

# **Modulation Methods Part 1**

## **CW and AM**

**William Sheets K2MQJ**

**Rudolf F. Graf KA2CWL**

The ability to transmit information on some form of carrier (usually electromagnetic radiation) depends on a process called modulation. The ability to generate the required amounts of energy at any frequency is of course necessary for transmission of intelligence via electromagnetic radiation, but if this energy cannot have information applied to it in some way, it is useless for most communications purposes. Parameters that can be controlled are the amplitude or power level, frequency of the radiation, and the phase of the waveform of the energy with respect to a known reference. In addition, some means of extracting this information from the transmitted radiation is necessary. This process is usually called demodulation or detection. For this discussion, we will assume that a carrier consisting of electromagnetic radiation in the radio frequency spectrum will be used. This can be any frequency, but we will assume it is on a frequency between 10 kHz and 300,000 MHz. These limits are those presently allocated for communications purposes. 10 kHz is low enough in frequency to be audible as a high pitched tone, if a headphone or speaker is used. Above 300,000 MHz, about as high as can be presently readily handled by microwave techniques, the radio spectrum becomes a region which is called the submillimeter region, and then above about 30,000,000 MHz, (10 microns wavelength) becomes the far infrared region of the spectrum. This radiation can be felt as heat rays. Visible light starts at about 430,000,000 MHz, (0.7 microns wavelength) perceived by the eye as red light. Lasers operate in the far infrared to visible spectrum and these can also be modulated. These frequencies allow almost unlimited modulation bandwidth and are used for fiber optical communications. Even though we will confine this discussion to radio frequencies, one should be aware that other forms of radiation can also be modulated, and the same theoretical concepts will apply, although the physical methods and techniques will be generally very much different than those used in the radio spectrum.

The simplest and oldest form of modulation is a digital type, that of turning on and off a source of energy. The energy (light, RF carrier, etc) is either present or absent. (See Fig 1) In old times lanterns with shutters were used. Then the Morse telegraph, which used a DC current that was turned on and off to form the dashes and dots of Morse Code. Later, radio waves were used to accomplish the same thing. A key is used to turn on and off a transmitter generating a continuous wave (CW) signal. Morse code is sent this way, and although this technique is not used as widely today, it remains one of the simplest and most efficient means of communication known. Only a very simple transmitter, even a very simple oscillator circuit with a single transistor can be employed. The inherently narrow bandwidth occupied by

the signal allows a very narrow band receiver to be used (20 to 100 Hz). This enables low power transmitters to send signals thousands of miles. Reception with a relatively simple receiver is possible. Radio amateurs do this quite often, and this activity is called “QRP operation”, where QRP is a CW shorthand for signifying reduced or lowered transmitter power. Worldwide contacts have been made with only a milliwatt of power in the HF region of the spectrum ( 2 – 30 MHz ), and often enough to be almost commonplace. Before modulation methods are discussed, one factor that limits potential performance of any given system should be discussed. This factor is the noise inherent in any physically realizable system. The limiting factor on how weak a signal can be and still be receivable depends on the receiver bandwidth, temperature, and type of modulation. In the following discussions some high school math is used (algebra and trigonometry). Sorry for the math, but there is really no better way to present the following discussion properly. Mathematics is a fascinating field and the language of science. If you really want to get into electronics, or other aspects of engineering or the physical sciences, you need mathematical proficiency to fully understand many theoretical and practical design concepts. If you would rather not follow the math, you will have to take our word for the figures and numbers we use.

