

Intuitive Guide to Principles of Communications <u>www.complextoreal.com</u>

All About Modulation

Basic Concepts, Signal Space, Constellations and Phase Shift Keying modulations (PSK, QPSK, OQPSK, M-PSK, π/4-QPSK, MSK, and GMSK)

Basic Concepts of modulation

Three kinds of modulations

Modulation is the process of facilitating the transfer of information over a medium. Sound transmission in air has limited range for the amount of power your lungs can generate. To extend the range your voice can reach, we need to transmit it through a medium other than air, such as a phone line or radio. The process of converting information (voice in this case) so that it can be successfully sent through a medium (wire or radio waves) is called modulation.

We begin our discussion of digital modulation by starting with the three basic types of digital modulation techniques. These are;

Amplitude-Shift Keying (ASK) Frequency-Shift Keying (FSK) Phase-Shift Keying (PSK)

All of these techniques vary a parameter of a sinusoid to represent the information which we wish to send. A sinusoid has three different parameters than can be varied. These are its amplitude, phase and frequency. Modulation is a process of mapping such that it takes your voice (as an example of a signal) converts it into some aspect of a sine wave and then transmits the sine wave, leaving the actual voice behind. The sine wave on the other side is remapped back to a near copy of your sound.

The medium is the thing through which the sine wave travels. So wire is a medium and so are air, water and space. The sine wave is called the carrier. The information to be sent, which can be voice or data is called the information signal. Once the carrier is mapped with the information to be sent, it is no longer a sine

wave and we call it the signal. The signal has the unfortunate luck of getting corrupted by noise as it travels.

In ASK, the amplitude of the carrier is changed in response to information and all else is kept fixed. Bit 1 is transmitted by a carrier of one particular amplitude. To transmit 0, we change the amplitude keeping the frequency constant. On-Off Keying (OOK) is a special form of ASK, where one of the amplitudes is zero as shown below.



Figure 1 - Baseband information sequence - 0010110010

 $ASK(t) = s(t)\sin(2\pi ft)$



Figure 2 - Binary ASK (OOK) signal

In FSK, we change the frequency in response to information, one particular frequency for a 1 and another frequency for a 0 as shown below for the same bit sequence as above. In the example below, frequency f1 for bit 1 is higher than f2 used for the 0 bit.

$$FSK(t) = \begin{cases} \sin(2\pi f_1 t) & \text{for bit } 1\\ \sin(2\pi f_2 t) & \text{for bit } 0 \end{cases}$$
1

2



Figure 3 - Binary FSK singnal

In PSK, we change the phase of the sinusoidal carrier to indicate information. Phase in this context is the starting angle at which the sinusoid starts. To transmit 0, we shift the phase of the sinusoid by 180°. Phase shift represents the change in the state of the information in this case.



Figure 4 - Binary PSK Carrier (Note the 180° phase shifts at bit edges)

ASK techniques are most susceptible to the effects of non-linear devices which compress and distort signal amplitude. To avoid such distortion, the system must be operated in the linear range, away from the point of maximum power where most of the non-linear behavior occurs. Despite this problem in high frequency carrier systems, Amplitude Shift Keying is often used in wire-based radio signaling, both with or without a carrier.

ASK is also combined with PSK to create hybrid systems such as Quadrature Amplitude Modulation (QAM) where both the amplitude and the phase are changed at the same time.

What is digital, what is analog?

There are three parts to a communications system.

1. The information, also called the baseband

3. The carrier

Information can be defined in two forms, digital or analog. Analog signal is considered continuous. Its signal amplitude can take on any number of values between the signal maximum and minimum. Voice is analog and can take any number of volume levels between its "dynamic-range" which is the range of volumes your vocal cords can produce. Digital devices convert analog voice to a digital signal by process of **sampling and quantization**. The analog signal is first sampled and then quantized in levels and then each level is converted to a binary number. For example, we may quantize your voice in 16 levels. Each of these levels can be represented by four bits.

Perhaps you remember when your telephone system went to the "tone" dialing. It went from being a pure analog system to a digital system based on sampling and quantization. Other examples of analog information are music and voice transmitted via FM and AM radio transmissions. Nearly everything else nowadays is digital.

The medium is *thing* the signal travels through. It can be air, space or wires of all sorts. Each of these mediums offers its own unique set of advantages and distortions that determine what is used as a carrier. A short wire in a chip for example may not need a carrier at all. A signal through space such as for satellite transmission may need a very high frequency carrier that can overcome space loss and other atmospheric losses.

If medium is the road taken, then carrier is the truck that carries the information hence we call it *Carrier*. It is a sinusoid in our case. Depending on the medium, it will have a frequency appropriate to the medium. It can be at light frequencies as in optical fiber or a microwave frequency as for mobile communications. An electromagnetic carrier can be of any frequency depending on the medium and the communication needs. Most mediums dictate what type of carrier (its frequency, amplitude) can propagate through it and the type of distortions it will suffer while traveling through it.

Anything that is wireless is analog – always. Wired signals can be digital or analog. Communications inside a computer are examples of pure digital communications, digital data over digital medium. LAN communications are digital data over analog medium. The AM and FM radios are examples of analog data over analog medium.

In general when we talk about a digital system, we are usually talking about digital information over an analog medium. However, there are exceptions. Pulse Coded modulation (PCM) is a form of modulation where there is no carrier, so that makes it a pure digital system.

The "Shift Keying" the second two terms in the name of these modulations imply that they are digital modulations, i.e. the information is also digital.

Signal Spaces and basis functions

The study of signal spaces provides us with a geometric method of conceptualizing the modulation process. In a physical space when we describe a vector by its coordinates (x, y); the vector is being described by a linear combination of two functions (1, 0) and (0, 1). Any vector can be written as a linear combination of these two functions. These functions are called **basis functions** and are orthogonal to each other.

Another example of such a family of functions are the unit width pulses separated in time shown below. Each of these is independent of others and clearly we can use these functions to create any random data sequence consisting of square pulses. Each one of these single pulses is a basis function. However, this is not a very efficient set of basis functions as it takes a large number of these functions to create a random signal.



