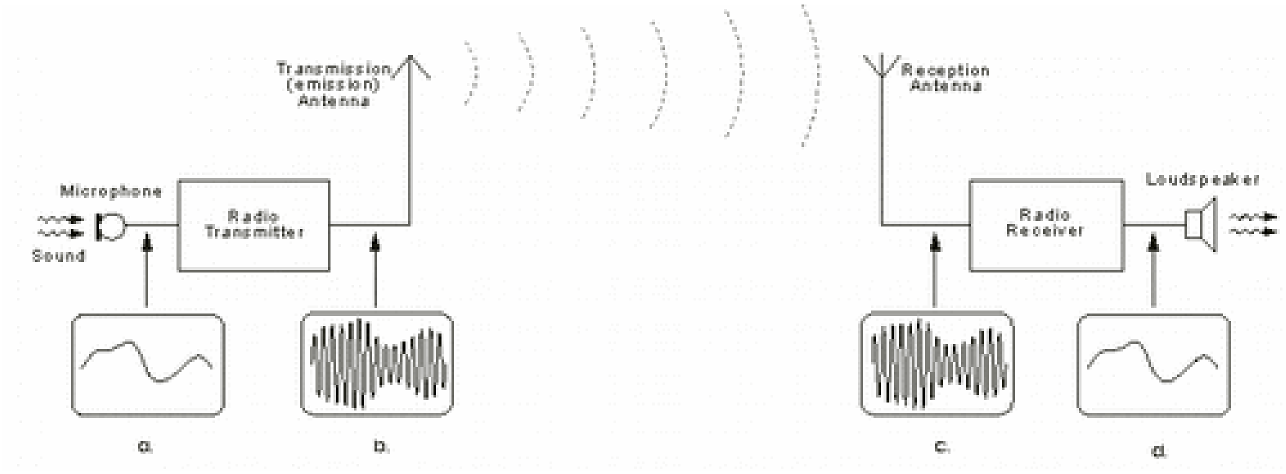
Amateur Wireless Station Operators License Exam

Study material 2017

CHAPTER 4

Chapter 4:Radio Receivers

There are many natural sources of radio waves. But in the later part of the 19th century, scientists figured out how to electronically generate radio waves using electric currents. Two components are required for radio communication: a transmitter and a receiver.



Radio receivers

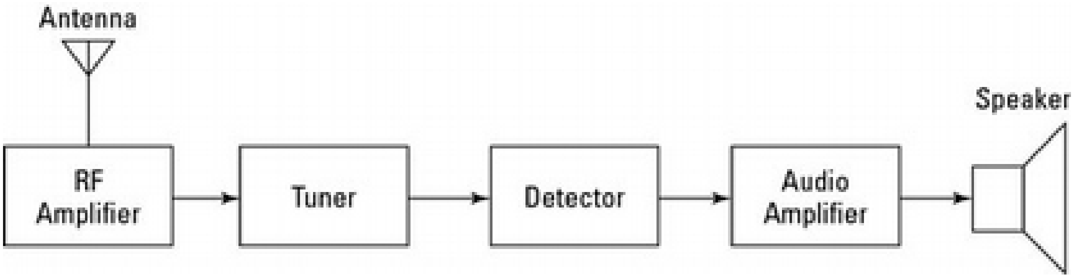
A radio receiver is the opposite of a radio transmitter. It uses an antenna to capture radio waves, processes those waves to extract only those waves that are vibrating at the desired frequency, extracts the audio signals that were added to those waves, amplifies the audio signals, and finally plays them on a speaker.

The pioneers of radio were Popov and Marconi, but the place of honor belongs to Nikola Tesla, who demonstrated wireless broadcasting in 1893, at the Franklin Institute.

Tesla's idea was to produce electromagnetic waves by means of oscillatory circuits and transmit them over an antenna. A receiver would then receive the waves with another antenna and oscillatory circuit being in resonance with the oscillatory circuit of the transmitter. This represented the groundwork of today's radio communications.

In 1904 John Flemming created the diode, and in 1907 Lee De Forest invented the triode. That year can be considered the birth of electronics, with the triode being the first electronic component used in a circuit for signal amplification.

Basic building blocks of a radio receiver.



Antenna: Captures the radio waves. Typically, the antenna is simply a length of wire. When this wire is exposed to radio waves, the waves induce a very small alternating current in the antenna.

RF amplifier: A sensitive amplifier that amplifies the very weak radio frequency (RF) signal from the antenna so that the signal can be processed by the tuner.

Tuner: A circuit that can extract signals of a particular frequency from a mix of signals of different frequencies. On its own, the antenna captures radio waves of all frequencies and sends them to the RF amplifier, which dutifully amplifies them all.

Unless you want to listen to every radio channel at the same time, you need a circuit that can pick out just the signals for the channel you want to hear. That's the role of the tuner.

The tuner usually employs the combination of an inductor (for example, a coil) and a capacitor to form a circuit that resonates at a particular frequency. This frequency, called the resonant frequency, is determined by the values chosen for the coil and the capacitor. This type of circuit tends to block any AC signals at a frequency above or below the resonant frequency.

You can adjust the resonant frequency by varying the amount of inductance in the coil or the capacitance of the capacitor. In simple radio receiver circuits, the tuning is adjusted by varying the number of turns of wire in the coil. More sophisticated tuners use a variable capacitor (also called a tuning capacitor) to vary the frequency.

