos 2\pi f_c t$ is made up of a positive frequency component plus a negative frequency component.
It is mathematically equivalent to write the real carrier wave as $c(t) = \mathrm{Re}\{c_+(t)\}$ This notation is commonly used, but the disadvantage is that this notation does not clearly show how the carrier wave is made up of positive and negative frequency components.

This completes the discussion of the carrier wave.

1.1.2 Modulated carrier wave with time-varying amplitude and phase

The communications signal $s(t)$ is a modified (or modulated) carrier wave and has 3 main attributes: frequency, amplitude and phase.

We wish to find a mathematical way to describe $s(t)$.

We will again use Euler's equation
$e^{j\theta} = cos\theta + jsin\theta$ and also $\cos\theta = [e^{j\theta} + e^{-j\theta}]/2$
as a starting point.

Recall the carrier wave with positive and negative frequencies is written
$c(t) = e^{j2\pi f_c t} + e^{-j2\pi f_c t}$

Text

To write the communications signal $s(t)$ we modify the carrier wave by changing its amplitude and phase

$$s(t) = a(t)e^{j\theta(t)} + a(t)e^{-j\theta(t)} = a(t)\cos\theta(t)$$

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 if there is no message, then theta(t) = 2 pi f_c t and a(t) = constant.

Often the general form of a radio signal (or any communications signal) is written as the real part of a positive complex exponential

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

This is mathematically identical to taking the projection of the helix in the figure above.

This method of writing $s(t)$ by taking the real part of a complex signal appears in many textbooks and articles, but is potentially confusing, since it appears that something (the imaginary part) is being removed from the positive complex exponential and then the mystery is what happened to that imaginary part. Instead, it may be better to think of adding the negative complex exponential to the positive complex exponential (adding the two helixes) to get the same result, and noticing that the imaginary part is cancelled out by the addition.

exercise: write this out in detail using e^jA = cos A + j sin A

$$s(t) = a(t)e^{j\theta(t)} + a(t)e^{-j\theta(t)} = a(t)\cos\theta(t)$$

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

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

2 pi radians in one second is one cycle per second or 1 Hz
one cycle per second is one cycle of a cos or sine wave per second

For the positive exponential, the frequency is positive and visa versa.

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 a constant carrier amplitude coefficient $a(t) = A_c$.

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

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

As mentioned in the previous section, 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

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

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 or (equivalently) as the sum of a positive frequency complex waveform $s_+(t)$ and a negative frequency complex waveform $s_-(t)$.

In the equation above, the complex exponentials are written in polar form showing amplitude and phase.

The radio signal may also be written $s(t) = \mathrm{Re}\{s_+(t)\} = \mathrm{Re}\{s_-(t)\}$.

but the best way to think about is that s(t) = s+(t) + s-(t)

1.1.3 Complex envelope

We define the *complex envelope* of the communications signal to be $\tilde{s}(t) = a(t)e^{j\phi(t)}$ The complex conjugate of the complex envelope is $a(t)e^{-j\phi(t)}$.

The communications signal is then written

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

$$s(t) = s_+(t) + s_-(t) = a(t)e^{j\phi(t)}e^{j2\pi f_c t} + a(t)e^{-j\phi(t)}e^{-j2\pi f_c t}$$

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

$$s_-(t) = \tilde{s}^*(t)e^{-j2\pi f_c t}$$

For the negative frequency component $s_-(t)$ we use the complex conjugate of 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 $\tilde{s}(t) = a(t)e^{j\phi(t)} = I(t) + jQ(t)$

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

$$\begin{aligned} s(t) &= a(t)\cos 2\pi f_c t \cos\phi(t) - a(t)\sin 2\pi f_c t \sin\phi(t) \\ &= I(t)\cos 2\pi f_c t - Q(t)\sin 2\pi f_c t \end{aligned}$$