Figure 5 - Ortho-normal basis set

Ideally we want as few basis functions as possible which when combined can create a large number of independent signals, both digital and analog. In general, basis functions should

- Have unit energy, such as the (1, 0) and the (0, 1) vectors and the above unit pulses.
- They should be orthogonal to every other function in the set, represented mathematically by

$$\int_{-\infty}^{+\infty} \phi_i(t)\phi_j(t) = \begin{cases} 1 & i=j\\ 0 & i\neq j \end{cases}$$
3

An important example of terrific basis functions is the pair of sine and cosine waves of unit amplitude. This special **basis set** is used as carriers in all real communications systems.



Figure 6 - Sine and cosine, two orthogonal functions are the basis set for all modern communications

The concept of I and Q Channels

Without worrying about what a signal is, let's just define it as a vector. Below you see two views of a signal space. One shows a signal in rectangular and the other in its polar form. We can describe the signal it in polar form by its magnitude and it phase (angle) or by its rectangular projections, such as s_{11} and s_{12} .





(b) Polar form

Figure 7 - Signal vector plotted on signal space

In Figure 7(a) the x and y-axis are called In-phase and Quadrature projections of the signal. Quantity s_{11} is I projection and s_{12} is the Q projection of the signal. Figure 7 (b) shows the same signal in polar form, with its length equal to its amplitude and the angle is equal to its phase. These are two canonical ways of representing signals.

The coefficients s_{11} represent the amplitude of I signal and s_{12} the amplitude of the Q signal. These amplitudes when plotted on the x and y axis respectively, give the signal vector. The angle the signal vector makes with the x-axis is the phase of this signal.

Magnitude of signal S =
$$\sqrt{I^2 + Q^2}$$
 4

7

Phase of the signal = $\tan^{-1} \frac{I}{Q}$

In itself, this is simple enough but gets confusing when related to modulation as we shall see.

Symbols, bits and bauds

A symbol is quite apart from a bit in concept although both can be represented by sinusoidal or wave functions. Where bit is the unit of information, the symbol is a unit of transmission energy. It is the representation of the bit that the medium transmits to convey the information. Imagine bits as widgets, and symbols as boxes in which the widgets travel on a truck. We can have one widget per box or we can have more. Packing of widgets (bits) per box (symbols) is what modulation is all about.

In communications, the analog signal shape, by pre-agreed convention, stands for a certain number of bits and is called a symbol.



Figure 8 – Digital information travels on analog carrier

A symbol is just a symbol. It can stand for any number of bits, not just one bit. The bits that it stands for are not being transmitted, what is transmitted is the symbol or actually the little signal packet shown above. The frequency of this packet is usually quite high. The 1 Hz signal shown above is just an abstraction.

A baud is same as the symbol rate of a communication system. So if we send 200 bauds, then we are send 200 symbols per second.

PSK modulations

BPSK

Let's imagine a ship lost at sea with no communication system. It sees an airplanes flying overhead and wants to communicate its plight to the airplane while it is overhead. The

captain marks two spots on each side of the mast as shown below. Now he holds a bright light and runs back and forth between the marked spots to signal a message. Spot to the right means a 1 and spot to the left means a 0. We assume that all airplanes seeing this know that what each light stands.



Figure 9 – Two signaling spots, a simple modulation system

This is a one dimensional signal, because the captain uses only one dimension (running from left to right) to indicate a symbol change.

The shining of the light is a symbol. There are two light positions, so those are two symbols. Let's give these two symbols names of s_1 and s_2 . Simplest thing is to have the symbol stand for just one bit. This method of transmitting information, i.e. the bits, is essentially a Binary Phase Shift Keying (BPSK) modulation. We utilize just one sinusoid as the basis function. We vary the phase of this signal to transmit information which is identical in concept to the example of shining the light from the deck. Each symbol is signaled by a change in position (really the phase) of the light as in this example. In BPSK we define two little packets of the cosine wave, one with zero phase and second one with a 180 degree different phase.

The BPSK signal is special. It lies totally in one axis, x-axis. It has no y-axis projection. The vector flip-flops on the x-axis depending on the value of the bit. Table 1 lists the two symbols and the signals used to represent them. (The carrier signal shown is for f = 1 Hz.) The I and Q amplitudes are the x and y projections computed by setting $f_c = 0$, and

 $\sqrt{\frac{2E_s}{T}} = 1$, then we get, I = 1 for the first symbol and -1 for the second symbol. Q

amplitude is zero for both symbols because sin of both 0° and 180° is zero.

Symbol	Bit	Expression	Modulation Signal at $f_c = 1$ Hz	I At $f_c = 0$	$\begin{array}{c} Q\\ At \ f_c=0 \end{array}$
S1	0	$\sqrt{\frac{2E_s}{T}}\cos(2\pi f_c t + 0)$		1	0
S2	1	$\sqrt{\frac{2E_s}{T}}\cos(2\pi ft + \pi)$		-1	0

Table 1 – Mapping rules for BPSK

Look at the modulation signals in the above table. The little signals shown are at fc = 1 Hz. They are there to give you a feel for what the transmitted signal looks like. Of course, in a real system these would be at much higher frequency. This is what is transmitted in response to the bits not the bits themselves. But what are those funny coefficients in front of the expressions above?

Recall from Tutorial 1 that energy of a signal is equal to

$$E_s = \frac{A^2 T}{2}$$
 5

So instead of writing a amplitude term to make the expression general, we write the equation in terms of energy, where $A = \sqrt{2E_s/T}$. When referring to carrier signals, we typically talk in terms of signal or bit energy, so it makes sense to write the equations in terms of energy, which is what this scaling factor is. Now we can scale the carrier signal for the power with which it is transmitted.

Creating a BPSK carrier

How would we send a bit sequence 0111 0101 0010 1011 using BPSK signaling technique? To transmit this sequence, we need 16 symbols since each BPSK symbol stands for one bit. These are

```
s1 s2 s2 s2 s1 s2 s1 s2 s1 s2 s1 s2 s1 s2 s1 s2 s2 s2
```

Now string together the appropriate symbol signal packets from Table 1 in the right order. Figure below is the modulated carrier that would be transmitted for this sequence if we use the mapping in Table I.