The noise power measured in watts in any bandwidth is given by the formula  $\text{Power} = KTB$ . See a physics textbook for the derivation of this equation if you are curious.  $K$  is Boltzmanns constant, which is equal to  $1.38 \times 10^{-23}$  joules/degree K,  $T$  is the absolute temperature in degrees Kelvin, and  $B$  is the bandwidth in cycles per second (Hz). One joule is equal to one watt for one second and is a measure of energy. At normal room temperature, (taken as 20 deg C or 68 deg F),  $T$  is 293 deg K. Multiplying this out, at room temperature in a 1 Hz bandwidth we have a noise power of  $4.04 \times 10^{-21}$  watts of power. The watt is inconveniently large for this work, so the milliwatt (.001 watt) is used instead. This noise level is then  $4.04 \times 10^{-18}$  milliwatts. Since, in RF systems we are usually dealing with very large variations in power levels, the decibel system is used to avoid inconveniently large or small numbers and ratios. Converting this power level to the more useful measurement of decibels referred to a milliwatt, remembering that a decibel is a logarithmic ratio of two power levels:

$$\text{dB} = 10 \log (P2/P1) \text{ for a ratio of } 4.04 \times 10^{-18} \text{ where } P2/P1 \text{ is the power ratio}$$

$$\text{dB} = 10 \log (4.04 \times 10^{-18}) = 10 \log 4.04 + (-18) \times 10 \log 10$$

$$\text{dB} = 10 ( 0.606 - 18 ) = -173.94 \text{ dBm ( very closely equals } -174 \text{ dBm)}$$

Note:         $\text{dB} = 10 \log (P2/P1)$  Where  $P2$  and  $P1$  are power levels  
                $\text{dB} = 20 \log (V2/V1)$  Where  $V2$  and  $V1$  are voltage levels

$\text{dBm} = \text{decibels with respect to 1 milliwatt reference power level}$   
 $0 \text{ dBm} = 1 \text{ milliwatt} = 0.223 \text{ volts RMS in a } 50 \text{ ohm system}$

As an example the following figures are typically those signal levels one would encounter in operating HF (2-30 MHz) SSB or CW amateur radio equipment. Figures have been rounded off, are for a 50 ohm impedance (the usual situation), and are approximate within a percent or so. The S meter readings are those that would be seen on a typical shortwave receiver signal strength (“S”) meter:

- 20 dBm = 22.3 millivolts, “pegs“ S meter, very strong signal
- 47 dBm = 1 millivolt (approx), S9 + 26 dB, strong signal level
- 60 dBm = 223 microvolts, S9 + 13 dB, a good signal
- 73 dBm = 50 microvolts, an S9 (average) signal
- 87 dBm = 10 microvolts, (S7+) a weaker but still decent signal
- 107 dBm = 1 microvolt, (S3) weak, SSB marginal, CW is OK
- 127 dBm = 0.1 microvolt, very weak, CW only readable

The dBm is independent of the resistance or impedance of the system, but the impedance must be specified for it to have any relation to actual voltages or currents. Since noise voltage is related to power and resistance, and power is  $V^2/R$ , then the noise voltage across a resistance is

$$V_{\text{noise}}^2 = KTB/R \quad \text{and} \quad V_{\text{noise}} = \sqrt{KTB} / \sqrt{R}$$

Since in any generator with a voltage  $V$  and internal resistance  $R$ , the maximum power available to the load occurs when  $R_{\text{load}} = R_{\text{generator}}$ . This is called the maximum power transfer theorem. The load power will be  $(V/2)^2/R$ , or  $V^2/4R$ . Then the noise voltage will be :

$$V_{\text{noise}} = \sqrt{4KTBR} \quad \text{where } K = 1.38 \times 10^{-23}$$

$T = \text{Temp deg K} ; \quad \text{note: deg K} = \text{deg C} + 273$   
 $B = \text{Bandwidth Hz}$   
 $R = \text{Resistance in ohms}$

Normally we use power levels in noise work as it is more convenient. In a system for example, the noise power level is inherently  $-174$  dBm in a 1 hz bandwidth. Considering a 10 kHz bandwidth typically used in an AM broadcast receiver, we could take the ratio of 10kHz to 1 Hz as 10000 to 1, which is a 40 db power ratio ( $10 \log 10000$ , or  $10 \times 4$  since the log of 10000 is 4, therefore a 10000 to 1 ratio, which is 40 dB). Adding 40 dB to  $-174$  dB gives  $-134$  dBm, or 134 dB below a milliwatt. In a 50 ohm system, 1 milliwatt equals 0.223 volts RMS across 50 ohms. Since:

$$DB = 10 \log (P2/P1) \quad \text{then } \log P2/P1 = DB/10 \quad \text{and } P2/P1 = \text{antilog} (DB/10)$$