Detector: Responsible for separating the audio information from the carrier wave. For AM signals, this can be done with a diode that just rectifies the alternating current signal. What's left after the diode has its way with the alternating current signal is a direct current signal that can be fed to an audio amplifier circuit. For FM signals, the detector circuit is a little more complicated.

Audio amplifier: This component's job is to amplify the weak signal that comes from the detector so that it can be heard. This can be done using a simple transistor amplifier circuit.

Of course, there are many variations on this basic radio receiver design. Many receivers include additional filtering and tuning circuits to better lock on to the intended frequency — or to produce better-quality audio output — and exclude other signals. Still, the above discussed are the basic elements found in most receiver circuits.

Types of receivers

Radio receivers may be divided into several categories as the following

(a) crystal (b) regenerative (c) super-regenerative (d) tuned-radio—frequency (e) super-heterodyne(f) synchrodyne.

(a) Crystal

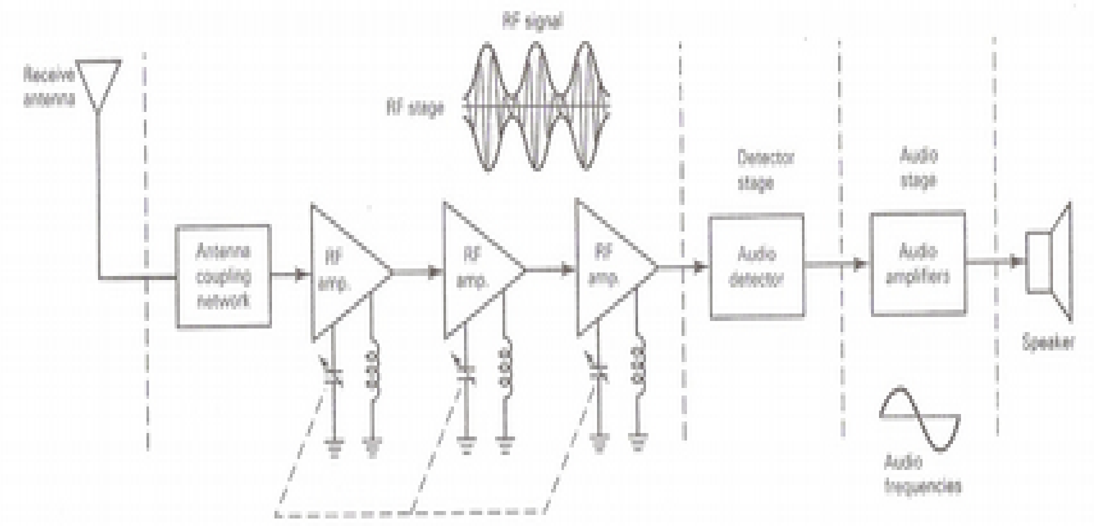
The simplest type of receiver (We studied about in chapter 3) employs a crystal such as galena for a detector. This, together with a suitable tuned circuit and a pair of headphones, forms a satisfactory local station receiving set. Its disadvantages are poor sensitivity and selectivity, together with low output. Modern developments have made available " fixed " germanium diodes which are being used satisfactorily as detectors.

(b) Regenerative receivers

The regenerative circuit (or regen) allows an electronic signal to be amplified many times by the same active device. It consists of an amplifying vacuum tube or transistor with its output connected to its input through a feedback loop, providing positive feedback. This circuit was widely used in radio receivers, called regenerative receivers, between 1915 and World War II. The regenerative receiver was invented in 1912 and patented in 1914 by American electrical engineer Edwin Armstrong when he was an undergraduate at Columbia University. The regenerative receiver was superseded by other receiver designs like the TRF and superheterodyne receivers and became obsolete, but regeneration (now called positive feedback) is widely used in other areas of electronics, such as in oscillators and active filters.

(d) TRF receiver

The T.R.F. (tuned radio frequency) receiver was among the first designs available in the early days when means of amplification by valves became available. The basic principle was that all RF stages simultaneously tuned to the received frequency before detection and subsequent amplification of the audio signal.



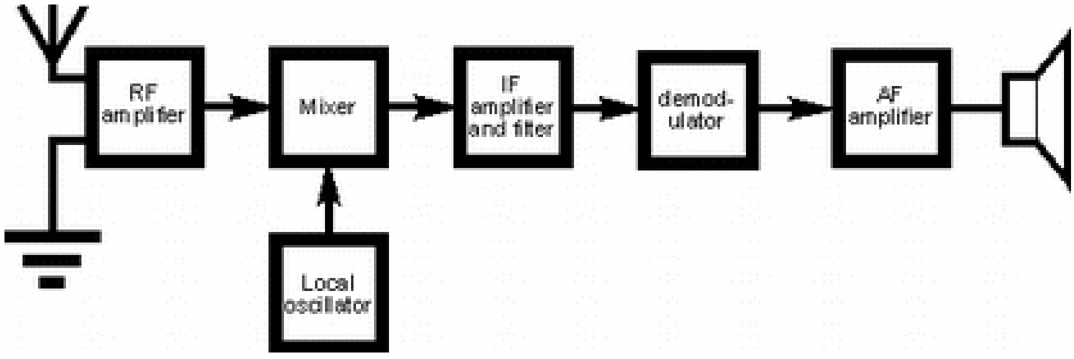
The principle disadvantages were (a) all RF stages had to track one another and this is quite difficult to achieve technically, also (b) because of design considerations, the received bandwidth increases with frequency. a further disadvantage (c) was the shape factor could only be quite poor.