Figure 10 - A BPSK signal for bit sequence 0111 0101 0010 1011

If you could catch the modulated carrier and look at it on a network analyzer, you would see the above. However, the above picture is at a carrier frequency of 1 Hz, which is not realistic. In real systems, the carrier frequency is very high and we would see a signal that covers a lot of cycles between each transition.

What is a transition? A transition is the time at which we switch from one symbol to the next. What happens at the transition boundary is different for various modulations and is quite an important thing. In the case of BPSK, at every bit transition the signal does a 180 degree phase shift.

We worry about what the signal does at transitions because of amplifier non-linearities. Amplifiers used in communications have a very hard time with sudden changes in signal amplitudes and introduce distortions. Since this makes it harder to decode the symbol, we try to control these transitions.

QPSK

Now imagine a different ship. Its captain thinks up a different signaling arrangement. Here the he has marked out four spots on the deck, to the East and West and North and South. He assigns four different combinations to each of the spots as shown below. He can send two bits, with each flash of the light. If he can do it in the same time period as the first ship, then this person would be able to communicate twice as fast.



Figure 11 – A two dimensional signaling system

By creating four signaling spots, he has added another dimension. This gives two basis functions, the East-West and the North-South movements. Now there are four different symbol positions possible and we can assign 2 bits to each unique symbol.

The dimensionality of a modulation is defined by the number of basis functions used. That makes QPSK a two-dimensional signal. Not because it sends two bits per symbol, but because it uses **two** independent signals (a sine and a cosine) to create the symbols. All PSK modulations we will discuss here are two-dimensional.

Now some light math –

QPSK signal is an extension of the BPSK signal. Both of these are a type of M-ary signals. We can write the process that describes the modulated signal in a polar form as

6

$$s_i(t) = A_c ps(t) \cos\left(2\pi f_c t + \frac{2\pi i}{M}\right)$$

Where $p_s(t)$ is the pulse shaping function. In digital phase modulation, the phase of the sinusoid is modified in response to a received bit. The changing phase is shown in blue. A sinusoid can go through a maximum of 2π phase change in one period. So the maximum phase we can change at any one time is 180°. We can use M quantized levels of 2π , to create a variety of PSK modulation. The variable i is a number from 1 to M. The allowed phases are given by

Modulation angles
$$\theta_i = \frac{2\pi i}{M}$$
 7

M stands for the order of the modulation. M = 2, makes this a BPSK, M = 4 is QPSK, M = 8, 8PSK and so on. Following diagram shows three of these modulations and their "constellations." A rotation of the second resulting in the third figure does not change the modulation, its power or performance. These modulation are called rotationally invariant.



Figure 12 – M-PSK modulations, a. BPSK, b. QPSK, c. also QPSK, d. 8PSK

For baseband PSK signals, we use a square pulse. The pulse has an amplitude of A. The energy in this pulse is equal to the power of the signal times the duration, T it lasts. Power is equal to A^2 with R = 1 ohm and T is the symbol time.

$$A = \frac{1}{T} = \frac{A^2 T}{2}$$
$$A = \sqrt{\frac{1}{T}}$$

.

Which gives this equation for the pulse; the pulse has this amplitude over a period T secs.

$$ps(t) = \sqrt{\frac{2}{T}} \qquad 0 \le t \le T$$
 8

Substitute this into equation 6, we get

$$s_i(t) = A_c \sqrt{\frac{2}{T}} \cos\left(2\pi f_c t + \frac{2\pi i}{M}\right) \qquad 9$$

The carrier amplitude A_{c} lets just set it at $\sqrt{E_s}$. Now we have the modulation equation of a general M-PSK signal.

$$s_{i}(t) = \sqrt{\frac{2E_{s}}{T}} \cos \left(\frac{2\pi f_{c}t}{\frac{1}{Changing}} + \frac{2\pi i}{\frac{M}{Changing}} \right) \qquad i = 0, 1, \dots, M \qquad 10$$

An arbitrary plot of this equation looks as shown below. At each tick which is the symbol time, there is a phase shift.



Figure 13 – An arbitrary modulated signal which shows phase shifts at each time tick

Here we see several phase shifts, some 90° as at t = 1 and others 180° as we see at t = 6. A modulated signal for BPSK has only 180° degree phase shifts, whereas a QPSK has 90° degrees and 180° degree phase shifts.

The modulation equation 10, is much more useful in understanding modulation. The first part is the magnitude or the amplitude of the signal and is a constant. The rest is a function of the phase angle just as we would like it to be since we are doing phase modulation.

Now let's expand this equation using the trigonometric identity

$$\cos(A+B) = CosACosB - \sin A \sin B$$

$$s_{i}(t) = \sqrt{\frac{2E_{s}}{T}} \cos\left(2\pi f_{c}t + \frac{2\pi i}{M} + \frac{\pi}{4}\right)$$

$$= \sqrt{\frac{2E_{s}}{T}} \left[\cos\left(2\pi f_{c}t\right)\cos\left(\frac{2\pi i}{M} + \frac{\pi}{4}\right) - \sin\left(2\pi f_{c}t\right)\sin\left(\frac{2\pi i}{M} + \frac{\pi}{4}\right)\right]$$
11

at $f_c = 0$, we get four baseband signals we will use for signaling. We have initialized the phases to start at 45°. This shift has no effect on the modulation.

