UNIT-I

Pulse characteristics in a Radar system.

Radar is an electromagnetic system for the detection and location of objects. It operates by transmitting a particular type of waveform, a pulse-modulated sine wave for example, and detects the nature of the echo signal. Radar is used to extend the capability of one's senses for observing the environment, especially the sense of vision. The value of radar lies not in being a substitute for the eye, but in doing what the eye cannot do Radar cannot resolve detail as well the eye, nor is it capable of recognizing the "color" of objects to the degree of sophistication which the eye is capable. However, radar can be designed to see through those conditions impervious to normal human vision, such as darkness, haze, fog, rain, and snow. In addition, radar has the advantage of being able to measure the distance or range to the object. This is probably its most important attribute.

An elementary form of radar consists of a transmitting antenna emitting electromagnetic radiation generated by an oscillator of some sort, a receiving antenna, and an energy-detecting device or receiver. A portion of the transmitted signal is intercepted by a reflecting object (target) and is reradiated in all directions. It is the energy reradiated in the back direction that is of prime interest to the radar. The receiving antenna collects the returned energy and delivers it to a receiver, where it is processed to detect the presence of the target and to extract its location and relative velocity. The distance to the target is determined by measuring the time taken for the radar signal to travel

To the target and back. The direction, or angular position, of the target may be determined from the direction of arrival of the reflected wave front. The usual method of measuring the direction of arrival is with narrow antenna beams. If relative motion exists between target and radar, the shift in the carrier frequency of the reflected wave (Doppler effect) is a measure of the target's relative (radial) velocity and may be used to distinguish moving targets from stationary objects. In radars which continuously track the movement of a target, a continuous indication of the rate of change of target position is also available.

The name radar reflects the emphasis placed by the early experimenters on a device to detect the presence of a target and measure its range. Radar is a contraction of the words radio detection and ranging. It was first developed as a detection device to warn of the approach of hostile aircraft and for directing antiaircraft weapons. Although a well-designed modern radar can usually extract more information from the target signal than merely range the measurement of range is still one of radar's most important functions. There seem to be no other competitive techniques which can measure range as well or as rapidly as can radar.

The most common radar waveform is a train of narrow, rectangular-shape pulses modulating a sine wave carrier. The distance, or range, to the target is determined by measuring the time T_R taken by the pulse to travel to the target and return. Since electromagnetic energy propagates at the speed of light c = 3 $x10^8$ m/s, the range R is $R=CT_R/2$. The factor 2 appears in the denominator because of the two-way propagation of radar. With the range in kilometers or nautical miles, and T_R in microseconds, Eq. above becomes

$$R(\text{km}) = 0.15T_R(\mu s) \qquad \text{or} \qquad R(\text{nmi}) = 0.081T_R(\mu s)$$

Each microsecond of round-trip travel time corresponds to a distance of 0.081 nautical mile, 0.093 statute mile, 150 meters, 164 yards, or 492 feet. Once the transmitted pulse is emitted by the radar, a sufficient length of time must elapse to allow any echo signals to return and be detected before the next pulse may be transmitted. Therefore the rate at which the pulses may be transmitted is determined by the longest range at which targets are expected. If the pulse repetition frequency is too high, echo signals from some targets might arrive after the transmission of the next pulse, and ambiguities in measuring range might result.

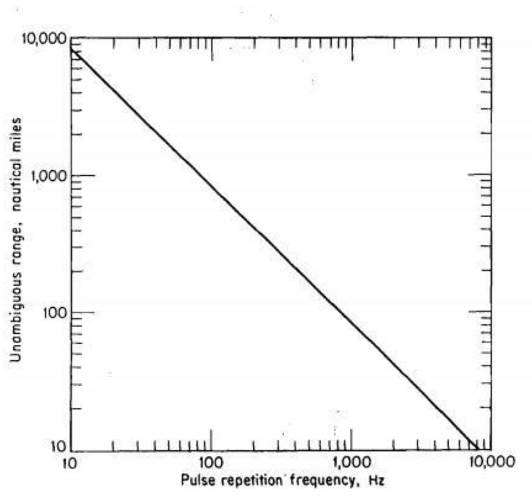


Fig.1.1 Plot of maximum unambiguous range as a function of the pulse repetition frequency

Echoes that arrive after the transmission of the next pulse are called second-time-around (multiple-time-around) echoes. Such an echo would appear to be at a much shorter range than the actual and could be misleading if it were not known to be a second-time-around echo. The range beyond which targets appear as second-time-around echoes is called the maximum unambiguous range and is

$$R_{unamb} = \frac{c}{2f_p}$$

Where f_p = pulse repetition frequency, in Hz. A plot of the maximum unambiguous range as a function of pulse repetition frequency is shown in figure1.1.

Radar Range Equation

The radar equation relates the range of radar to the characteristics of the transmitter, receiver, antenna, target, and environment. It is useful not just as a means for determining the maximum distance from the radar to the target, but it can serve both as a tool for under-standing radar operation and as a basis for radar design.

If the power of the radar transmitter is denoted by P_t , and if an isotropic antenna is used (one which radiates uniformly in all directions), the power density (watts per unit area) at a distance R from the radar is equal to the transmitter power divided by the surface area $4\pi R^2$ of an imaginary sphere of radius R, or

$$\text{Power density from isotropic antenna} = \frac{P_t}{4\pi R^2}$$

Radars employ directive antennas to channel, or direct, the radiated power P_t into some particular direction. The gain G of an antenna is a measure of the increased power radiated in the direction of the target as compared with the power that would have been radiated from an isotropic antenna. It may be defined as the ratio of the maximum radiation intensity from the subject antenna to the radiation intensity from a lossless, isotropic antenna with the same power input. (The radiation intensity is the power radiated per unit solid angle in a given direction.) The power density at the target from an antenna with a transmitting gain G is

$$\text{Power density from directive antenna} = \frac{P_t G}{4\pi R^2}$$

The target intercepts a portion of the incident power and reradiates it in various directions.