(e) Super-heterodyne Receiver

The super-heterodyne principle was devised in 1918 by U.S. Army Major Edwin Armstrong in France during World War I. He invented this receiver as a means of overcoming the deficiencies of early vacuum tube triodes used as high-frequency amplifiers in radio direction finding equipment. Unlike simple radio communication, which only needs to make transmitted signals audible, direction-finders measure the received signal strength, which necessitates linear amplification of the actual carrier wave.

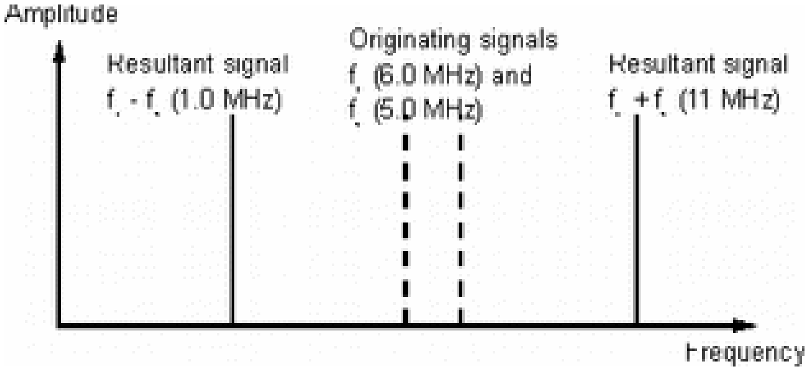
One of the most common forms of radio receiver is the superhet or super-heterodyne radio receiver. Virtually all broadcast radio receivers, as well as televisions, short wave receivers and commercial radios use the super-heterodyne principle as the basis of their operation.

Basic block diagram of a super-heterodyne receiver



Mixing and the superhet receiver

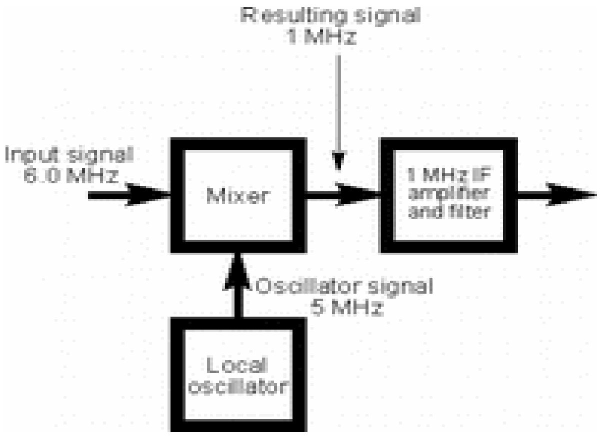
The idea of the super heterodyne receiver revolves around the process of mixing. Here RF mixers are used to multiply two signals together. (This is not the same as mixers used in audio desks where the signals are added together). When two signals are multiplied together the output is the product of the instantaneous level of the signal at one input and the instantaneous level of the signal at the other input. It is found that the output contains signals at frequencies other than the two input frequencies. New signals are seen at frequencies that are the sum and difference of the two input signals, i.e. if the two input frequencies are f1 and f2, then new signals are seen at frequencies of (f1+f2) and (f1-f2). To take an example, if two signals, one at a frequency of 5 MHz and another at a frequency of 6 MHz are mixed together then new signals at frequencies of 11 MHz and 1 MHz are generated.



Concept of the super-heterodyne receiver

In the superhet radio, the received signal enters one inputs of the mixer. A locally generated signal (local oscillator signal) is fed into the other. The result is that new signals are generated. These are applied to a fixed frequency intermediate frequency (IF) amplifier and filter. Any signals that are converted down and then fall within the pass-band of the IF amplifier will be amplified and passed on to the next stages. Those that fall outside the pass-band of the IF are rejected. Tuning is accomplished very simply by varying the frequency of the local oscillator. The advantage of this process is that very selective fixed frequency filters can be used and these far out perform any variable frequency ones. They are also normally at a lower frequency than the incoming signal and again this enables their performance to be better and less costly.

To see how this operates in reality take the example of two signals, one at 6 MHz and another at 6.1 MHz. Also take the example of an IF situated at 1 MHz. If the local oscillator is set to 5 MHz, then the two signals generated by the mixer as a result of the 6 MHz signal fall at 1 MHz and 11 MHz. Naturally the 11 MHz signal is rejected, but the one at 1 MHz passes through the IF stages. The signal at 6.1 MHz produces a signal at 1.1 MHz (and 11.1 MHz) and this falls outside bandwidth of the IF so the only signal to pass through the IF is that from the signal on 6 MHz.

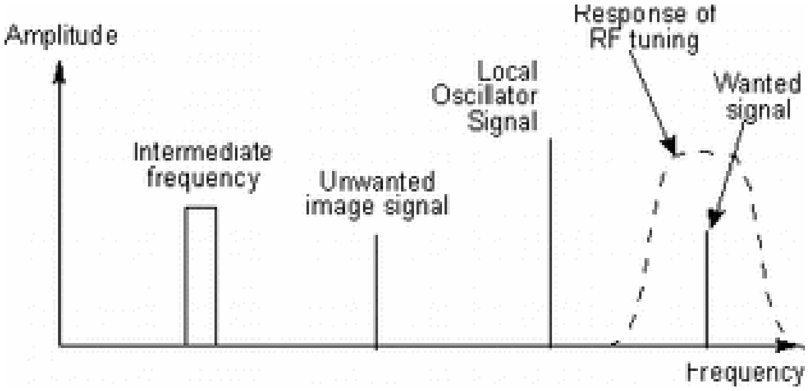


Basic superhet or super-heterodyne radio receiver concept

If the local oscillator frequency is moved up by 0.1 MHz to 5.1 MHz then the signal at 6.1 MHz will give rise to a signal at 1 MHz and this will pass through the IF. The signal at 6 MHz will give rise to a signal of 0.9 MHz at the IF and will be rejected. In this way the receiver acts as a variable frequency filter, and tuning is accomplished.