Now come the I and the Q channels, remember the basis set.

$$\phi_1(t) = \cos \omega_c t$$
$$\phi_2(t) = \sin \omega_c t$$

Any two signals created through the scaled versions of these basis signals are also orthogonal. So let's scale these and call them I and Q channels.

$$I = \sqrt{\frac{2E_s}{T}} \cos(2\pi f_c t)$$

$$Q = \sqrt{\frac{2E_s}{T}} \sin(2\pi f_c t)$$
12

The above are clearly orthogonal because we just multiplied the basis functions with a constant. Now multiply them with the angle part from Eq. 11 part also. For i = 0, 1, 2, 3 and M = 4, a QPSK constellation, the second part is also a constant. So the two equations remain orthogonal.

$$I = \sqrt{\frac{2E_s}{T}} \cos\left(2\pi f_c t\right) \left(\cos\left(\frac{\pi}{4}\right) or \cos\left(\frac{3\pi}{4}\right) or \cos\left(\frac{5\pi}{4}\right) or \cos\left(\frac{7\pi}{4}\right)\right)$$
$$Q = \sqrt{\frac{2E_s}{T}} \sin\left(2\pi f_c t\right) \left(\sin\left(\frac{\pi}{4}\right) or \sin\left(\frac{3\pi}{4}\right) or \sin\left(\frac{5\pi}{4}\right) or \sin\left(\frac{7\pi}{4}\right)\right)$$

Actually true orthogonality requires that f_c be a integer multiple of T/M but for large carrier frequencies, this is not so important. Now we can express the modulation equation as

$$s(t) = \underbrace{\sqrt{\frac{2E_s}{T}}\cos\left(\theta(t)\right)\cos\left(2\pi f_c t\right)}_{1} - \underbrace{\sqrt{\frac{2E_s}{T}}\sin\left(\theta(t)\right)\sin\left(2\pi f_c t\right)}_{2} 13$$

This is called the *quadrature form* of the modulation equation. The two signals are orthogonal. The colored part is the amplitude of the I and the Q channels. The values of

these is the same as the x and y-axis projection of the signal of energy $\sqrt{E_s}$. So a phase modulated signal can now be seen as a combination of two quadrature signals, the amplitude of which changes in response to the phase change. The modulating signal can be seen as a vector with I and Q as its x and y components.

Although we created the BPSK modulated signal by stringing together the appropriate packets of signals, in real systems, we can not create a modulated carrier this way. What we have at our disposal are oscillators that can produce continuous sines and cosines. We can not just use a certain part of the signal as if it was sitting on a shelf for us to grab the needed piece. We need a way to create a signal packet of a particular phase when needed out of a free-running sine or cosine. This is where **Quadrature Modulation** with I and Q channels come into play. I and Q channels are not just concepts but also how modulators are designed. However, the signal created by I and Q channels is not what is transmitted, it is the sum or the difference (makes no difference as long the polar form is consistent) of these two, and that is the real modulated signal.

How the bits are mapped to the possible phases can be done in many ways, especially when you have a lot of symbols. Try this with 8PSK: The bit combinations are 001, 000, 100, 101, 111, 110, 010, 011. There are many different ways of doing this. The best way to do this is to number them such that each adjacent phase means just one bit difference. So that when a phase mistake is made and the most likely one is the nearest phase, then only one bit is decoded incorrectly. This is called Gray coding and is always applied in PSK. In QPSK we can do this perfectly. In other higher order (M > 4) PSK modulations, perfect gray ordering is not always possible.

For QPSK, we have four symbols, each stands for two bits. Nominally we start the first at 45 and then change phase by 90 each time to get the next symbol. The I and Q values are computed by setting $f_c = 0$, and

 $\sqrt{\frac{2E_s}{T}} = \sqrt{2}$. For the first symbol which lies in the first quadrant, the I and Q values are

Symbol	Bits	S(t)	Phase, (Deg.)	Mod. Signal At fc = 1	Ι	Q
S 1	00	$\sqrt{\frac{2E_s}{T}}\cos(2\pi f_c t + \pi/4)$	45°		1	1
S2	01	$\sqrt{\frac{2E_s}{T}}\cos(2\pi f_c t + 3\pi/4)$	135°		-1	1

S 3	11	$\sqrt{\frac{2E_s}{T}}\cos(2\pi f_c t + 5\pi/4)$	225°	 -1	-1
S4	10	$\sqrt{\frac{2E_s}{T}}\cos(2\pi f_c t + 7\pi/4)$	315°	1	-1

Table 2 – Mapping rules for QPSK

Constellation of QPSK

A constellation is a plot of the symbols on the rectangular space. We create this for PSK by first drawing a circle of radius = $\sqrt{E_s}$. Since we like the I and Q channels to have amplitudes of 1, then the radius of the circle becomes 1.414. Now compute the modulation angle, which is 360° divided my M. For PSK that is 90°. So that's four points each 90° apart on the circle. If the modulation is 16 OSK, the points or symbols would be 22.5° apart. Now compute the x and y projections for each symbol. These are the amplitudes of the I and Q channels. Once you know that, you can create I and Q channels and the real modulated signal which is a sum of I and Q channels.

Visually the constellation diagram which is what this picture is called, shows the phases of the symbols and their relationship to each other. The x-axis projection for each symbol is the I channel *amplitude* and y-axis projection is the Q channel *amplitude*. Each signal is shown with a little packet of signal that goes with it. The constellation diagrams is always done at baseband, i.e. $f_c = 0$. So the signal is just a point. But I have added the little packet of modulated signal at $f_c = 1$, just so you can get an intuitive feel for what we are doing. Depending of the f_c , it is these little packets that are transmitted.



Figure 14 – Constellation points are the tips of the modulating signal

QPSK example

Signal s1 shows a string of symbols, numbers from 0 to 4 signifying a random symbol that needs to be transmitted. We map these into I and Q by using the figure above. For

example, the second integer in signal s1 below is 2. Its I channel value is 1 and its Q value is -1.



Figure 15 – QPSK, s1 – arbitrary integer stream indicating a bit stream of 2 bits per integer, s2 – the I channel mapping, s3. the Q channel mapping

Symbol	0	3	2	0	3	3	2	3	3	2
I Data	1	-1	1	1	-1	-1	1	-1	-1	1
Q Data	1	-1	-1	1	-1	-1	-1	-1	-1	-1



Figure 16 – QPSK modulation s4 a cosine wave of frequency 1 Hz, s5 – s4 multiplied by s2, s6 – a sine wave of frequency 1 Hz, s7 – s6 multiplied by s3, s8 – add I and Q channels to get the real modulated signal

Signal s4 is the I carrier or the cosine wave of frequency 1. Signal s5, multiplies I channel values from signal s2 with carrier signal of s4. Signal s6 is the Q carrier or the sine wave of frequency 1. Signal s7 multiplies the carrier s6 with Q(t) of signal s3. Signal s7 and signal s5 are Q and I channel respectively. We add I and Q to get the real modulated signal in s8.