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

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

$$\mathrm{Re}\{[I(t) + jQ(t)][\cos 2\pi f_c t + j sin 2\pi f_c t]\}$$

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

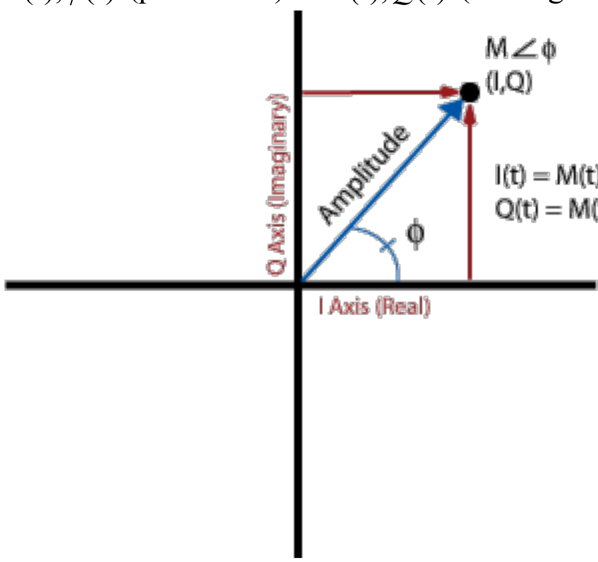
$$= a(t)\cos[2\pi f_c t + \phi(t)]$$

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 or the sum of a positive complex exponential and a negative complex exponential.

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, $M(t) = a(t)$.

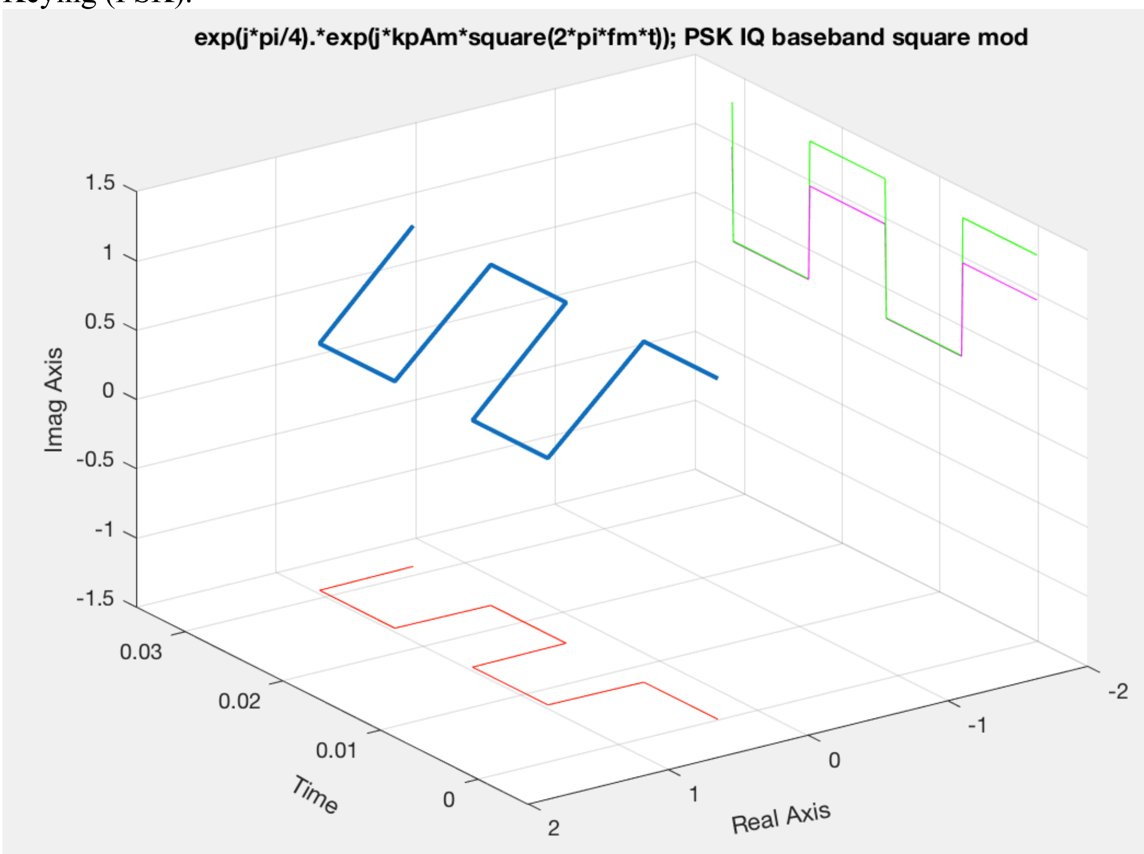


$I(t), Q(t)$ are functions of the message signal(s), where the exact function depends on the modulation type. A simple example is to consider the message signal to be a stereo (2-channel) music signal written as $m_L(t), m_R(t)$ and choose $m_L(t) = I(t), m_R(t) = Q(t)$