Image responses

The basic concept of the super-heterodyne receiver appears to be fine, but there is a problem. There are two signals that can enter the IF. With the local oscillator set to 5 MHz and with an IF it has already been seen that a signal at 6 MHz mixes with the local oscillator to produce a signal at 1 MHz that will pass through the IF filter. However if a signal at 4 MHz enters the mixer it produces two mix products, namely one at the sum frequency which is 10 MHz, whilst the difference frequency appears at 1 MHz. This would prove to be a problem because it is perfectly possible for two signals on completely different frequencies to enter the IF. The unwanted frequency is known as the image. Fortunately it is possible to place a tuned circuit before the mixer to prevent the signal entering the mixer, or more correctly reduce its level to an acceptable value.



Fortunately this tuned circuit does not need to be very sharp. It does not need to reject signals on adjacent channels, but instead it needs to reject signals on the image frequency. These will be separated from the wanted channel by a frequency equal to twice the IF. In other words with an IF at 1 MHz, the image will be 2 MHz away from the wanted frequency.

Advantages of super-heterodyne receivers

The advantages of super-heterodyne receiver are many. An obvious advantage is that by reducing to lower frequency, lower frequency components can be used, and in general, cost is proportional to frequency. RF gain at 40 GHz is expensive, IF gain at 1 GHz is cheap as dirt.

The second advantage is in the superior sensitivity that we almost take for granted. Filtering out unwanted signals at IF is a much easier job than filtering them out at RF, because the desired bandwidth is much higher after the signal is mixed down.

Further advantage in that many components can be designed for a fixed frequency (and even shared between different receiver designs), which is easier and cheaper than designing wide band components.

Basic performance characteristics of a radio receiver

Sensitivity, noise, selectivity, and ***fidelity*** are important receiver performance characteristics. These characteristics will be useful to you when performing receiver tests. They can help you to determine whether a receiver is working or not or in comparing one receiver to another.

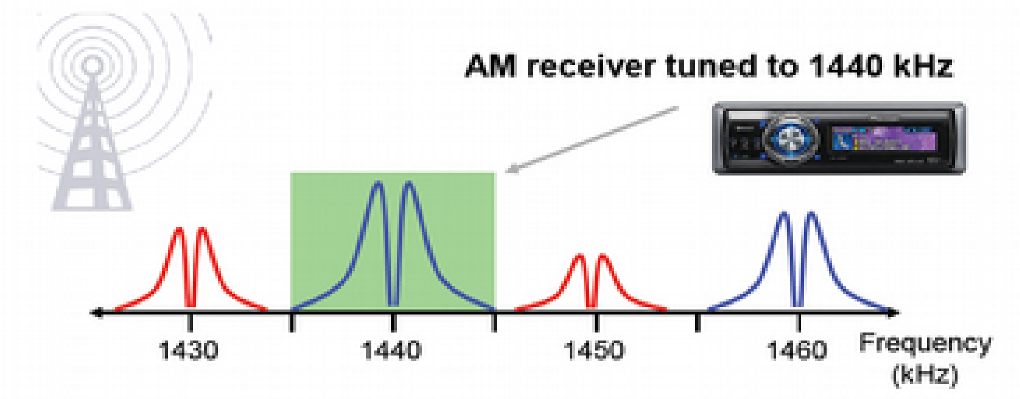
Sensitivity:
The ability of a receiver to reproduce weak signals is a function of the sensitivity of a receiver. The weaker a signal that can be applied to a receiver and still produce a certain value of signal output, the better the sensitivity rating. Sensitivity of a receiver is measured under standardized conditions. It is expressed in terms of the signal voltage, usually in the microvolts that must be

applied to the antenna input terminals to give an established level of the output. The output may be an AC or DC voltage measured at the detector output or a power measurement (measured in decibels or watts) at the loudspeaker or headphone terminals.

Noise:

All receivers generate a certain amount of noise, which you must take into account when measuring sensitivity. Receiver noise may originate from the atmosphere (lightning) or from internal components(transistors, tubes). Noise is the limiting factor of sensitivity. You will find sensitivity is the value of input carrier voltage (in microvolts) that must be applied from the signal generator to the receiver input to develop a specified output power.

