microwaves&rf

JACK BROWNE
Technical Contributor

Basics of Modulation and Demodulation

Radio waves can carry audio, video, and digital information over great distances by using changes in a carrier wave's amplitude, frequency, or phase to represent the information being transmitted.

Information can be sent from a transmitter to a receiver by means of modulation and demodulation, respectively, whether those signals are light waves moving through optical cables, radio waves through metallic cables, or radio waves propagating through the air. The electromagnetic (EM) waves that transport the information are referred to as carrier signals, while the information they carry may be in the form of audio, video, or data.

By changing the amplitude, frequency, or phase, or a combination of the three signal characteristics, information can be added as modulation to a signal. Due to the increased amount of information for transmission and reception, signal-modulation techniques have advanced in their capabilities to handle more data for a given amount of occupied bandwidth, although they have also grown more complex in the process.

Modulation of a radio wave can be performed by varying one or more of its signal components—amplitude, frequency, or phase—while keeping its other signal components constant. (Pulse modulation is yet another form of modulation, without a carrier, in which pulses with precisely known characteristics are transmitted and details can be learned about a target by receiving the reflected pulses from the target.)

AM & FM

The simplest form of carrier modulation, amplitude modulation (AM), has long been the basis for sending audio information to listeners with radios operating at carrier frequencies from about 535 to 1,605 kHz in the commercial broadcast band. AM is also used for maritime communications and navigation, as well as aircraft navigation, at carrier frequencies from 30 to 535 kHz.

In AM radio broadcasts, the amplitudes of the lower and upper sidebands of the center frequency of a broadcast channel are modulated with the audio content from a radio station, to be demodulated at the receiver of a listener. The lower and upper sidebands extend out from the carrier frequency, usually occupying a total bandwidth of about 25% around the carrier frequency. The audio content from a received AM radio wave can be recovered or demodulated by using a diode to rectify the signals and extract the audio content, or else via filtering to separate the high-frequency carrier signal from the audio content.

In frequency modulation (FM), the frequency of the carrier signal is varied as a function of the message or information. As with AM, audio content is the most commonly transmitted information using FM, such as in commercial FM broadcast radios operating on channels from 88 to 108 MHz. FM can be created by applying message signals directly to a voltage-controlled oscillator (VCO), so that the VCO's output will be a function of the input signal.

Phase-modulation and -demodulation techniques are more complex than modulation and demodulation based on amplitude and frequency. However, they provide the benefit of higher data rates for the amount of bandwidth consumed. Phase modulation is the basis for many digital modulation formats, in which a modulated signal is divided into in-phase (0 deg.) and quadrature (90 deg.) signal components. In contrast to sending video or audio information, digitized information can be easily transmitted by means of digital modulation formats since the modulated information need not be sent continuously in time, but can be sent in bursts or staggered with time and reconstructed at the receiver and demodulator.

KEYING IN ON DIGITAL MODULATION

Digital modulation relies on digital signal processing, such as digital-to-analog converters (DACs) at a receiver and analog-to-digital converters (ADCs) at a transmitter to transform analog information (e.g., audio or video) into a digital form that can then be represented by varying the characteristics of a carrier wave. The three fundamental types of digital modulation—amplitude-shift-keying (ASK), frequency-shift-keying (FSK), and phase-shift-keying (PSK) modulation—use changes in amplitude, frequency, and phase to represent digital bits "0" and "1."

In ASK, the signal amplitude is varied as a function of the information to be transmitted, and all other parameters of the signal remain constant. When sending digital information, one amplitude represents a 0 digital bit while a higher or lower amplitude represents the 1 bit. Waveforms with ASK have the rapidly changing amplitude levels representing a digital bit stream.

In FSK, two different frequencies are used to represent the digital 0 and 1 values. The shift in frequencies in FSK is implemented in different ways, notably in a noncoherent or coherent format. In noncoherent FSK, discontinuities exist between the frequencies that represent the digital bits. Termed "mark" and "space" frequencies, they are used as kinds of frequency gaps to separate the bit-representing frequencies. In coherent FSK, the changes in bit-representing frequencies are instantaneous, without phase discontinuities between the frequencies.

In PSK, the phase of the carrier is discretely changed to denote the different digital bits. The phase can be changed in relation to a reference phase, such as using 0 deg. for a 0 digital bit and 180 deg. to represent the digital 1 bit, or if a difference of 180 deg. is used to denote different digital bits, one of the bits may be represented by a relative phase of –90 deg. and the other by +90 deg. In such a simple, biphase modulation format, the two phase angles of the carrier represent two digital bits of information, so that the modulation rate is equal to the bit rate. But if a greater number of phase angles is used, the bit rate can be increased in parallel with an increasing number of phase angles.

In a quadrature phase-modulation format such as quadrature phase shift keying (QPSK), where four phase angles are used to represent the digital bits, two bits of digital information are able to be carried with each phase angle, so that the information can be represented as 00, 01, 10, and 11. Similarly, if eight phase angles are employed in the phase-modulation scheme, then three digital bits can be represented by each of eight possible phase angles. In turn, the bit rate will increase as the number of phase angles grows in the phase-modulation scheme.

As a result, many digital-modulation formats based on changes in phase attempt to represent the greatest number of digital bits possible by variations in phase, so as to support the highest bit rates possible. This performance parameter of a modulated waveform, spectral efficiency, refers to the number of bits that can be transmitted during a given period of time and for a given portion of bandwidth, usually measured as b/s/Hz.

ASK can be affected by nonlinearities in a system—for example, any form of nonlinear distortion like nonlinear amplification—so it is essential that components with extremely linear performance be used to preserve the amplitude characteristics of a transmitted and received signal. FSK, on the other hand, requires high frequency stability in a system's signal sources, such as VCOs used for local oscillators (LOs) in receiver and transmitters. To maintain high frequency stability, oscillators in FSK systems are typically stabilized by means of phase-locked loops (PLLs) to synchronize the frequency and phase of the system's frequency sources to a common reference source. In addition, PSK depends on tight phase tolerances in a system, such as the lengths of transmission lines, where variations can mean increasing phase errors with increasing transmission frequencies.

Editor's Note: This is part one of a two-part article on the basics of modulation and demodulation. The next installment will examine some of the more complex forms of digital modulation, and explain the use of the time domain and pulsed signals in systems employing pulse modulation (e.g., military radars and automotive collision-avoidance systems). Part 2 will also review the types of hardware needed for each type of modulation/demodulation format, and which modulator/demodulator performance parameters are most critical to achieving good communications-systems performance with high spectral efficiency.