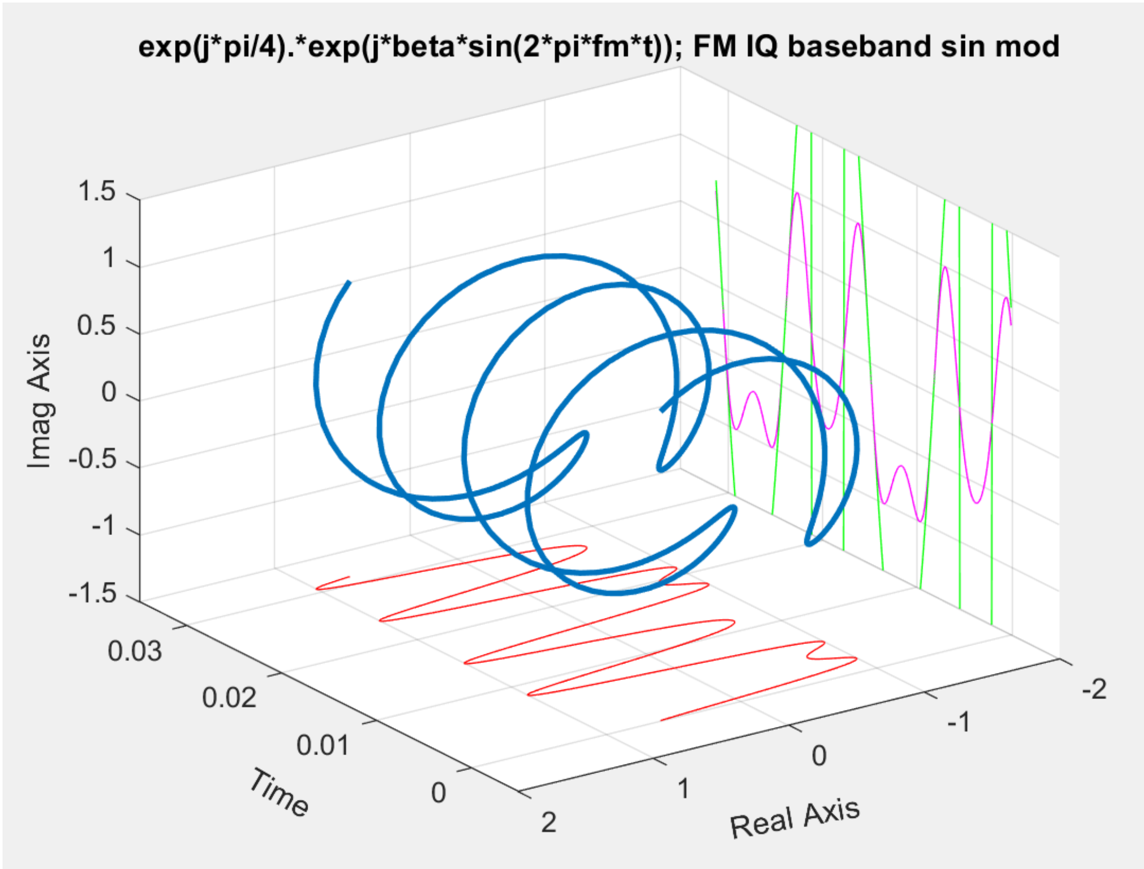
$I(t) + jQ(t) = a(t)e^{j\phi(t)} = \tilde{s}(t)$ is the so-called *complex baseband* signal that is a function of the message. We previously also called this the complex envelope. For a stereo music signal, the complex baseband signal is $m_L(t) + jm_R(t)$. The radio signal $s(t) = a(t)\cos[2\pi f_c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

The complex baseband signal represents the time varying information represented by $a(t), \phi(t)$ which is transported by a carrier wave from one point to another. The complex baseband signal $\tilde{s}(t) = I(t) + jQ(t) = a(t)e^{j\phi(t)}$ can also be represented as a 3-D waveform with projection $I(t)$ onto the real axis and projection $Q(t)$ onto the imaginary axis