The measure of the amount of incident power intercepted by the target and reradiated back in the direction of the radar is denoted as the radar cross section σ, and is defined by the relation

$$\text{Power density of echo signal at radar} = \frac{P_t G}{4\pi R^2}\frac{\sigma}{4\pi R^2}$$

The radar cross section σ has units of area. It is a characteristic of the particular target and is a measure of its size as seen by the radar. The radar antenna captures a portion of the echo power. If the effective area of the receiving antenna is denoted A_e, the power P_r, received by the radar is

$$P_r = \frac{P_t G}{4\pi R^2}\frac{\sigma}{4\pi R^2} A_e = \frac{P_t G A_e \sigma}{(4\pi)^2 R^4}$$

The maximum radar range R_{max} is the distance beyond which the target cannot be detected. It occurs when the received echo signal power P, just equals the minimum detectable signal S_{min},

Therefore

$$R_{max} = \left[\frac{P_t G A_e \sigma}{(4\pi)^2 S_{min}}\right]^{1/4}$$

This is the fundamental form of the radar equation. Note that the important antenna parameters are the transmitting gain and the receiving effective area.

Antenna theory gives the relationship between the transmitting gain and the receiving effective area of an antenna as

$$G = \frac{4\pi A_e}{\lambda^2}$$

Since radars generally use the same antenna for both transmission and reception, Eq. can be substituted into Eq. above, first for A_e, then for G, to give two other forms of the radar equation

$$R_{max} = \left[\frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 S_{min}}\right]^{1/4}$$

$$R_{max} = \left[\frac{P_t A_e^2 \sigma}{4\pi \lambda^2 S_{min}}\right]^{1/4}$$

Radar Block Diagram and Operation:

The operation of typical pulse radar may be described with the aid of the block diagram shown in firure1.2. The transmitter may be an oscillator, such as a magnetron, that is "pulsed" (turned on and on) by the modulator to generate a repetitive train of pulses. The magnetron has probably been the most widely used of the various microwave generators for radar. A typical radar for the detection of aircraft at ranges of 100 or 200 nmi might employ a peak power of the order of a megawatt, an average power of several kilowatts, a pulse width of several microseconds, and a pulse repetition frequency of several hundred pulses per second.

The waveform generated by the transmitter travels via a transmission line to the antenna, where it is radiated into space. A single antenna is generally used for both transmitting and receiving. The receiver must be protected from damage caused by the high power of the transmitter. This is function of the duplexer. The duplexer also serves to channel the returned echo signals to the receiver and not to the transmitter. The duplexer might consist of two gas-discharge devices, one known as a TR (transmit-receive) and the other an ATR (anti-transmit- receive). The TR protects the receiver during transmission and the ATR directs the echo signal to the receiver during reception. Solid-state ferrite circulators and receiver protectors with gas-plasma TR devices and/or diode limiters are also employed as duplexers.

The receiver is usually of the super heterodyne type. The first stage might be a low-noise RF amplifier, such as a parametric amplifier or a low-noise transistor. However, it is not always desirable to employ a low-noise first stage in radar. The receiver input can simply be the mixer stage, especially in military radars that must operate in a noisy environment. Although a receiver with a low-noise front-end will be more sensitive, the mixer input can have greater dynamic range, less susceptibility to overload, and less vulnerability to electronic interference.

The mixer and local oscillator (LO) convert the RF signal to an intermediate frequency (IF). A " typical" IF amplifier for an air-surveillance radar might have a center frequency of 30 or 60 MHz and a bandwidth of the order of one megahertz. The IF amplifier should be designed as a notched filter; i.e., its frequency-response function H (f) should maximize the peak-signal-to-mean-noise-power ratio at the output. This occurs when the magnitude of the frequency-response function $|H(f)|$ is equal to the magnitude of the echo signal spectrum $|S(f)|$, and the phase spectrum of the matched filter is the negative of the phase spectrum of the echo signal . In a radar whose signal waveform approximates a rectangular pulse, the conventional IF filter bandpass characteristic approximates a matched filter when the product of the IF bandwidth B and the pulse width τ is of the order of unity, that is, $B\tau \approx 1$.

After maximizing the signal-to-noise ratio in the IF amplifier, the pulse modulation is extracted by the second detector and amplified by the video amplifier to a level where it can be properly displayed, usually on a cathode-ray tube (CRT).

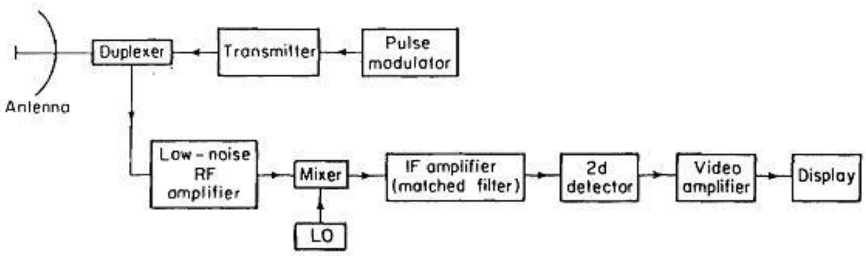


Fig.1.2 Block diagram of a pulse radar

Timing signals are also supplied to the indicator to provide the range zero. Angle information is obtained from the pointing direction of the antenna. The most common form of cathode-ray tube display is the plan position indicator, or PPI which maps in polar coordinates the location of the target in azimuth and range. This is an intensity-modulated display in which the amplitude of the receiver output modulates the electron-beam intensity (z axis) as the electron beam is made to sweep outward from the centre of the tube. The beam rotates in angle in response to the antenna position.