The following diagrams show the two conventional ways of doing modulation.

Method I – The polar form of modulation



Figure 17 - Method II – The quadrature form of modulation using I and Q channels

Method 1 is most straight forward but requires multiplications and square roots. Method II, the quadrature method is easier to implement in hardware so is the predominant method used in digital modulation.

These are called the canonical forms, which just means that this is the conventional way of doing this type of modulation.

You need to recognize that in quadrature modulation, channel I and Q are not transmitted. Only the real signal is transmitted.

Modulation Index of a QPSK signal

The modulation signal can also be written like this.

$$s(t) = A_c \cos\left[\omega_c t + D_p m(t)\right]$$
 14

where m(t) is the information signal, varying between +1 and -1 with a rectangular pulse shape. D_p is called the phase sensitivity. It is equal to the peak phase deviation over one symbol. Its units are radians per volt. Plotting above equation, we get the modulated signal as a function of the phase sensitivity.

The phase sensitivity of a modulated signal is related to the traditional modulation index as

$$h = \frac{2\Delta\theta}{\pi} = \frac{2D_p}{\pi}$$
 15

For a true, carrier suppressed PSK signal, the modulation index is equal to 1 since the peak phase variation is 90 degrees. PSK signals are all carrier suppressed and all have a modulation index of 100%.



Figure 18 – The modulation index of a PSK signal is 100%

Shaping the pulse to reduce bandwidth

The square pulses shown here are not practical to send. They are hard to create and require a lot of bandwidth. So in their lieu we send shaped pulses that convey the same information but use smaller bandwidths and have other good properties such as intersymbol interference rejection. One of the most common pulse shaping is called "root raised cosine". This pulse shaping has a parameter called the roll-off which controls the shape and the bandwidth of the signal.

Some common pulse shaping methods are

- Root Raised cosine (used with QPSK)
- Half-sinusoid (used with MSK)
- Gaussian (used with GMSK)

19

• Quadrature partial response (used with QPR)

In Figure 19 we see the time domain traces of a QPSK I and Q channel where the signal has been shaped by root raised cosine pulses. The general bit shape can be seen easily.



Figure 19 - Root-raised cosine shaped pulses

We can draw a constellation diagram of this signal by sampling this signal every 20 samples, and then plotting the measured I values against Q values. We get the following constellation diagram.



Figure 20 – Constellation diagram of a root-raised cosine shaped signal

Note that the RRC constellation diagram has a scatter around the ideal points. This is an inevitable consequence of any pulse shaping and relates to increased bit error rate for the signal.

A Gaussian pulse shape replaces the square pulses with a Gaussian pulse. Here is a time domain trace for a Gaussian pulse shaped signal.



Figure 21 – The time domain signal shaped by Gaussian pulses.

This is how we do it in hardware.



Figure 22 – Hardware implementation of the MPSK modulator

Summary: The Serial to parallel converter takes the bit stream coming in at a bit rate of Rb and splits it into two streams, each of half the bit rate. Depending on the dual bit pattern coming in, I and Q amplitudes are set from a table lookup function. Each of these is then individually modulated by a sine or a cosine wave of carrier frequency ω after being shaped into a root raised cosine pulse. These are added together to get the transmitted signal.

Constant Envelope modulation

QPSK is part of a class of signals called **constant-envelope** signals. There is no rigorous definition of a constant envelope signal. One definition is; when sampled at the symbol rate, the sampled value of the amplitude is constant. Another is that there are no discontinuous phase changes. Yet another is that the maximum and minimum amplitude attained by the signal over one period is constant. The sine wave is an ideal constant envelope signal.

Constant envelope signals suffer less distortion in high power amplifiers and are preferred for wireless applications. The reason is that amplifiers work by changing a signal's amplitude, either increasing or decreasing it. To increase a signal's power is to increase its amplitude. A non-linear amplifier changes the signal amplitude by differing amount depending on the instantaneous amplitude of the signal. The more the amplitude of a signal varies, the more non-linear amplification occurs and this results in a distorted signal. QPSK is not technically a constant envelope because of its discontinuous phase shifts but is considered nearly so. A shaped-pulse signal is also not constant envelope but is nearly so.



Figure 23 – FSK is definitely a constant envelope modulation.



Figure 24 - ASK is definitely not a constant envelope modulation

Offset QPSK

Offset QPSK is a minor but important variation on QPSK. In Offset QPSK, the Q channel is shifted by half a symbol time so that I and Q channel signals do not transition at the same time. The result of this simple change is that phase shifts at any one time are limited and hence offset QPSK is more "constant-envelope" than straight QPSK. In high power amplifiers and for certain satellite applications, Offset QPSK offers better performance. Although in a linear channel its bit error rate is the same as QPSK, in non-linear applications, its BER is lower when operating close to the saturation point of the transmitting amplifier. Offset QPSK (OQPSK) is also called staggered QPSK (SQPSK).



Figure 25 – QPSK modified to become OQPSK



Figure 26 – I and Q channel mappings of an Offset QPSK signal, the symbol transitions do not occur at the same time.

Unlike QPSK, I and Q channels of an OQPSK signal do not transition at the same time. One consequence of this is that when we look at the constellation diagram of the OQPSK, the symbol transitions occur only to neighbors. This means that the transitions are never more than 90°. At any symbol change, for either I or Q channel, only one axis can change at a time, either I or the Q but not both. (At any transition, only I or the Q changes but not both.) In constellation-speak, if the signal was in the right upper quadrant, the next signal can only go to either the lower right quadrant or to upper left quadrant but not across. Note how this is different from QPSK, where all transitions can occur.