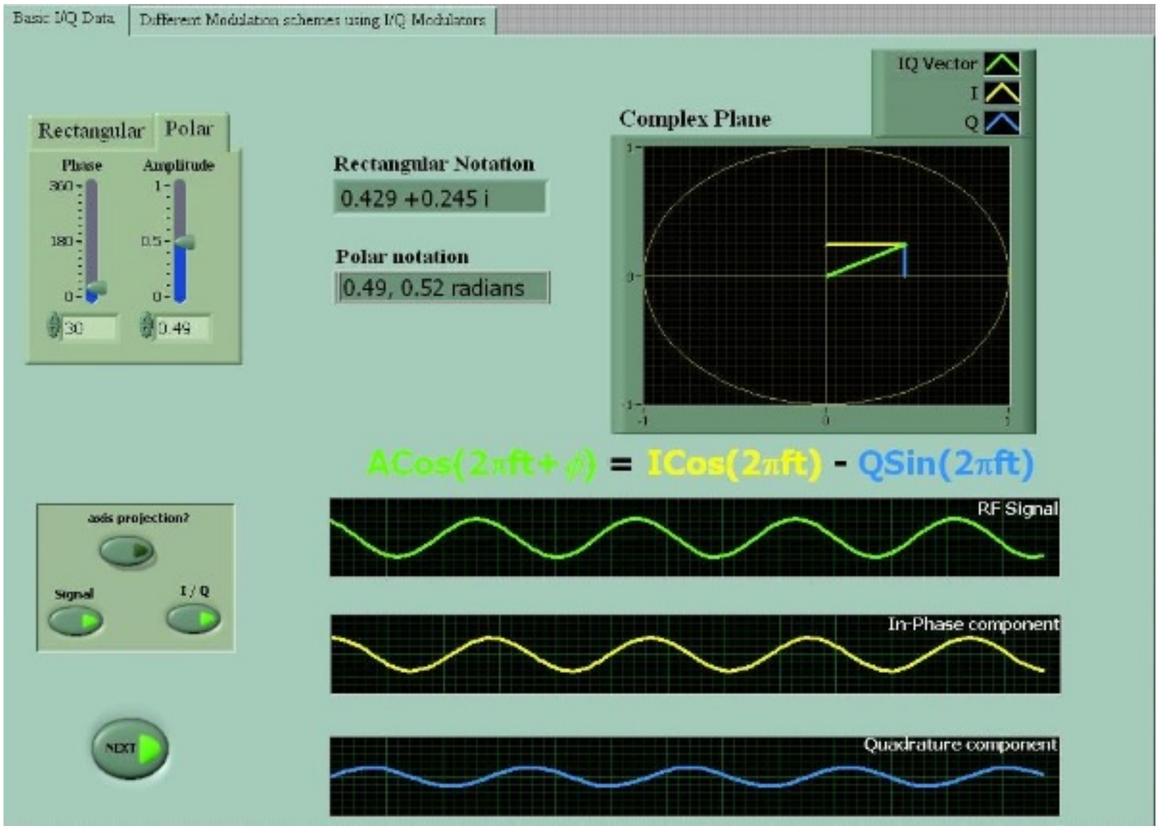
For example, if the amplitude is constant and the phase has only two possible values $-\pi/2$ and $\pi/2$ to represent a to represent a 101010 … data sequence, then the complex baseband signal appears like a bipolar square wave, and the projections $I(t), Q(t)$ are also square waves. In a future chapter, we will recognize this waveform as Phase Shift Keying (PSK).



In another example, if the amplitude is constant and the phase is varied in a sinusoidal manner $\phi(t) = \beta \sin 2\pi f_m t$ at some frequency f_m to represent a 101010 … data sequence (as will be seen in a later chapter), then the 3-D waveform appears as in the figure below. In this example, the helix rotates in one direction, slows down and then rotates in the opposite direction, following the sinusoidal variation in phase. The amount of rotation in each direction depends on β. The projections $I(t), Q(t)$ are periodic patterns which is not easy to interpret without also seeing the 3-D helix. In a future chapter, we will recognize this waveform as Frequency Shift Keying (FSK).



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.



Question: write an expression for both a, ϕ as a function of I, Q

1.1.4 Summary

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

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

The real passband signal is equivalent to the sum of the positive and negative complex passband signals

$$s(t)=s_+(t)+s_-(t)=a(t)e^{j\phi(t)}e^{j2\pi f_c t}+a(t)e^{-j\phi(t)}e^{-j2\pi f_c t}$$
$$s_+(t)=\tilde{s}(t)e^{j2\pi f_c t}$$
$$s_-(t)=\tilde{s}^*(t)e^{-j2\pi f_c t}$$

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.

1.1.5 Frequency shifting

From the definition of a (real) communications signal with information in $a(t),\phi(t)$

$$s(t)=s_+(t)+s_-(t)$$
$$=\frac{a(t)}{2}e^{j\phi(t)}e^{j2\pi f_c t}+\frac{a(t)}{2}e^{-j\phi(t)}e^{-j2\pi f_c t}$$
$$=a(t)\cos[2\pi f_c t+\phi(t)]$$

we observe that it consists of the complex envelope multiplied by a positive complex exponential plus the complex conjugate of the complex envelope multiplied by a negative complex exponential.
We now consider communications signals in the frequency domain.
First recall the Fourier transform.