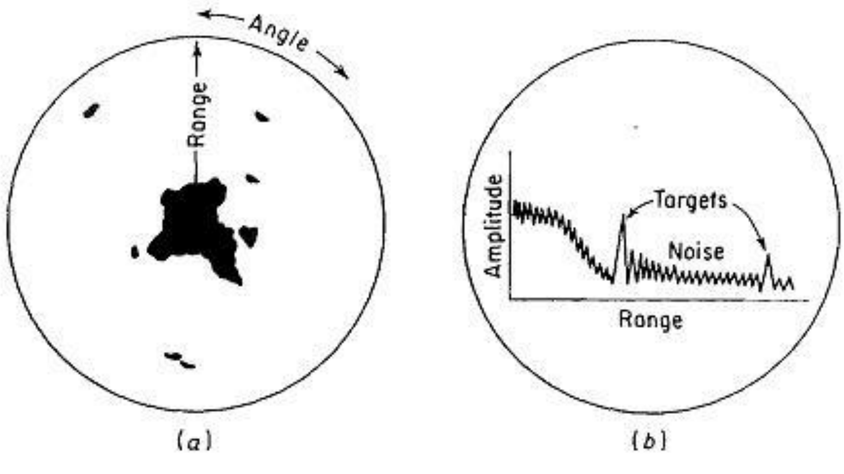


Fig.1.3 (a) PPI presentation displaying range vs. angle (intensity modulation); (h) A-scope presentation displaying amplitude vs. range (deflection modulation).

A B-scope display is similar to the PPI except that it utilizes rectangular, rather than polar, coordinates to display range vs. angle. Both the B-scope and the PPI, being intensity modulated, have limited dynamic range. Another form of display is the A-scope, shown in figure 1.3, which plots target amplitude (y axis) vs. range (x axis), for some fixed direction. This is a deflection-modulated display. It is more suited for tracking-radar application than for surveillance radar.

Radar Frequencies.

Conventional radars generally have been operated at frequencies extending from about 220 MHz to 35 GHz, a spread of more than seven octaves. These are not necessarily the limits, since radars can be, and have been, operated at frequencies outside either end of this range. Sky wave HF over-the-horizon (OTH) radar might be at frequencies as low as 4 or 5 MHz, and ground wave HF radars as low as 2MHz. At the other end of the spectrum, millimeter radars have operated at 94 GHz. Laser radars operate at even higher frequencies.The place of radar frequencies in the electromagnetic spectrum is shown in firure1.4. Some of the nomenclature employed to designate the various frequency regions is also shown.

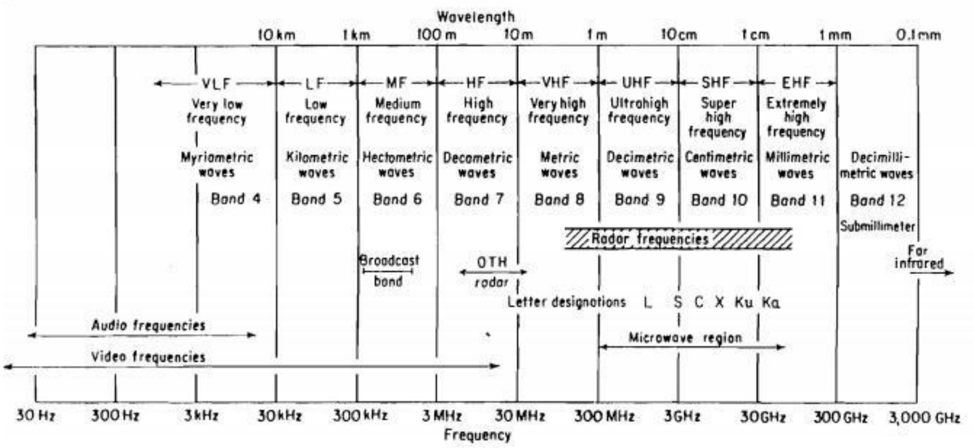


Fig.1.4 Radar frequencies and the electromagnetic spectrum

Early in the development of radar, a letter code such as S, X, L etc., was employed to designate radar frequency bands. Although its original purpose was to guard military secrecy, the designations were maintained, probably out of habit as well as the need for some convenient short nomenclature. This usage has continued and is now an accepted practice of radar engineers.

Table 1.1 lists the radar-frequency letter-band nomenclature adopted by the IEEE. These are related to the specific bands assigned by the International Telecommunications Union for radar. For example, although the nominal frequency range for L band is 1000 to 2000 MHz, an L-band radar is thought of as being confined within the region from 1215 to 1400 MHz since that is the extent of the assigned band. Letter-band nomenclature is not a substitute for the actual numerical frequency limits of radars. The specific numerical frequency limits should be used whenever appropriate, but the letter designations of Table 1.1 may be used whenever a short notation is desired substitute for the actual numerical frequency limits of radars. The specific numerical frequency limits should be used whenever appropriate, but the letter designations of Table 4.1 may be used whenever a short notation is desired.

Table 1.1 Standard radar-frequency letter-band nomenclature

Band designation	Nominal frequency range	Specific radiolocation (radar) bands based on ITU assignments for region 2
HF	3-30 MHz	
VHF	30-300 MHz	138-144 MHz
		216-225
UHF	300-1000 MHz	420-450 MHz
		890-942
L	1000-2000 MHz	1215-1400 MHz
S	2000-4000 MHz	2300-2500 MHz
		2700-3700
C	4000-8000 MHz	5250-5925 MHz
X	8000-12,000 MHz	8500-10,680 MHz
K_u	12.0-18 GHz	13.4-14.0 GHz
		15.7-17.7
K	18-27 GHz	24.05-24.25 GHz
K_a	27-40 GHz	33.4-36.0 GHz
mm	40-300 GHz	

Applications of Radar:

Radar has been employed on the ground, in the air, on the sea, and in space. Ground-based radar has been applied chiefly to the detection, location, and tracking of aircraft or space targets. Shipboard radar is used as a navigation aid and safety device to locate buoys, shore lines, and other ships as well as for observing aircraft. Airborne radar may be used to detect other aircraft, ships, or land vehicles, or it may be used for mapping of land, storm avoidance, terrain avoidance, and navigation. In space, radar has assisted in the guidance of spacecraft and for the remote sensing of the land and sea.

The major user of radar, and contributor of the cost of almost all of its development, has been the military: although there have been increasingly important civil applications, chiefly for marine and air navigation. The major areas of radar application, in no particular order of importance, are described below.