Figure 27 – OQPSK and QPSK constellation diagrams, (1) OQPSK, (2) QPSK



(a) OQPSK – All phase shifts are 90° .



(b) QPSK - Note the 180° phase shift.

Figure 28 – The phase jumps at the symbol transition for OQPSK are smaller. (Note that the figures above are not of the same scale in time.)

Figure 28 compares the OQPSK signal with a QPSK signal. Note that the OQPSK signal never transitions more than 90°. QPSK on the other hand goes through phase change of 180° for some transitions. The larger transitions are a source of trouble for amplifiers and to be avoided if possible. In satellite transmission, QPSK reigns supreme, it is easy to build and operate. Military often uses OQPSK because of its need to use low power radios and minimum adjacent channel interference issues.

How OQPSK differs from QPSK: The Q channel of OQPSK is delayed by a half a symbol time, staggering the two quadrature channels.

Minimum Shift Keying (MSK)

Although MSK is often classified as FM modulation, it is also related to offset-QPSK owing to the dual nature of FSK and PSK modulations. OQPSK is created from QPSK by delaying Q channel by half a symbol from I channel. This delay reduces the phase shifts the signal goes through at any one time and results in an amplifier-friendly signal.

MSK can be derived from OQPSK by making one further change - OQPSK I and Q channels use square root-raised cosine pulses. For MSK, change the pulse shape to a half-cycle sinusoid. Figure 29 shows a MSK pulse signal and then multiplication by the carrier. The red curve is the carrier signal, and the blue the MSK pulse shape and the black the multiplication of the pulse shape and the carrier giving the modulated carrier.



(a) MSK pulse and carrier for a 1 bit

(b) MSK pulse and carrier for a 0 bit

Figure 29 – MSK pulse shaping is a half-sine wave shown in blue, positive for a 1 and negative for a 0.

The carrier signal expression for MSK is

$$c(t) = a(t)\sin\left(\frac{\pi}{2T}t\right)\cos\left(\frac{\pi}{T}t\right) + a(t)\sin\left(\frac{\pi}{2T}t\right)\sin\left(\frac{\pi}{T}t\right)$$

with the underlined portion, the half-sinusoid pulse shape. Figure 29 shows how MSK pulses look compared to QPSK square pulses. Remember in QPSK, the square pulse itself equates to a discrete phase. In MSK, the shape is continuous changing, so there is no discrete jump in the modulated signal at the symbol edge as there is QPSK. For this reason the modulated signal in Fig. 32 has no discontinuities as compared to M-PSK signals.



Figure 30 – MSK pulse, each pulse is a half cycle sine wave.

The dashed line is the QPSK I and Q channel symbols and the solid lines show how these have been shaped by the half sine wave. The I and Q channels are computed by

$$MSKI(t) = QPSKI(t) \left| \sin\left(\frac{\pi t}{2T}\right) \right|$$
$$MSKQ(t) = QPSKQ(t + .5T) \left| \sin\left(\frac{\pi (t + .5T)}{2T}\right) \right|$$

The I and Q channels are then multiplied by the carrier, cosine for the I channel and sine for the Q channel. Note that the period of pulse shape is twice that of the symbol rate.

$$MSKcarrI(t) = QPSKI(t) \left| \sin\left(\frac{\pi t}{2T}\right) \right| \cos\left(\frac{\pi t}{T}\right)$$
$$MSKcarrQ(t) = QPSKQ(t) \left| \sin\left(\frac{\pi t}{2T}\right) \right| \sin\left(\frac{\pi (t+.5T)}{T}\right)$$



Figure 31 – MSK I and Q modulated carriers.

Adding I and Q components gives the MSK carrier of Figure 31. Compare this carrier to a QPSK carrier. This one has much smoother phase shifts at the symbol boundaries. This results in lower side lobes which is an advantageous property for wireless signals since it results in less adjacent signal interference.

Minimum Shift Keying (MSK) is also called continuous phase (CP) Frequency Shift Keying (FSK). MSK is a class of continuous phase modulations. These are particularly suited to media which uses non-linear amplifiers.

FSK is the digital version of analog Frequency Modulation (FM) and MSK is also a form of FSK, where modulation index is equal to .5 which results in a minimum frequency separation such that the modulation frequencies are still orthogonal. (See FM tutorial)



Figure 32 – MSK modulated carrier



Figure 33 – MSK modulator block diagram

Figure 32 shows the modification made to the QPSK modulator to create the MSK signal. Only the pulse shaping has been changed. The half cycle time shift of the OQPSK stays.

How MSK differs from QPSK: MSK is generally considered a FSK modulation but it is exactly the same as OQPSK except that it uses a half-sinusoid for pulse shaping instead of root-raised cosine pulses.

Gaussian MSK (GMSK)

We created MSK by applying a half sinusoid to the square pulse. By using a Gaussian pulse shape, the result can be improved even further. The modulation obtained this way is called GMSK.

GMSK is used in several mobile systems around the world. Global Speciale Mobile (GSM), Digital European Cordless Telephone (DECT), Cellular Digital Packet Data (CDPD), DCS1800 (Digital Communications System in the 1800 MHz band) in Europe, and GSM-based PCS1900 (Personal Communications Services in the 1900 MHz band) in the U.S. uses GMSK.

Recall that the root-raised cosine pulse has a roll off factor, α . The roll-off factor determines how sharply the pulse rolls off to zero energy. A Gaussian pulse similarly has a BT factor that determines how sharply it rolls off. A BT of .3 is used commonly.

MSK and GMSK, both being related to FM modulation, can both be created two ways, 1. as a PSK signal and 2. as a FSK signal. Both are most commonly implemented as a FSK technique.

The Gaussian pulse shape used instead of the half-sinusoid or the root raised cosine is given by

$$g(t) = \frac{1}{2T} \left(Q \left(2\pi B_b \left(\frac{t - .5T}{\sqrt{\ln 2}} \right) - Q \left(2\pi B_b \left(\frac{t + .5T}{\sqrt{\ln 2}} \right) \right) \right)$$

where

$$Q(t) = \int_{t} \frac{1}{\sqrt{2}} e^{-(x^2/2)} dx$$

Quantity B_b is signal bandwidth. BT factor is equal to this number times the symbol time, T.