This means that we divide the dB ratio by 10 and find the inverse log of the result, in this case 13.4. Since we want the voltage ratio, which is the square root of the power ratio for a given resistance, we can divide the logarithm by two. This gives 6.7. Finding the antilog of this will give the voltage ratio that 134 dB represents

Antilog (6.7) =  $5.01 \times 10^6$  or a 5.01 million to one ratio.

This means that  $-134 \text{ dBm} = 0.223/5.01 \times 10^6$ . This comes out to 0.045 microvolts across 50 ohms, and this would be the noise power level in a perfect receiver with a 10 kHz bandwidth. This is theoretically the minimum detectable signal (MDS), if this is assumed to be when the received signal power equals the noise power. (This is only an assumption, as techniques exist for detecting signals below the noise, and the MDS also depends in the signal processing used in the receiver.) A good Morse code operator can usually copy a weak signal that is at the receiver noise level. However, receivers are not perfect. Good receivers used for VHF-UHF work may have noise figures of 1 dB, which means that the receiver noise level is 1 dB above ideal. A typical HF receiver has a 10 to 20 dB noise figure, so the signal detectable in a 10 kHz bandwidth in this case would be 10 to 20 dB higher (A 3 to 10 times voltage ratio). External and atmospheric noise limits reception anyway, so noise figures lower than 15 dB or so are of dubious advantage in an HF receiver, especially below 20 MHz. (Strong signal performance is generally more important in the HF region). This would then be 10 times 0.045 microvolts, or 0.45 microvolt. However, for voice work at least a 6 dB signal to noise ratio is needed for barest intelligibility, with 10 dB being more like it. This raises the minimum input signal to the 1 to 1.5 microvolt level for copy of a voice signal, such as that from an AM medium wave or short wave station. You would probably not listen to this program for a long time, as it would be quite noisy, Another 10 to 20 dB signal level would be needed for comfortable copy, depending on how badly you wanted to listen to it. This brings the signal level up to 5 to 15 microvolts for reasonable reception. The important thing is the signal to noise ratio, and not just the signal level. In noisy reception areas stronger signals are needed. For any system, the bandwidth is important in optimizing the quality of the received signal. Too wide, we get more noise and poorer signal to noise ratio. Too narrow, we may lose some of the information in the signal, or introduce distortion.

In the case of the Morse CW signal, the necessary bandwidth can be estimated by examining the signal. (See Fig 1) At a speed of 25 words per minute (A fairly rapid, but comfortable speed typical of experienced CW operators) this would be about 125 Morse characters per minute, assuming an average 5 letter word. This is roughly 1 letter and space per 500 milliseconds. Taking a worst case, the Morse code symbol for the number 5 has five consecutive dots, and can be considered as a square wave with 5 complete cycles in half a second. This is equivalent to a 10 Hz square wave. A square wave consists of frequencies that are mainly fundamental, and the third, and fifth harmonics (odd) of the fundamental. If the square wave is asymmetrical, typical for Morse Code as there are dots, dashes, and spaces, there are second and fourth (even) harmonics also. Although it is an approximation, a

square wave decent enough to be copiable as a Morse Code signal consists of harmonics up to at least the fifth. Therefore a minimum bandwidth of 50 to 100 Hz would be needed in this example, for 25 words per minute speed of transmission. This allows for some tuning error and short term receiver drift. More than this, the signal to noise ratio will start to decrease. Less bandwidth will cause loss of the higher harmonics and rounding of the waveforms to where the signal would be difficult copy unless the sending speed were reduced. If speeds of 5 words per minute were employed, bandwidth could be reduced accordingly at the expense of speed of transmission. This is why very weak signal CW work is done at slow transmission speeds, to allow narrow bandwidth and an increase in effective receiving sensitivity.