Air Traffic Control (ATC): Radars are employed throughout the world for the purpose of safely controlling air traffic en route and in the vicinity of airports. Aircraft and ground vehicular traffic at large airports are monitored by means of high-resolution radar. Radar has been used with GCA (ground-control approach) systems to guide aircraft to a safe landing in bad weather. In addition, the microwave landing system and the widely used ATC radar-beacon system are based in large part on radar technology.

Aircraft Navigation: The weather-avoidance radar used on aircraft to outline regions of precipitation to the pilot is a classical form of radar. Radar is also used for terrain avoidance and terrain following. Although they may not always be thought of as radars, the radio altimeter (either FM/CW or pulse) and the Doppler navigator are also radars. Sometimes ground-mapping radars of moderately high resolution are used for aircraft navigation purposes.

Ship Safety: Radar is used for enhancing the safety of ship travel by warning of potential collision with other ships, and for detecting navigation buoys, especially in poor visibility. In terms of numbers, this is one of the larger applications of radar, but in terms of physical size and cost it is one of the smallest. It has also proven to be one of the most reliable radar systems. Automatic detection and tracking equipments (also called plot extractors) are commercially available for use

with such radars for the purpose of collision avoidance. Shore-based radar of moderately high resolution is also used for the surveillance of harbors as an aid to navigation.

Space: Space vehicles have used radar for rendezvous and docking, and for landing on the moon. Some of the largest ground-based radars are for the detection and tracking of satellites. Satellite-borne radars have also been used for remote sensing as mentioned below.

Remote Sensing: All radars are remote sensors; however, as this term is used it implies the sensing of geophysical objects, or the "environment." For some time, radar has been used as a remote sensor of the weather. It was also used in the past to probe the moon and the planets (radar astronomy). The ionospheric sounder, an important adjunct for HF (short wave) communications, is a radar. Remote sensing with radar is also concerned with Earth resources, which includes the measurement and mapping of sea conditions, water resources, ice cover, agriculture, forestry conditions, geological formations, and environmental pollution. The platforms for such radars include satellites as well as aircraft.

Law Enforcement: In addition to the wide use of radar to measure the speed of automobile traffic by highway police, radar has also been employed as a means for the detection of intruders.

Military: Many of the civilian applications of radar are also employed by the military. The traditional role of radar for military application has been for surveillance, navigation, and for the control and guidance of weapons. It represents, by far, the largest use of radar.

Prediction of range performance

The simple form of the radar equation expressed the maximum radar range R_{max}, in terms of radar and target parameters:

$$R_{max} = \left[\frac{P_t G A_e \sigma}{(4\pi)^2 S_{min}}\right]^{1/4}$$

where P_t = transmitted power, watts
G = antenna gain
A_e = antenna effective aperture, m^2
σ = radar cross section, m^2
S_{min} = minimum detectable signal, watts

All the parameters are to some extent under the control of the radar designer, except for the target cross section σ. The radar equation states that if long ranges are desired, the transmitted power must be large, the radiated energy must be concentrated into a narrow beam (high transmitting antenna gain), the received echo energy must be collected with a large antenna aperture (also synonymous with high gain), and the receiver must be sensitive to weak signals.

In practice, however, the simple radar equation does not predict the range performance of actual radar equipments to a satisfactory degree of accuracy. The predicted values of radar range are usually optimistic. In some cases the actual range might be only half that predicted. Part of this discrepancy is due to the failure of Eq. above to explicitly include the various losses that can occur throughout the system or the loss in performance usually experienced when electronic equipment is operated in the field rather than under laboratory-type conditions. Another important factor that must be considered in the radar equation is the statistical or unpredictable nature of several of the parameters. The minimum detectable signal S_{min} and the target cross section σ are both statistical in nature and must be expressed in statistical terms.

Other statistical factors which do not appear explicitly in equation but which have an effect on the radar performance are the meteorological conditions along the propagation path and the performance of the radar operator, if one is employed. The statistical nature of these several parameters does not allow the maximum radar range to be described by a single number. Its specification must include a statement of the probability that the radar will detect a certain type of target at a particular range.

Minimum detectable signal

The ability of a radar receiver to detect a weak echo signal is limited by the noise energy that occupies the same portion of the frequency spectrum as does the signal energy. The weakest signal the receiver can detect is called the minimum detectable signal. The specification of the minimum detectable signal is sometimes difficult because of its statistical nature and because the criterion for deciding whether a target is present or not may not be too well defined.

Detection is based on establishing a threshold level at the output of the receiver. If the receiver output exceeds the threshold, a signal is assumed to be present. This is called threshold detection. Consider the output of a typical radar receiver as a function of time (figure 1.5). This might represent one sweep of the video output displayed on an A-scope. The envelope has a fluctuating appearance caused by the random nature of noise. If a large signal is present such as at A in figure 2.1, it is greater than the surrounding noise peaks and can be recognized on the basis of its amplitude.

Thus, if the threshold level were set sufficiently high, the envelope would not generally exceed the threshold if noise alone were present, but would exceed it if a strong signal were present. If the signal were small, however, it would be more difficult to recognize its presence. The threshold level must be low if weak signals are to be detected, but it cannot be so low that noise peaks cross the threshold and give a false indication of the presence of targets.

The voltage envelope of figure 2.1 is assumed to be from a matched-filter receiver. A matched filter is one designed to maximize the output peak signal to average noise (power) ratio. It has a frequency-response function which is proportional to the complex conjugate of the signal spectrum.The ideal matched-filter receiver cannot always be exactly realized in practice, but it is possible to approach it with practical receiver circuits. A matched filter for a radar transmitting a rectangular-shaped pulse is usually characterized by a bandwidth B approximately the reciprocal of the pulse width τ, or $B\tau \approx 1$.