Selectivity:



Selectivity is the degree of distinction made by the receiver between the desired signal and unwanted signals. You will find the better the ability of the receiver to reject unwanted signals, the better its selectivity. The degree of selection is determined by the sharpness of resonance to which the frequency-determining circuits have been engineered and tuned. You usually measure selectivity by taking a series of sensitivity readings. As you take the readings, you step the input signal along a band of frequencies above and below the circuit resonance of the receiver; for example, 100 kilohertz below to 100 kilohertz

Signal-to-Noise (S/N) ratio :

The Signal-to-Noise Ratio (S/N) (SNR) in a receiver is the signal power in the receiver divided by the mean noise power of the receiver. All receivers require the signal to exceed the noise by some amount. Usually if the signal power is less than or just equals the noise power it is not detectable. For a signal to be detected, the signal energy plus the noise energy must exceed some threshold value. S/N is a required minimum ratio, if N is increased, then S must also be increased to maintain that threshold. The acceptable minimum Signal-to-Noise ratio (or think of it as Signal above Noise) for a receiver depends on the intended use of the receiver. For instance, a receiver that had to detect a single radar pulse would probably need a higher minimum S/N than a receiver that could integrate numerous radar pulses (increasing the total signal energy) for detection with the same probability of false alarms. Receivers with human operators using a video display may function satisfactorily with low minimum S/N because a skilled operator can be very proficient at picking signals out of a noise background

Automatic volume control

Automatic volume control AVC, or automatic gain control AGC, as it is sometimes called, is a system for keeping the volume from the loudspeaker at a constant level, no matter what the strength of the high-frequency signal strength may be. Manual volume control - ie, by hand - is usually fitted even on sets provided with automatic volume control, since although it may be desirable that all stations should come in at the same strength on the loudspeaker, yet one obviously wants to be able to control that strength to suit personal requirements.

The main object of AVC is to make it unnecessary to keep altering the manual volume control as one passes from station to station. In the ordinary way, one may readily receive a local station at tremendous strength, while a few degrees away on the dial another station may be comparatively weak. As one tunes still farther round the dial, a very powerful station may produce a most unpleasant blare. There is an obvious advantage if all the stations can be brought in at full loudspeaker strength, but no louder and no weaker. For Tuned Radio Frequency TRF or straight sets this is done by altering the sensitivity of the whole set, either on the high-frequency or on the low-frequency side, or on both; while in the case of super-heterodyne, detectors and intermediate frequency amplifiers gain may be adjusted to increase or decrease the sensitivity of the set as may be required.

Squelch

An circuit that is important to proper receiver behavior is called "squelch" or muting. The function of this circuit is to mute or silence the audio output of the receiver in the absence of the desired radio signal. When the desired signal is lost (due to multi-path dropout, excessive distance, loss of power to the transmitter, etc.) the "open" receiver may pick up another signal or background radio "noise." Typically, this is heard as "white" noise and is often much louder than the audio signal from the desired source.

The traditional squelch circuit is an audio switch controlled by the radio signal level using a fixed or manually adjustable threshold (level). When the received signal strength falls below this level the output of the receiver is muted. Ideally, the squelch level should be set just above the background radio noise level or at the point where the desired signal is becoming too noisy to be acceptable. Higher settings of squelch level require higher received signal strength to unmute the receiver. Since received signal strength decreases as transmission distance increases, higher squelch settings will decrease the operating range of the system.

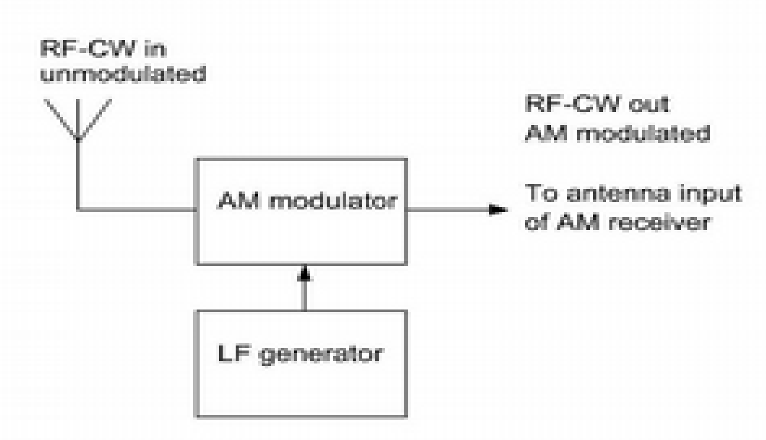
One type of the standard squelch circuit is referred to as "noise squelch." This technique relies on the fact that the audio from undesirable radio noise has a great deal of high frequency energy compared to a typical audio signal. The noise squelch circuit compares the high frequency energy of the received signal to a reference voltage set by the squelch adjustment. In this system the squelch control essentially determines the "quality" of signal (signal-to-noise ratio) required to unmute the receiver. This allows operation at lower squelch settings with no likelihood of noise if the desired signal is lost.

Another type is known as "tone-key" or "tone-code" squelch. It enables the receiver to identify the desired radio signal by means of a supra- or sub-audible tone that is generated in the transmitter and sent along with the normal audio signal. The receiver will unmute only when it picks up a radio signal of adequate strength and also detects the presence of the tone-key. This effectively prevents the possibility of noise from the receiver when the desired transmitter signal is lost, even in the presence of a (non-tone-key) interfering signal at the same frequency.

CW

A continuous wave or continuous waveform (CW) is an electromagnetic wave of constant amplitude and frequency; of infinite duration. Continuous wave is also the name given to an early method of radio transmission, in which a carrier wave is switched on and off. Information is carried in the varying duration of the on and off periods of the signal, for example by Morse code in early radio. In early wireless telegraphy radio transmission, Morse code is still used by radio amateurs for communication.