In practice, many receivers for amateur radio CW use 200 to 400 Hz bandwidth as it allows for more comfortable tuning by the operator and for some receiver drift, and less costly filtering. Even with 400 Hz bandwidth and a 20 dB noise figure, the minimum discernable signal level is around 0.1 microvolts, depending on the operators skill and hearing acuity. In most cases external noise will be the limit anyway. A 0.5 microvolt signal is typically comfortable copy. Contrast this with the 5 to 15 microvolt figure needed for AM or 2 to 5 microvolts for SSB for marginal copy, and you can readily see the advantages of CW techniques using Morse code or other forms of slow speed digital modulation in weak signal work. In this era of cheap and powerful computers, the internet, cell phones and sophisticated equipment, simplicity still has a place of importance. It is a sobering fact and also somewhat amusing to note that the use of plain old (obsolete....?) Morse Code, 1940 era radio technology, and a skilled operator, can give reliable and dependable emergency communications when all else is knocked out. Only a simple transmitter, a shortwave receiver, and a length of wire strung up between two trees or other supports are needed to get a station on the air. A 12 volt auto battery will do for power. In emergency situations, communications might be impossible using much more sophisticated equipment, whose operation depends on a vulnerable infrastructure destroyed or rendered inoperable in a natural disaster, or made useless and/or inaccessible during a lockdown, terrorist, or national emergency. Do not count on using the internet, the telephone system, or your cell phones at these times.

The next form of modulation that evolved was probably amplitude modulation, called AM. In this case the amplitude of the signal is modulated in some way by the waveform of the intelligence to be transmitted. In this case the envelope of the transmitted AM signal is a replica of the modulating waveform (See Fig 2). In the usual case, the carrier is a sinusoidal waveform, and the modulation is audio or data. The modulating waveform can be represented as a superposition of harmonically related sine wave components (Fourier's Theorem). The amplitude of the carrier waveform is modulated by the modulation and a mixing action takes place.

The carrier waveform can be represented as:

$V_c(t) = A \sin \omega_c T$  where  $\omega_c = \text{freq. radians/second} = 2\pi \times \text{Frequency in Hz}$   
 $A = \text{peak amplitude of sinewave in volts}$   
 $V_c(t) = \text{Instantaneous voltage of carrier}$   
 $T = \text{time}$

And if a waveform is available, having an amplitude that swings between zero and  $V_m$  volts described as:

$V_m(t) = 1 + M \sin \omega_m T$  where  $V_m(t) = \text{total modulating signal}$   
 $V_s(t) = \text{modulating signal}$   
 $\omega_m = \text{modulating frequency rad/sec}$   
 $T = \text{time}$   
 $M = \text{relative amplitude of modulation}$   
 $(M \text{ is } 0 \text{ minimum to } 1 \text{ maximum})$

then this signal can be used to modulate a carrier signal.

If these two signals are mixed (multiplied together) in a modulator circuit, which produces an output which is proportional to the mathematical product of the input signals, the resultant output is an amplitude modulated signal. We will assume this circuit has a gain of unity for simplicity. Then, multiplying the two signals we get an output signal as follows:

$$V_c(t) \times V_m(t) = A \sin \omega_c T + AM (\sin \omega_c T)(\sin \omega_m T) = \text{resultant signal}$$

Using a trigonometric identity from your high school trigonometry book that which says that the product of two sines of two angles is as follows:

$$\sin X \sin Y = \frac{1}{2} \cos (X-Y) + \frac{1}{2} \cos (X+Y)$$

For simplicity, assume  $A = M = 1$

(This will result in a 1 volt carrier with a 1 volt peak modulating signal)

Substituting, in the trigonometric identity,  $X = \omega_c$  and  $Y = \omega_m$ ,  $A \ \& \ B = 1$

$$V_c(t) \times V_m(t) = \sin \omega_c T + \frac{1}{2} \cos (\omega_c - \omega_m) T + \frac{1}{2} (\cos \omega_c + \omega_m) T$$

This says that we have three components in the resulting signal:

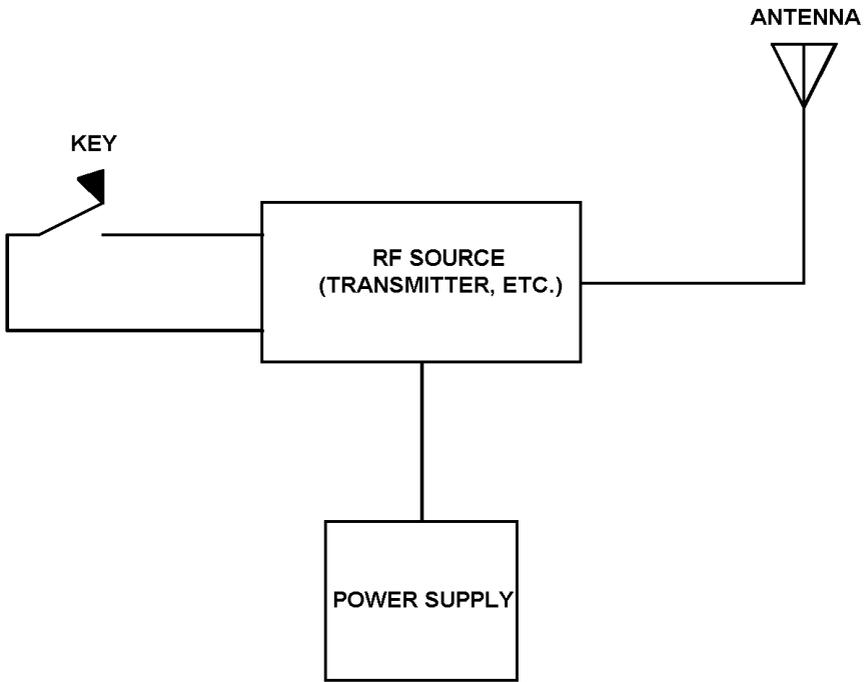
- 1)  $\sin \omega_c T$ , which is a unit level sinewave signal at the carrier frequency
- 2)  $\frac{1}{2} \cos (\omega_c - \omega_m) T$ , which is a half unit level cosinusoidal signal at a frequency equal to the difference between the carrier frequency and the modulating signal frequency. This is called the lower sideband

3)  $\frac{1}{2} \cos (\omega c + \omega m)T$ , which is a half unit level cosinusoidal signal at a frequency equal to the sum of the carrier frequency and the modulating signal frequency. This is called the upper sideband