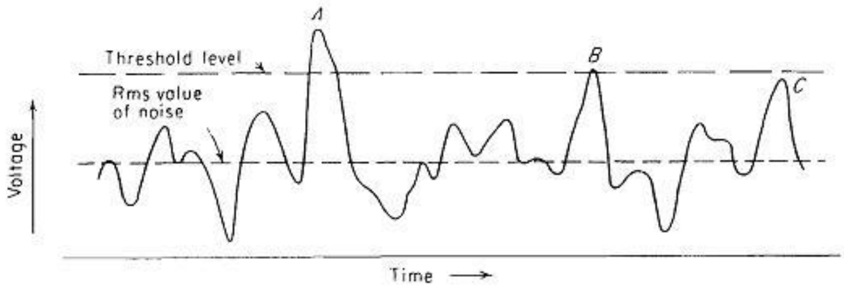


Fig.1.5 Typical envelope of tile radar receiver output as a function of time A, and B and C represent signal plus noise. A and B would be valid detections, but C is a missed detection.

The output of a matched-filter receiver is the cross correlation between the received waveform and a replica of the transmitted waveform. Hence it does not preserve the shape of the input waveform. (There is no reason to wish to preserve the shape of the received waveform so long as the output signal-to-noise ratio is maximized.)

Receiver Noise:

Since noise is the chief factor limiting receiver sensitivity, it is necessary to obtain some means of describing it quantitatively. Noise is unwanted electromagnetic energy which interferes with the ability of the receiver to detect the wanted signal. It may originate within the receiver itself, or it may enter via the receiving antenna along with the desired signal. If the radar were to operate in a perfectly noise-free environment so that no external sources of noise accompanied the desired signal, and if the receiver itself were so perfect that it did not generate any excess noise, there would still exist an unavoidable component of noise generated by the thermal motion of the conduction electrons in the ohmic portions of the receiver input stages. This is called thermal noise, or Johnson noise, and is directly proportional to the temperature of the ohmic portions of the circuit and the receiver bandwidth. The available thermal-noise power generated by a receiver of bandwidth B_n, (in hertz) at a temperature T (degrees Kelvin) is equal to kTB_n where k = Boltzmann's constant = 1.38×10^{-23} J/deg. If the temperature T is taken to be 290 K, which corresponds approximately to room temperature (62˚F), the factor kT is 4×10^{-21} W/Hz of bandwidth. If the receiver circuitry were at some other temperature, the thermal-noise power would be correspondingly different

A receiver with a reactance input such as a parametric amplifier need not have any significant ohmic loss. The limitation in this case is the thermal noise seen by the antenna and the ohmic losses in the transmission line. For radar receivers of the superheterodyne type, the receiver bandwidth is approximately that of the intermediate-frequency stages. It should be cautioned that the bandwidth B, of Eq. is not the 3-dB, or half-power, bandwidth commonly employed by electronic engineers. It is an integrated bandwidth and is given by

$$B_n = \frac{\int_{-\infty}^{\infty} |H(f)|^2 \, df}{|H(f_0)|^2}$$

where H(f) = frequency-response characteristic of IF amplifier (filter) and f_o = frequency of maximum response (usually occurs at midband). When H (f) is normalized to unity at midband (maximum-response frequency), H (f_o) = 1.

The bandwidth B_n is called the noise bandwidth and is the bandwidth of an equivalent rectangular filter whose noise-power output is the same as the filter with characteristic H (f). The 3-dB bandwidth is defined as the separation in hertz between the points on the frequency-response characteristic where the response is reduced to 0.707 (3 dB) from its maximum value. The 3-dB bandwidth is widely used, since it is easy to measure. The measurement of noise bandwidth however, involves a complete knowledge of the response characteristic H (f). The frequency-response characteristics of many practical radar receivers are such that the 3-dB and the noise bandwidths do not differ appreciably. Therefore the 3-dB bandwidth may be used in many cases as an approximation to the noise bandwidth.

The noise power in practical receivers is often greater than can be accounted for by thermal noise alone. The additional noise components are due to mechanisms other than the thermal agitation of the conduction electrons. The exact origin of the extra noise components is not important except to know that it exists. No matter whether the noise is generated by a thermal mechanism or by some other mechanism, the total noise at the output of the receiver may be considered to be equal to the thermal-noise power obtained from an" ideal”

$$F_n = \frac{N_o}{kT_0 B_n G_a} = \frac{\text{noise out of practical receiver}}{\text{noise out of ideal receiver at std temp } T_0}$$

where N_o = noise output from receiver, and G_a = available gain. The standard temperature T is taken to be 290 K, according to the Institute of Electrical and Electronics Engineers definition. The noise N_o is measured over the linear portion of the receiver input-output characteristic, usually at the output of the IF amplifier before the nonlinear second detector. The receiver bandwidth B_n is that of the IF amplifier in most receivers. The available gain G_a is the ratio of the signal out S_o to the signal in S_i, and kT_oB_n is the input noise N_i in an ideal receiver. Equation above may be rewritten as

$$F_n = \frac{S_i/N_i}{S_o/N_o}$$

The noise figure may be interpreted, therefore, as a measure of the degradation of signal-to-noise-ratio as the signal passes through the receiver.

Rearranging Eq. the input signal may be expressed as

$$S_i = \frac{kT_0 B_n F_n S_o}{N_o}$$

If the minimum detectable signal S_{min}, is that value of S_i corresponding to the minimum ratio of output (IF) signal-to-noise ratio $(S_o/N_o)_{min}$ necessary for detection, then

$$S_{\min} = kT_0 B_n F_n \left(\frac{S_o}{N_o}\right)_{\min}$$

Substituting Eq. discussed above into Eq. earlier results in the following form of the radar equation:

$$R_{\max}^4 = \frac{P_t G A_e \sigma}{(4\pi)^2 k T_0 B_n F_n (S_o/N_o)_{\min}}$$

Consider an IF amplifier with bandwidth B_{IF} followed by a second detector and a video amplifier with bandwidth B_V, (figure 1.6). The second detector and video amplifier are assumed to form an envelope detector, that is, one which rejects the carrier frequency but passes the modulation envelope. To extract the modulation envelope, the video bandwidth must be wide enough to pass the low-frequency components generated by the second detector, but not so wide as to pass the high-frequency components at or near the intermediate frequency. The video bandwidth B_V, must be greater than $B_{IF}/2$ in order to pass all the video modulation. Most radar receivers used in conjunction with an operator viewing a CRT display meet this condition and may be considered envelope detectors. Either a square-law or a linear detector may be assumed since the effect on the detection probability by assuming one instead of the other is usually small.