Figure 34 – Modulated I and Q GMSK carriers



Figure 35 – Add I and Q GMSK carriers to obtain the composite carrier

The GMSK modulated carrier is even better at transitions than MSK and this is the main reason it is used as a standard in some cellular systems.

How GMSK differs from MSK: GMSK is nearly always implemented as a FM modulation. However conceptually it is same as MSK except instead of half-sinusoid as a pulse shape a Gaussian pulse shape is used instead.

8-PSK

Imagine once more the man on the ship, he figures he's got the space on the deck, so why not add more signaling positions. He marks out a circle and doubles the number of places where he will stand to flash up the signal. We can see the problem right away, how is the airplane going to make out where he is standing. But never mind, he goes ahead with his plan. Here is how his new signaling positions look



Figure 36 – 8-PSK uses eight different unique signals

He assigns bit values to each of the eight positions as shown. Note that each set of bits is just one bit different from its neighbor. So if the airplane does make an error in reading his position, most likely this will result in only one bit being misinterpreted. The eight positions are created with x and y distances or by phases of sines and cosines in communications.



Figure 37 – 8PSK constellation diagram and the I and Q channel amplitudes

We have two basis functions again, a sine and a cosine and each configuration has a different phase to indicate a specific bit pattern. We use four different phase values, namely $\pi/8$, $3\pi/8$, $5\pi/8$ and $7\pi/8$. Each of these phase shifts is 45 degrees apart. Each of these is applied to the sine and the cosine to give us a total of eight values.

(Assume
$$\sqrt{\frac{2E_s}{T}} = 1.414$$
)

Symbol	Bits	Expression	Phase	Signal, fc = 1	Ι	Q
S1	000	$s_1(t) = \sqrt{\frac{2E_s}{T}}\cos(\omega t)$	0°		1.414	0

S2	001	$s_2(t) = \sqrt{\frac{2E_s}{T}} \cos(\omega t + \pi/4)$	45°	1.0	1.0
S3	011	$s_3(t) = \sqrt{\frac{2E_s}{T}} \cos(\omega t + \pi/2)$	90°	0	1.414
S4	010	$s_4(t) = \sqrt{\frac{2E_s}{T}} \cos(\omega t + 3\pi/4)$	135°	-1.0	1.0
S5	110	$s_5(t) = \sqrt{\frac{2E_s}{T}} \cos(\omega t + 5\pi/8)$	180°	1.414	0.0
S6	111	$s_6(t) = \sqrt{\frac{2E_s}{T}} \cos(\omega t + 7\pi/8)$	225°	-1.0	-1.0
S7	101	$s_7(t) = \sqrt{\frac{2E_s}{T}} \cos(\omega t + 9\pi/8)$	270°	0	-1.414



Table 3 – 8-PSK signals

8-PSK step-by-step

Step 1 - The bits stream to send is: 100 111 111 111 111 001





Let's name these bit packets for convenience.

Symbol sequence: s8 s6 s6 s6 s6 s2

Map each symbol to I and Q using the amplitudes in Table 4.

	s8	s6	s6	s6	s6	s2
Ι	-0.707	0.707	0.707	0.707	0.707	1.414
Q	-1.414	-1.414	-1.414	-1.414	-1.414	0.707

Each of these three-bit packets are mapped to levels of I and Q channel from the table above.

After the complex multiplication with cosine and sine wave of the carrier frequency we get the signals shown in Figure 39. When I and Q are added together, you get the composite signal in Figure 40.

33



Figure 39 – 8-PSK mapping of I and Q for sequence in Figure 36. Note that both I and Q can take on 4 different values.



Figure 39 –I and Q modulated sequences



Figure 40 –Adding I and Q to obtain the composite carrier

8-PSK transmitted signal shows smaller phase transitions (on the average) then QPSK which is a good thing but since the signals are also less distinctly different from each other, makes 8-PSK prone to higher bit errors. Why then would we want to use 8-PSK? Because, we can pack more bits per symbol, with each symbol transmitted, we can convey three bits. The throughput of 8-PSK is 50% better than QPSK which can transmit just 2 bits per symbol as compared to 3 for 8-PSK. 8-PSK is the first of the bandwidth-efficient modulations.

π /4-QPSK – a variation on both QPSK and 8-PSK

This a variation of QPSK that mimics 8-PSK. Like QPSK, $\pi/4$ -QPSK transmits two bits per symbol. So only four carrier signals are needed but this is where the twist comes in. In QPSK we have four signals that are used to send the four twobit symbols. In $\pi/4$ -QPSK we have eight signals, every alternate symbol is transmitted using a $\pi/4$ shifted pattern of the QPSP constellation. Symbol A uses a signal on Path A as shown below and the next symbol, B, even if it is exactly the same bit pattern uses a signal on Path B. So we always get a phase shift even when the adjacent symbols are exactly the same.

The constellation diagram looks similar to 8-PSK. Note that a 8-PSK constellation can be broken into two QPSK constellations as show below. In $\pi/4$ -QPSK, one symbol is transmitted on the A constellation and the next one is transmitted using the B constellation. Even though on a network analyzer, the constellation looks like 8-PSK, this modulation is strictly a form of QPSK with same BER and bandwidth. Although the symbols move around, they always convey just 2 bits per symbol.



Figure 41 - $\pi/4$ -QPSK constellation mimics 8-PSK but it is two QPSK constellations that are phase shifted.

Step-by-step $\pi/4$ -QPSK

We wish to transmit the following bit sequence. We divide the bit sequence into 2-bit pieces just as we would do for QPSK

Bit sequence: 00 00 10 00 01 11 11 00 01 00

Transmit the first symbol using the A constellation shown in Figure 41 and the next symbol uses the B constellation. For each 2-bit, the I and Q values are the signal coordinates as shown below.