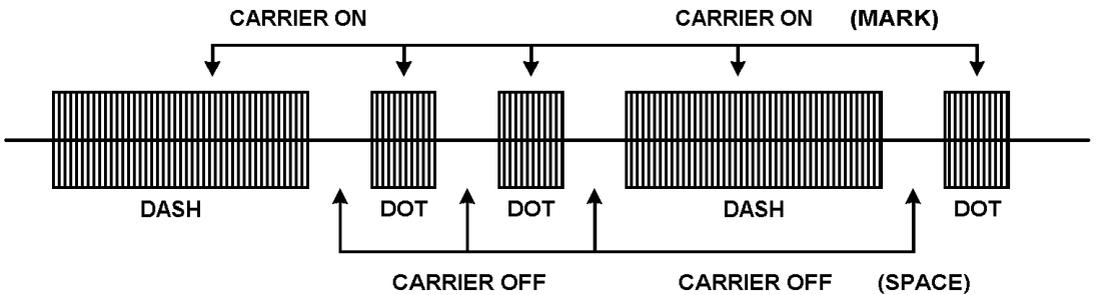
The three signals produced are the carrier, the lower sideband, and the upper sideband. A cosinusoidal waveform is just a sine waveform shifted in phase by 90 degrees, so at  $T = 0$  it is maximum, falling to zero at  $\omega T = 90$  degrees. A sinewave starts at zero at  $T=0$  and has a maximum at 90 degrees. Note that the two sidebands are one half that of the carrier in amplitude, and are different in frequency from the carrier by the modulating frequency. There is also a 90 degree phaseshift. The ratio of the modulating signal to its peak value at full modulation is called the modulation index and is denoted by the letter  $M$ .  $M$  has a value between zero (no modulation) and 1 (maximum modulation) If  $M$  exceeds 1 this is called overmodulation and results in distortion. But the important thing to see is that the total signal bandwidth needed to pass these three components is twice the modulation frequency. It does not depend on the value of  $M$ . Therefore, for a standard AM broadcast signal with a maximum modulating frequency of 5 kHz, a 10 kHz bandwidth is required in the receiver. Also note, since the carrier term is simply a constant amplitude sine wave, it carries no intelligence and its amplitude is constant. Now comes the big kicker: Note that the amplitude of each sideband is only half of that of the carrier even when  $M = 1$ . Therefore the power in each sideband when  $M=1$  is only one quarter that of the carrier. Since there are two sidebands, there is a total sideband energy of only half that of the carrier. Since these sidebands are identical, differing only in frequency by twice the modulating frequency, they both carry the same information and are redundant from an information viewpoint. The sidebands contain only 1/3 the total signal power generated by the transmitter, but they carry all the information, and really, only one is needed, the other being redundant. The modulating system must supply this sideband energy, half the power of the carrier signal if  $M = 1$ . The modulating power needed is equal to one half  $M$  squared. A 1000 watt AM carrier for example, needs 500 watts of audio to fully modulate it. Well, then why not generate the AM signal at low level and amplify it? Not very efficient. Since the total peak amplitude of the signal is twice that of the carrier, a peak power of 4000 watts is present in a 1000 watt AM signal. Therefore a power amplifier used for AM must be capable of delivering 4 times the carrier power on modulation peaks. The 4000 watt amplifier is delivering only a 1000 watt carrier, and seldom operates at full power except on modulation peaks. The overall efficiency is then low. As is always true in real life, you do not get something for nothing. The alternative to a 500 watt modulator and a 1000 watt RF amplifier in this case is a low level audio amplifier and a 4000 watt RF amplifier running inefficiently. Not that this is so bad, because at high power levels it has the advantage of eliminating the expensive and heavy 500 watt modulation transformer needed to couple the audio energy to the transmitter power amplifier. No matter how you look at it or do it, AM is a ripoff from an efficiency standpoint. But it is simple to do, has fairly good audio fidelity, and still has better weak signal performance over certain other modulation methods. It is easily received with a simple low cost receiver, and is not critical as to receiver mistuning. AM is still used

**worldwide for short, medium, and long wave broadcasting, and for air to ground VHF voice communications.**

**It was realized in the early days of radio that since only one of the sidebands is needed, why bother to transmit the carrier and the other sideband? The carrier doesn't "carry" anything, as both the sidebands are RF and can be radiated by an antenna. Getting rid of the carrier and one sideband gets rid of 5/6 of the radiated power with no loss of information. So the transmitter power can be effectively increased by a factor of six, since all the energy can be placed in the transmitted sideband. Furthermore, the receiver bandwidth can be reduced by a factor of two. This gives a total of 8 dB transmitter gain and 3 dB receiver sensitivity, or 11 dB improvement in signal to noise ratio. The likelihood of interference to or from other signals is also reduced by using half the bandwidth, and channel capacity of a frequency band can be doubled, since each signal needs only half the bandwidth of an AM signal. This modified form of AM modulation is called single sideband, or SSB. This will be discussed in the next part of this article.**