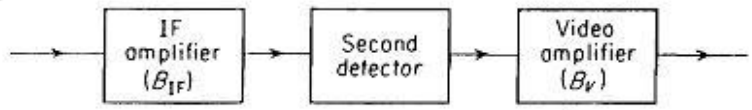


Fig.1.6 Envelope detector

The noise entering the IF filter (the terms filter and amplifier are used interchangeably) is assumed to be gaussian, with probability-density function given by

$$p(v) = \frac{1}{\sqrt{2\pi\psi_0}} \exp \frac{-v^2}{2\psi_0}$$

where p(v) dv is the probability of finding the noise voltage v between the values of v and v+dv, Ψ_0 is the variance of v is taken to be zero. If gaussian noise were passed through a narrowband IF filter-one whose bandwidth is small compared with the mid frequency-the probability density of the envelope of the noise voltage output is shown by Rice to be

$$p(R) = \frac{R}{\psi_0} \exp\left(-\frac{R^2}{2\psi_0}\right)$$

where R is the amplitude of the envelope of the filter output. Equation above is a form of the Rayleigh probability-density function. The probability that the envelope of the noise voltage will lie between the values of V_1 and V_2 is

$$\text{Probability } (V_1 < R < V_2) = \int_{V_1}^{V_2} \frac{R}{\psi_0} \exp\left(-\frac{R^2}{2\psi_0}\right) dR$$

The probability that the noise voltage envelope will exceed the voltage threshold V_T is

$$\text{Probability } (V_T < R < \infty) = \int_{V_T}^{\infty} \frac{R}{\psi_0} \exp\left(-\frac{R^2}{2\psi_0}\right) dR$$
$$= \exp\left(-\frac{V_T^2}{2\psi_0}\right) = P_{fa}$$

Whenever the voltage envelope exceeds the threshold, target detection is considered to have occurred, by definition. Since the probability of a false alarm is the probability that noise will cross the threshold, Eq. above gives the probability of a false alarm, denoted P_{fa}.

The average time interval between crossings of the threshold by noise alone is defined as the false-alarm time T_{fa},

$$T_{fa} = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} T_k$$

where T_k is the time between crossings of the threshold V_T by the noise envelope, when the slope of the crossing is positive. The false-alarm probability may also be defined as the ratio of the duration of time the envelope is actually above the threshold to the total time it could have been above the threshold, or

$$P_{fa} = \frac{\sum_{k=1}^{N} t_k}{\sum_{k=1}^{N} T_k} = \frac{\langle t_k \rangle_{av}}{\langle T_k \rangle_{av}} = \frac{1}{T_{fa} B}$$

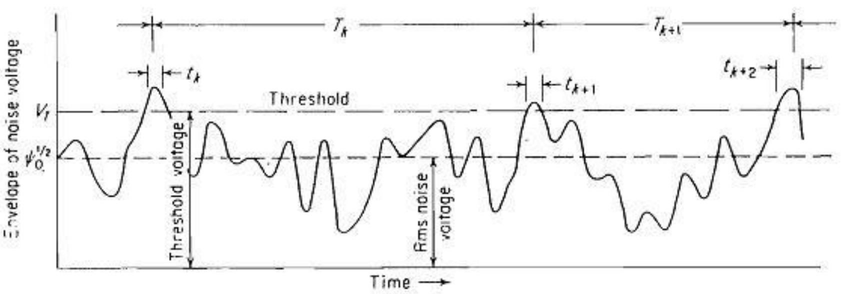


Fig.1.7 Envelope of receiver output illustrating false alarms due to noise.

where t_k and T_k are defined in firure 1.7. The average duration of a noise pulse is approximately the reciprocal of the bandwidth B, which in the case of the envelope detector is B_{IF}. Equating equations discussed above we get

$$T_{fa} = \frac{1}{B_{IF}} \exp \frac{V_T^2}{2\psi_0}$$

A plot of Eq. above is shown in firure 2.4, with $V^2_T / 2\,\Psi_0$ as the abscissa. If, for example, the bandwidth of the IF amplifier were 1 MHz and the average false-alarm time that could be tolerated were 15 min, the probability of a false alarm is 1.11×10^{-9} . From Eq. above the threshold voltage necessary to achieve this false-alarm time is 6.45 times the rms value of the noise voltage.

The false-alarm probabilities of practical radars are quite small. The reason for this is that the false-alarm probability is the probability that a noise pulse will cross the threshold during an interval of time approximately equal to the reciprocal of the bandwidth. For a 1-MHz bandwidth, there are of the order of 10^6 noise pulses per second. Hence the false-alarm probability of any one pulse must be small ($< 10^{-6}$) if false-alarm times greater than 1 s are to be obtained.

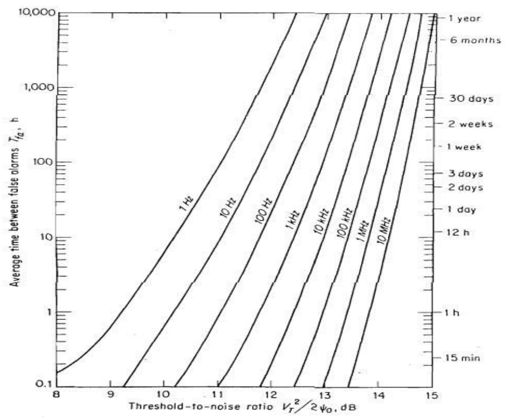


Fig.1.8 Average time between false alarms as a function of the threshold level V_T and the receiver bandwidth B; Ψ_o is the mean square noise voltage.