Symb ol	Bits	Symbol ID	I coordina te	Q coordinat e
1	00	A1	.707	.707
2	00	B1	0	1
3	10	A4	.707	707
4	00	B1	0	1
5	01	A2	707	.707
6	11	B2	-1	0
7	11	A3	707	707
8	00	B1	0	1
9	01	A2	707	.707
10	00	B1	0	1

Table 5 - $\pi/4$ -QPSK symbols mapping to I and Q

The I and Q channels for a $\pi/4$ -QPSK signal are shown below in Figure 42. Note that there are five possible levels (1, .707. 0 -.707, -1) and I and the Q channel show this variation in response to the symbols.



Figure 42 – I and Q mapping of $\pi/4$ -QPSK symbols

Step 1	1 - N	lap b	oits t	O SY	ymbol	ls
--------	-------	-------	--------	------	-------	----

Bits	00	00	10	00	01	11	11	00	01	00
Symbols	A1	B 1	A4	B 1	A2	B 2	A3	B 1	A2	B 3

Step 2 - Multiply the I and Q with a carrier (in the example below, the carrier frequency is 1 Hz.) and you get an 8-PSK signal constellation.



Figure 43 – of $\pi/4$ -QPSK symbols traverse over a 8-PSK constellation

The constellation diagram is a path that the symbols have traced in time as we can see in the above diagram of just the symbols of this signal. The path stars with symbol A1, then goes to B1 which is on path B. From here, the next symbol A2 is back on Path A. Each transition, we see above goes back and forth between Path A and B.



Figure 44 – $\pi/4$ -QPSK modulated I and Q Channels



Figure 45 – $\pi/4$ -QPSK modulated carrier

What is the advantage of doing this? On the average, the phase transitions are somewhat less than a straight QPSK and this does two things, one is that the side lobes are smaller so less adjacent carrier interference. Secondly the response to Class C amplifiers is better. This modulation is used in many mobile systems.

There is also a modification to this modulation where a differential encoding is added to the bits prior to modulation. (More about differential encoding in Tutorial 2.) When differential coding is added, the modulation is referred to as $\pi/4$ -DQPSK.

16-PSK

We can keep on subdividing the signal space into smaller regions. Doing so one more time for 8-PSK so that each signal is now only 22.5° apart, gives us 16-PSK. This will give 16 signals or symbols, so each symbol can convey 4 bits. Bit rate is now four times that of BPSK for the same symbol rate. The following figures show the 16-PSK signal at various stages during modulation.





(a) 16-PSK symbol mapping to I and Q channels. (Now the signal has 8 levels.)



(b) 16-PSK modulated I and Q Channels



(c) 16-PSK modulated carrier

Figure 46 – 16-PSK modulated signal

We can see where this is going. We can keep on increasing bits per symbol this way. However, 16-PSK is rarely used. Despite the fact that 16-PSK is bandwidth efficient is that it has higher bit error rate than a common modulation from the class of **Quadrature Amplitude Modulation** called 16-QAM which has the same bit efficiency.

16 QAM

The modulation equation for QAM is a variation of the one used for PSK. The generalized PSK allows changing both the amplitude and the pghase. In PSK all points lie on a circle so the I and Q values are related to eachother. PSK signals are constant envelope because of this. All points have the same amplitude. If we allow the amplitude to change from sysmbol to symbol, then we get a modulation called quadrature amplitude modulation (QAM). It can be considered a linear combination of two DSB-SC signal. So it is a AM and a PM modulation at the same time.

$$s(t) = \underbrace{\sqrt{\frac{2E_s}{T}}\cos\left(\theta(t)\right)\cos\left(2\pi f_c t\right)}_{I} - \underbrace{\sqrt{\frac{2E_s}{T}}\sin\left(\theta(t)\right)\sin\left(2\pi f_c t\right)}_{Q}$$

This equation can be used to create a hybrid type of modulation that varies both the amplitude and the phase. Let's say that M = 16, so that we have 16 symbols, each representing a four bit word. We can lay these out in a circle but they would be too close and the error rate is likely to be high. So how about this constellation?



Figure 47 – 16QAM constellation in the I-Q plane

It turns out that this constellation which has 16 points all spread in the x-y plane, instead of 16 points all on the circle, performs better in some situations. In QAM, the signal points lie in rectangle instead of a circle.

In M-QAM, and this one is for M = 16, we vary not just the phase of the symbol but also the amplitude. In PSK, all symbols sat on a circle so they all had the same amplitude. Here the points closer to the axes have lesser amplitudes and hence energy than some others. We can compute the x and y axis values of each of these points and depending on the total power we want, we can set the value of a. For typical constellation, set a = 1. If we call the symbols integers then they range from 0 to 15. We show a sequence of random integers up 15 in signal s1 below that we will use these to create a 16QAM signal.

Table 4 – 16 QAM Mapping

Symbol	Bits	Expression	Phase	Ι	Q
S1	0000	$s_1(t) = \sqrt{\frac{2E_s}{T}} \cos(2\pi f_c t + \pi/4)$	0°	1	1

(Fill in the rest of the table.)

The first integer in s1 is 0. Its x coordinate is 3 shown in s2. its y coordinate is also 3. The second integer is 15. Its x-coordinate is -3 and its y coordinate is -1. Having created the I and the Q channels this way (it is a table look up function, also called baseband processing), we can now multiply these signals with the cosine and the sine wave carriers. Then add (or subtract) the two and you have the modulated carrier shown in s6.



Figure 48 – Generation of a 16QAM signal

64 QAM

These figures belong to a 64 QAM signal. There are 64 symbols, which means the incoming integers are from 0 to 63, each representing 5 bits. Using a table lookup, we create the I and Q channels, and then the modulated signal from thereon as before.



Figure 49 – Generation of a 64QAM signal

More Example signals

8PSK





Figure 50 – Generation of a 8PSK signal



Figure 51 – Generation of a 16PSK signal

Copyright 1998, 2006 Charan Langton

Thanks to Rolando Menendez for his help and corrections.

I can be contacted at <u>mntcastle@earthlink.net</u>

Other tutorials at www.complextoreal.com