CW reception in an AM receiver with a Beat Frequency Oscillator (BFO)



Usually, CW is transmitted as an unmodulated carrier. To make it audible, a BFO (Beat Frequency Oscillator) is used. The frequency of the BFO is a few hundred hertz lower or higher than the frequency of the CW signal. In a detector, a beat tone is generated. That is an audio tone with a tone height equal to the difference of the frequencies of the CW signal and the BFO oscillator. This works excellent, sensitivity is very good and all CW signals can be seperated very good as they have different tone heights, as long as their transmit frequencies are different. A disadvantage is that you can receive only a limited frequency range, namely within an audible bandwidth around the BFO frequency. And if the frequency of your receiver is drifting a little, the audio tone changes too.

Types of Modulations

As you will recall, modulation is the process of changing amplitude or frequency or phase of a carrier wave in accordance with the intensity of the signal. Accordingly, there are three basic types of modulation, namely ;

(i) Amplitude modulation (ii) Frequency Modulation (iii) phase modulation

In India, both the amplitude modulation and the frequency modulation are used in radio broadcasting. However, in television transmission, frequency modulation is used for sound signal and amplitude modulation for picture signal.

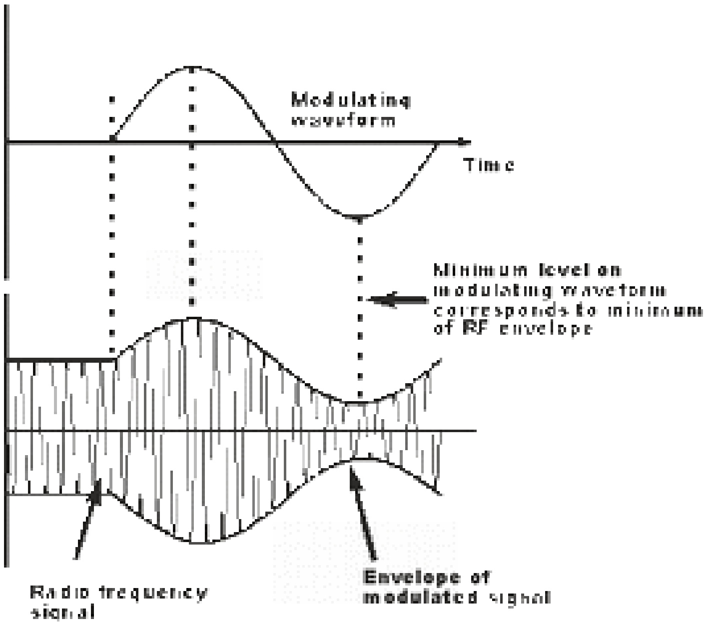
Amplitude Modulation

In order that a steady radio signal or “radio carrier” can carry information it must be changed or modulated in one way so that the information can be conveyed from one place to another. There are a number of ways in which a carrier can be modulated to carry a signal - often an audio signal and the most obvious way is to vary its amplitude.

Amplitude Modulation has been in use since the very earliest days of radio technology. The first recorded instance of its use was in 1901 when a signal was transmitted by a Canadian engineer named Reginald Fessenden. To achieve this, he used a continuous spark transmission and placed a carbon microphone in the antenna lead. The sound waves impacting on the microphone varied its resistance and in turn this varied the intensity of the transmission. Although very crude, signals were audible over a distance of a few hundred metres. The quality of the audio was not good particularly as a result of the continuous rasping sound caused by the spark used for the transmission.
Later, continuous sine wave signals could be generated and the audio quality was greatly improved. As a result, amplitude modulation, AM become the standard for voice transmissions.
Currently amplitude modulation is primarily used for broadcasting, but it is still used for some forms of two way radio communications. Its main radio communications use is for local aviation related VHF two way radio links. It is used for ground to air radio communications as well as two way radio links for ground staff as well.

How AM is generated?

When an amplitude modulated signal is created, the amplitude of the signal is varied in line with the variations in intensity of the sound wave. In this way the overall amplitude or envelope of the carrier is modulated to carry the audio signal. Here the envelope of the carrier can be seen to change in line with the modulating signal.



Amplitude Modulation, AM

Amplitude modulation, AM is the most straightforward way of modulating a signal. Demodulation, or the process where the radio frequency signal is converted into an audio frequency signal is also very simple. An amplitude modulation signal only requires a simple diode detector circuit. The circuit that is commonly used has a diode that rectifies the signal, only allowing the one half of the alternating radio frequency waveform through. A capacitor is used to remove the radio frequency parts of the signal, leaving the audio waveform. This can be fed into an amplifier after which it can be used to drive a loudspeaker. As the circuit used for demodulating AM is very cheap, it enables the cost of radio receivers for AM to be kept low. (Remember the crystal radio in the previous chapter)

Amplitude modulation advantages & disadvantages :
Amplitude modulation has several advantages and disadvantages. These mean that it is used in particular circumstances where its advantages can be used to good effect.

ADVANTAGES	DISADVANTAGES
It is simple to implement	An amplitude modulation signal is not efficient in terms of its power usage
It can be demodulated using a circuit consisting of very few components	It is not efficient in terms of its use of bandwidth, requiring a bandwidth equal to twice that of the highest audio frequency
AM receivers are very cheap as no specialized components are needed.	An amplitude modulation signal is prone to high levels of noise because most noise is amplitude based and obviously AM detectors are sensitive to it.

In view of its characteristics advantages and disadvantages, amplitude modulation is being used less frequently. However it is still in widespread use for broadcasting on the long, medium and short wave bands as well as for a number of mobile or portable communications systems including some aircraft communications.

Amplitude modulation related signals :
Amplitude modulation forms the basis of a number of forms of signal apart from the basic mode.