Consider sine-wave signals of amplitude A to be present along with noise at the input to the IF filter. The frequency of the signal is the same as the IF midband frequency f_{IF}. The output or the envelope detector has a probability-density function given by

$$p_s(R) = \frac{R}{\psi_0} \exp\left(-\frac{R^2 + A^2}{2\psi_0}\right) I_0\left(\frac{RA}{\psi_0}\right)$$

where I_0 (Z) is the modified Bessel function of zero order and argument Z. For Z large, an asymptotic expansion for I_0 (Z) is

$$I_0(Z) \approx \frac{e^Z}{\sqrt{2\pi Z}}\left(1 + \frac{1}{8Z} + \cdots\right)$$

When the signal is absent, A = 0, the probability-density function for noise alone. Equation above is sometimes called the Rice probability-density function. The probability that the signal will be detected (which is the probability of detection) is the same as the probability that the envelope R will exceed the predetermined threshold V_T. The probability of detection P_d, is therefore

$$P_d = \int_{V_T}^{\infty} p_s(R)\, dR = \int_{V_T}^{\infty} \frac{R}{\psi_0} \exp\left(-\frac{R^2 + A^2}{2\psi_0}\right) I_0\left(\frac{RA}{\psi_0}\right) dR$$

This cannot be evaluated by simple means, and numerical techniques or a series approximation must be used. A series approximation valid when RA/ Ψ_0 >> 1, A >> I R - AI, and terms in A^{-3} and beyond can be neglected is

$$P_d = \frac{1}{2}\left(1 - \operatorname{erf}\frac{V_T - A}{\sqrt{2\psi_0}}\right) + \frac{\exp[-(V_T - A)^2/2\psi_0]}{2\sqrt{2\pi}(A/\sqrt{\psi_0})} \times \left[1 - \frac{V_T - A}{4A} + \frac{1 + (V_T - A)^2/\psi_0}{8A^2/\psi_0} - \cdots\right]$$

where the error function is defined as

$$\operatorname{erf} Z = \frac{2}{\sqrt{\pi}}\int_0^Z e^{-u^2}\, du$$

Equation discussed above may be converted to power by replacing the signal to -rms-noise-voltage ratio with the following:

$$\frac{A}{\psi_0^{1/2}} = \frac{\text{signal amplitude}}{\text{rms noise voltage}} = \frac{\sqrt{2}(\text{rms signal voltage})}{\text{rms noise voltage}} = \left(2\,\frac{\text{signal power}}{\text{noise power}}\right)^{1/2} = \left(\frac{2S}{N}\right)^{1/2}$$

Integration of radar pulses

Many pulses are usually returned from any particular target on each radar scan and can be used to improve detection. The number of pulses n_B returned from a point target as the radar antenna scans through its beamwidth is

$$n_B = \theta_B f_p \quad \theta_B f_p$$

where θ_B = antenna beamwidth, deg

f_p = pulse repetition frequency, Hz

$\dot{\theta}_s$ = antenna scanning rate, deg/s

ω_m = antenna scan rate, rpm

Typical parameters for a ground-based search radar might be pulse repetition frequency 300 Hz, 1.5˚ beam width, and antenna scan rate 5 rpm (30˚/s). These parameters result in 15 hits from a point target on each scan. The process of summing all the radar echo pulses for the purpose of improving detection is called integration. Many techniques might be employed for accomplishing integration. All practical integration techniques employ some sort of storage device. Perhaps the most common radar integration method is the cathode-ray-tube display combined with the integrating properties of the eye and brain of the radar operator.

Integration may be accomplished in the radar receiver either before the second detector (in the IF) or after the second detector (in the video). A definite distinction must be made between these two cases. Integration before the detector is called predetection, or coherent, integration, while integration after the detector is called postdetection, or noncoherent, integration. Predetection integration requires that the phase of the echo signal be preserved if full benefit is to be obtained from the summing process. On the other hand, phase information is destroyed by the second detector; hence postdetection integration is not concerned with preserving RF phase. For this convenience, postdetection integration is not as efficient as predetection integration.

If n pulses, all of the same signal-to-noise ratio, were integrated by an ideal predetection integrator, the resultant, or integrated, signal-to-noise (power) ratio would be exactly n times that of a single pulse. If the same n pulses were integrated by an ideal postdetection device, the resultant signal-to-noise ratio would be less than n times that of a single pulse. This loss inintegration efficiency is caused by the nonlinear action of the second detector, which converts some of the signal energy to noise energy in the rectification process. The comparison of predetection and postdetection integration may be briefly summarized by stating that although postdetection integration is not as efficient as predetection integration, it is easier to implement in most applications. Post detection integration is therefore preferred, even though the integrated signal-to-noise ratio may not be as great. An alert, trained operator viewing a properly designed cathode-ray tube display is a close approximation to the theoretical postdetection integrator.

The efficiency of postdetection integration relative to ideal predetection integration has been computed by Marcum when all pulses are of equal amplitude. The integration efficiency may be defined as follows:

where n = number of pulses integrated

$(S/N)_1$ = value of signal-to-noise ratio of a single pulse required to produce given probability of detection (for $n = 1$)

$(S/N)_n$ = value of signal-to-noise ratio per pulse required to produce same probability of detection when n pulses are integrated

$$E_i(n) = \frac{(S/N)_1}{n(S/N)_n}$$

Radar Cross Section of Targets:

The radar cross section of a target is the (fictional) area intercepting that amount of power which when scattered equally in all directions, produces an echo at the radar equal to that from the target; or in other terms,

$$\sigma = \frac{\text{power reflected toward source/unit solid angle}}{\text{incident power density}/4\pi} = \lim_{R \to \infty} 4\pi R^2 \left| \frac{E_r}{E_i} \right|^2$$

where R = distance between radar and target

E_r = reflected field strength at radar

E_i = strength of incident field at target

This equation is equivalent to the radar range equation. For most common types of radar targets such as aircraft, ships, and terrain, the radar cross section does not necessarily bear a simple relationship to the physical area, except that the larger the target size, the larger the cross section is likely to be Scattering and diffraction are variations of the same physical process. When an object scatters an electromagnetic wave, the scattered field is defined as the difference between the total field in the presence of the object and the field that would exist if the object were absent (but with the sources unchanged). On the other hand, the diffracted field is the total field in the presence of the object. With radar backscatter, the two fields are the same, and one may talk about scattering and diffraction interchangeably.