Fourier analysis shows that any complex waveform can be resolved into sinusoidal waveforms of a <u>fundamental frequency</u> and a number of <u>harmonic frequencies</u>. 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:

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

The Fourier transform has frequency-shifting properties that will be used often in everything that follows. The properties are

$$s(t) \Leftrightarrow S(f)$$
$$s(t)e^{j2\pi f_1 t} \Leftrightarrow S(f - f_1)$$
$$s(t)e^{-j2\pi f_1 t} \Leftrightarrow S(f + f_1)$$

To apply these properties, recall the definitions

$$s(t) = s_+(t) + s_-(t) = a(t)e^{j\phi(t)}e^{j2\pi f_c t} + a(t)e^{-j\phi(t)}e^{-j2\pi f_c t}$$
$$s_+(t) = \tilde{s}(t)e^{j2\pi f_c t}$$
$$s_-(t) = \tilde{s}^*(t)e^{-j2\pi f_c t}$$
$$s(t) = \tilde{s}(t)e^{j2\pi f_c t} + \tilde{s}^*(t)e^{-j2\pi f_c t}$$

Transforming these definitions into the frequency domain

$$S(f) = S_+(f) + S_-(f) = \tilde{S}(f - f_c) + \tilde{S}^*(f + f_c)$$

Thus we see that the real communications signal is made up of the complex baseband signal (containing the message information) shifted by the carrier frequency. The shape of the spectrum of the complex baseband signal $\tilde{S}(f)$ is the same as the shape of the spectrum of the positive frequency signal $S_+(f)$.

The Fourier transform of a real signal $s(t)$ will always be symmetrical around zero Hz, i.e. the shape of the Fourier transform (the spectrum) in the negative frequency range will be the mirror image of the Fourier transform (spectrum) in the positive frequency range.

The Fourier transform of a complex signal such as $s_+(t)$ or $\tilde{s}(t)$ will have an asymmetrical spectrum, i.e the spectrum at negative frequencies will not be the mirror image of the spectrum at positive frequencies. For $s_+(t)$ the spectrum is zero at negative frequencies but not zero at positive frequencies.

A review of how the shifting properties arise is given below.

$$S(f) = \int_{t=-\infty}^{\infty} s(t)e^{-j2\pi ft}dt$$
$$s(t) \Leftrightarrow S(f)$$
$$\delta(t) \Leftrightarrow 1$$
$$1 \Leftrightarrow \delta(f)$$
$$e^{j2\pi f_1 t} \Leftrightarrow \delta(f - f_1)$$
$$e^{-j2\pi f_1 t} \Leftrightarrow \delta(f + f_1)$$

proof

$$S(f) = \int_{t=-\infty}^{\infty} s(t)e^{-j2\pi ft}dt = \int_{t=-\infty}^{\infty} e^{-j2\pi f_1 t}e^{-j2\pi f_1 t}dt$$
$$= \int_{t=-\infty}^{\infty} e^{-j2\pi(f - f_1)t}dt = \delta(f + f_1)$$
$$s(t)e^{j2\pi f_1 t} \Leftrightarrow S(f) \otimes \delta(f - f_1) = S(f - f_1)$$
$$s(t)e^{-j2\pi f_1 t} \Leftrightarrow S(f) \otimes \delta(f + f_1) = S(f + f_1)$$

proof

$$S(f) \otimes H(f) = \int_{\alpha=-\infty}^{\infty} S(\alpha)H(f - \alpha)d\alpha$$
$$S(f) \otimes \delta(f) = \int_{\alpha=-\infty}^{\infty} S(\alpha)\delta(f - \alpha)d\alpha = S(f)$$
$$S(f) \otimes \delta(f - f_1) = \int_{\alpha=-\infty}^{\infty} S(\alpha)\delta(f - f_1 - \alpha)d\alpha = S(f - f_1)$$

For a software radio system 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

Recall the discrete Fourier transform is defined as a Fourier transform operating on a sampled periodic signal

$$S(f = kf_0) = S_k = \int_{t=0}^{T_0 = nT_s} s(t)e^{-j2\pi ft}dt \,\Big|_{t=nT_s, f=kf_0}$$
$$= \sum_{n=0}^{N-1} s(nT_s)e^{-j2\pi kf_0 nT_s} = \sum_{n=0}^{N-1} s_n e^{-j2\pi kn/N}$$

where $T_s f_0 = 1/N$ and N is the number of samples in both the time domain and frequency domain.

For a software radio system receiving real passband signals in the 2.4 GHz range (e.g. WiFi) with sampling rate in the 100 MHz range, the “complex baseband” signal may contain many communications signals in a 100 MHz bandwidth.