These signal formats are typically generated by removing or suppressing the carrier, and then utilizing the sidebands. These formats are defined in greater detail in further pages of this tutorial.

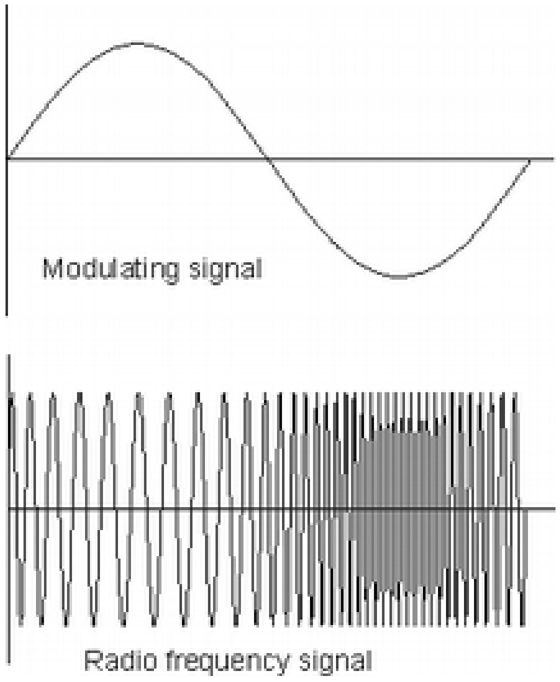
DESCRIPTION	ABBREVIATION	ITU(international telecommunication Union) DESIGNATION
Double sideband, full carrier	DSB full carrier	A3E
Single sideband reduced carrier	SSB reduced carrier	R3E
Single sideband full carrier	SSB full carrier	H3E
Single sideband suppressed carrier	SSBSC	J3E
Independent sideband	ISB	B8E
Vestigial sideband	VSB	C3F

While amplitude modulation is one of the simplest and easiest forms of signal modulation to implement, it is not the most efficient in terms of spectrum efficiency and power usage. As a result, the use of amplitude modulation is falling in preference to other analogue modes such as frequency modulation, and a variety of digital modulation formats. Yet despite this decrease, amplitude modulation is in such widespread use, especially for broadcasting, and many amplitude modulation signals can still be heard on the various long, medium and short wavebands where they will undoubtedly be heard for many years to come.

Frequency modulation.

FM is widely used for a variety of radio communications applications. FM broadcasts on the VHF bands still provide exceptionally high quality audio, and FM is also used for a variety of forms of two way radio communications, and it is especially useful for mobile radio communications, being used in taxis, and many other forms of vehicle.
in view of its widespread use, frequency modulation, FM, is an important form of modulation, despite many forms of digital transmission being used these days.
FM, frequency modulation has been in use for many years. However its advantages were not immediately apparent. In the early days of wireless, it was thought that a narrower bandwidth was required to reduce noise and interference. As FM did not perform well under these conditions, AM predominated and FM was not used. However, Edwin Armstrong, an American engineer looked at the use of wide-band FM for broadcasting and introduced the idea against the trend of the thinking of the time.
Since its first introduction the use of frequency modulation, FM has grown enormously. Now wideband FM is still regarded as a very high quality transmission medium for high quality broadcasting. FM, frequency modulation is also widely used for communications where it is resilient to variations in signal strength.

How FM signals are generated?
The most obvious method of applying modulation to a signal is to superimpose the audio signal onto the amplitude of the carrier. However this is by no means the only method which can be employed. It is also possible to vary the frequency of the signal to give frequency modulation or FM. It can be seen below that the frequency of the signal varies as the voltage of the modulating signal changes.



The amount by which the signal frequency varies is very important. This is known as the deviation and is normally quoted as the number of kilo Hertz deviation. As an example the signal may have a deviation of ±3 kHz. In this case the carrier is made to move up and down by 3 kHz.

Narrow-band FM, NBFM, and Wide-band FM, WBFM :
The level of deviation is important in many aspects. It obviously is important in determining the bandwidth of the overall signal. As a result the deviation used for FM is different between different applications. Broadcast stations in the VHF portion of the frequency spectrum between 88.5 and 108 MHz use large values of deviation, typically ±75 kHz. This is known as wide band FM (WBFM). These signals are capable of supporting high quality transmissions, but occupy a large amount of bandwidth. Usually 200 kHz is allowed for each wide-band FM transmission. For radio communications purposes less bandwidth is used. Narrow-band FM, NBFM often uses deviation figures of around ±3 kHz or possibly slightly more. As quality is not as important for radio communications applications, the much narrower bandwidth has advantages in terms of radio spectrum efficiency.

Improvement in Signal to Noise Ratio :
It has already been mentioned that FM can give a better signal to noise ratio than AM when wide bandwidths are used. The amplitude noise can be removed by limiting the signal to remove it. In fact the greater the deviation the better the noise performance. When comparing an AM signal to an FM one an improvement equal to 3 D2 is obtained where D is the deviation ratio.

Highlights
Frequency modulation is used in a wide variety or radio communications applications from broadcasting to two way radio communications links as well as mobile radio communications. It possesses many advantages over amplitude modulation and this is the reason for its widespread use. Nowadays, many digital forms of radio communications are being introduced, but despite this the use of frequency modulation, FM will undoubtedly continue for many years to come in many areas of radio communications.

Advantages of frequency modulation
There are many advantages to the use of frequency modulation. These have meant that it has been widely used for many years, and will remain in use for many years.