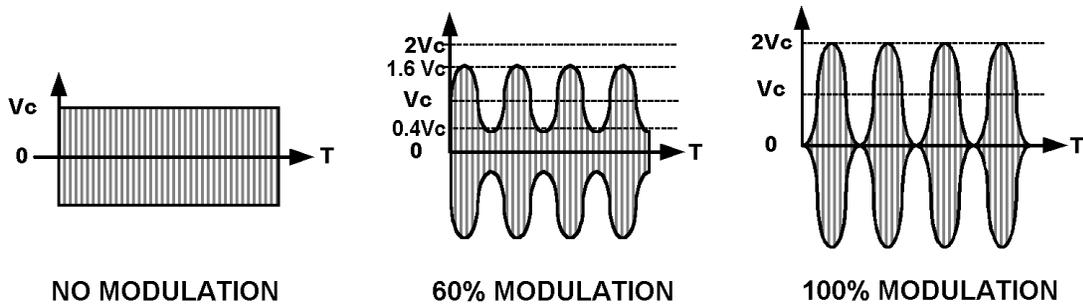
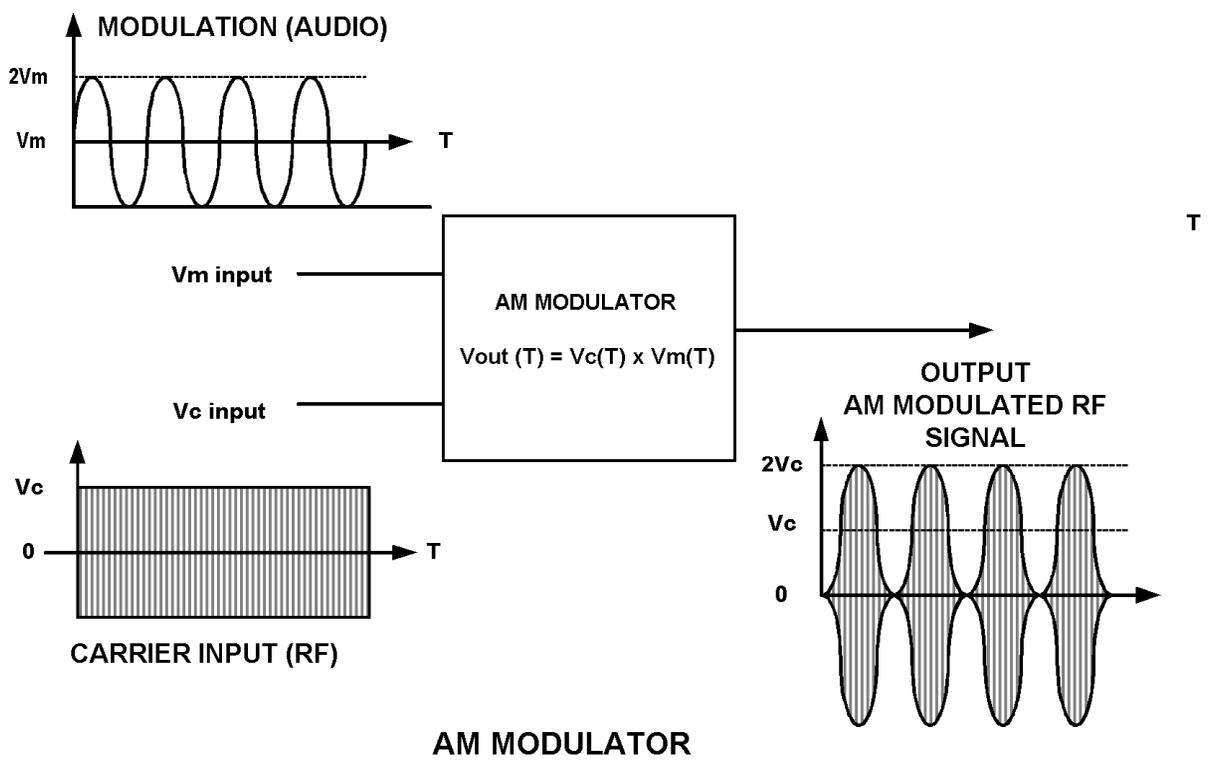


**SIMPLE CW (CONTINUOUS WAVE) TRANSMITTER**



**SIGNAL PRODUCED BY CW TRANSMITTER**

**FIGURE 1  
CONTINUOUS WAVE (CW) TRANSMISSION**



MODULATOR OUTPUT WAVEFORMS

FIG 2 AMPLITUDE MODULATION (AM)