- **Resilient to noise:** One of the main advantages of frequency modulation that has been utilized by the broadcasting industry is the reduction in noise. As most noise is amplitude based, this can be removed by running the signal through a limiter so that only frequency variations appear. This is provided that the signal level is sufficiently high to allow the signal to be limited.

- **Resilient to signal strength variations:** In the same way that amplitude noise can be removed, so too can any signal variations. This means that one of the advantages of frequency modulation is that it does not suffer audio amplitude variations as the signal level varies, and it makes FM ideal for use in mobile applications where signal levels constantly vary. This is provided that the signal level is sufficiently high to allow the signal to be limited.
- **Does not require linear amplifiers in the transmitter:** As only frequency changes are required to be carried, any amplifiers in the transmitter do not need to be linear.
- **Enables greater efficiency than many other modes:** The use of non-linear amplifiers, e.g. class C, etc means that transmitter efficiency levels will be higher - linear amplifiers are inherently inefficient.

Dis-advantages of frequency modulation

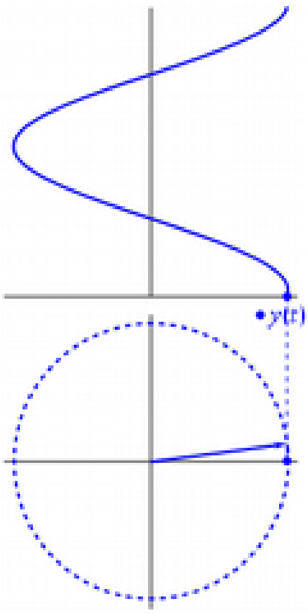
There are a number of dis-advantages to the use of frequency modulation. Some are can be overcome quite easily, but others may mean that another modulation format is more suitable.

- **Requires more complicated demodulator:** One of the minor dis-advantages of frequency modulation is that the demodulator is a little more complicated, and hence slightly more expensive than the very simple diode detectors used for AM. Also requiring a tuned circuit adds cost. However this is only an issue for the very low cost broadcast receiver market.
- **Some other modes have higher data spectral efficiency:** Some phase modulation and quadrature amplitude modulation formats have a higher spectral efficiency for data transmission that frequency shift keying, a form of frequency modulation. As a result, most data transmission system use PSK and QAM.
- **Sidebands extend to infinity either side:** The sidebands for an FM transmission theoretically extend out to infinity. To limit the bandwidth of the transmission, filters are used, and these introduce some distortion of the signal.

Even In spite of the above said dis-advantages, frequency modulation - it is still widely used for many broadcasts and radio communications applications. However with more systems using digital formats, phase and quadrature amplitude modulation formats are on the increase.

Phase modulation (PM)

Phase modulation (PM) is a modulation pattern that encodes information as variations in the instantaneous phase of a carrier wave. Modification in phase according to low frequency yields phase modulation. Unlike frequency modulation (FM), phase modulation is not widely used for broadcasting stations. This is because PM requires more complex receiving hardware, and there can be ambiguity problems in determining whether, for example, the signal has changed phase by +180° or -180°. But the phase modulation systems are mainly used for digital communication devises.



Phase modulation is also the basis for many forms of digital modulation based around phase shift keying, PSK which is a form of phase modulation. As various forms of phase shift keying are the favored form of modulation for digital or data transmissions, this makes phase modulation particularly important.

Phase modulation basics

Before looking at phase modulation it is first necessary to look at phase itself. A radio frequency signal consists of an oscillating carrier in the form of a sine wave is the basis of the signal. The instantaneous amplitude follows this curve moving positive and then negative, returning to the start point after one complete cycle - it follows the curve of the sine wave. This can also be represented by the movement of a point around a circle, the phase at any given point being the angle between the start point and the point on the waveform as shown.

Phase modulation works by modulating the phase of the signal, i.e. changing the rate at which the point moves around the circle. This changes the phase of the signal from what it would have been if no modulation was applied. In other words the speed of rotation around the circle is modulated about the mean value. To achieve this it is necessary to change the frequency of the signal for a short time. In other words when phase modulation is applied to a signal there are frequency changes and vice versa. Phase and frequency are inseparably linked as phase is the integral of frequency. Frequency modulation can be changed to phase modulation by simply adding a CR network to the modulating signal that integrates the modulating signal. As such the information regarding sidebands, bandwidth and the like also hold true for phase modulation as they do for frequency modulation, bearing in mind their relationship.

Forms of phase modulation

Although phase modulation is used for some analogue transmissions, it is far more widely used as a digital form of modulation where it switches between different phases. This is known as phase shift keying, PSK, and there are many flavors of this. It is even possible to combine phase shift keying and amplitude keying in a form of modulation known as quadrature amplitude modulation, QAM.

The list below gives some of the forms of phase shift keying that are used:
PM - Phase Modulation
PSK - Phase Shift Keying
BPSK - Binary Phase Shift Keying
QPSK - Quadrature Phase Shift Keying
8 PSK - 8 Point Phase Shift Keying
16 PSK - 16 Point Phase Shift Keying
QAM - Quadrature Amplitude Modulation
16 QAM - 16 Point Quadrature Amplitude Modulation
64 QAM - 64 Point Quadrature Amplitude Modulation
MSK - Minimum Shift Keying
GMSK - Gaussian filtered Minimum Shift Keying

These are just some of the major forms of phase modulation that are widely used in radio communications applications today. With today's highly software adaptable radio communications systems, it is possible to change between the different types of modulation to best meet the prevailing conditions.