In theory, the scattered field, and hence the radar cross section, can be determined by solving Maxwell's equations with the proper boundary conditions applied. Unfortunately, the determination of the radar cross section with Maxwell's equations can be accomplished only for the most simple of shapes, and solutions valid over a large range of frequencies are not easy to obtain. The radar cross section of a simple sphere is shown in figure 2.5 as a function of its circumference measured in wavelengths ($2\pi a/\lambda$, where a is the radius of the sphere and λ is the wavelength). The region where the size of the sphere is small compared with the wavelength ($2\pi a/\lambda \ll 1$) is called the Rayleigh region, after Lord Rayleigh who, in the early 1870's first studied scattering by small particles. Lord Rayleigh was interested in the scattering of light by microscopic particles, rather than in radar. His work preceded the original electromagnetic echo experiments of Hertz by about fifteen years.The Rayleigh scattering region is of interest to the radar engineer because the cross sections of raindrops and other meteorological particles fall within this region at the usual radar frequencies. Since the cross section of objects within the Rayleigh region varies as λ^{-4}, rain and clouds are essentially invisible to radars which operate at relatively long wavelengths (low frequencies).

The usual radar targets are much larger than raindrops or cloud particles, and lowering the radar frequency to the point where rain or cloud echoes are negligibly small will not seriously reduce the cross section of the larger desired targets. On the other hand, if it were desired to actually observe, rather than eliminate, raindrop echoes, as in a meteorological or weather-observing radar, the higher radar frequencies would be preferred.

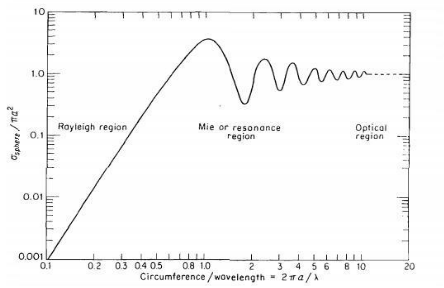


Fig.1.8 Radar cross section of the sphere. a = radius; λ = wavelength

At the other extreme from the Rayleigh region is the optical region, where the dimensions of the sphere are large compared with the wavelength ($2\pi a/\lambda >> 1$). For large $2\pi a/\lambda$, the radar cross section approaches the optical cross section a^2. In between the optical and the Rayleigh region is the Mie, or resonance, region. The cross section is oscillatory with frequency within this region. The maximum value is 5.6 dB greater than the optical value, while the value of the first null is 5.5 dB below the optical value. (The theoretical values of the maxima and minima may vary according to the method of calculation employed.) The behavior of the radar cross sections of other simple reflecting objects as a function of frequency is similar to that of the sphere.

Transmitter Power:

The power P_t in the radar equation is called by the radar engineer the peak power. The peak pulse power as used in the radar equation is not the instantaneous peak power of a sine wave. It is defined as the power averaged over that carrier-frequency cycle which occurs at the maximum of the pulse of power. (Peak power is usually equal to one-half the maximum instantaneous power.) The average radar power P_{av}, is also of interest in radar and is defined as the average transmitter power over the pulse-repetition period. If the transmitted waveform is a train of rectangular pulses of width τ and pulse-repetition period $T_p = 1/f_p$, the average power is related to the peak power by

$$P_{av} = \frac{P_t \tau}{T_p} = P_t \tau f_p$$

The ratio P_{av}/P_t, τ/T_p, or τf_p is called the duty cycle of the radar. A pulse radar for detection of aircraft might have typically a duty cycle of 0.001, while a CW radar which transmits continuously has a duty cycle of unity. Writing the radar equation in terms of the average power rather than the peak power, we get

$$R_{max}^4 = \frac{P_{av} G A_e \sigma n E_i(n)}{(4\pi)^2 k T_0 F_n (B_n \tau)(S/N)_1 f_p}$$

The bandwidth and the pulse width are grouped together since the product of the two is usually of the order of unity in most pulse-radar applications. If the transmitted waveform is not a rectangular pulse, it is sometimes more convenient to express the radar equation in terms of the energy $E_\tau = P_{av}/f_p$ contained in the transmitted waveform:

$$R_{max}^4 = \frac{E_\tau G A_e \sigma n E_i(n)}{(4\pi)^2 k T_0 F_n (B_n \tau)(S/N)_1}$$

In this form, the range does not depend explicitly on either the wavelength or the pulse repetition frequency. The important parameters affecting range are the total transmitted energy nE_τ, the transmitting gain G, the effective receiving aperture A_e, and the receiver noise figure F_n.

Pulse repetition frequency and range ambiguities

The pulse repetition frequency (prf) is determined primarily by the maximum range at which targets are expected. If the prf is made too high the likelihood of obtaining target echoes from the wrong pulse transmission is increased. Echo signals received after an interval exceeding the pulse-repetition period are called multiple-time-around echoes. They can result in erroneous or confusing range measurements. Consider the three targets labeled A. B, and C in figure. 2.6(a). Target A is located within the maximum unambiguous range R_{unamb} of the radar, target B is at a distance greater than R_{unamb} but less than $2R_{unamb}$ while target C is greater than 2Runarnh but less than $3R_{unamb}$. The appearance of the three targets on an A-scope is sketched in figure 2.6(b). The multiple-time-around echoes on the A-scope cannot be distinguished from proper target echoes actually within the maximum unambiguous range. Only the range measured for target A is correct; those for Band C are not.

One method of distinguishing multiple-time-around echoes from unambiguous echoes is to operate with a varying pulse repetition frequency. The echo signal from an unambiguous range target will appear at the same place on the A-scope on each sweep no matter whether the prf is modulated or not. However, echoes from multiple-time-around targets will be spread over a finite range as shown in firure 1.9 (c). The prf may be changed continuously within prescribed limits or it may be changed discretely among several predetermined values. The number of separate pulse repetition frequencies will depend upon the degree of the multiple-time targets. Second-time targets need only two separate repetition frequencies in order to be resolved.

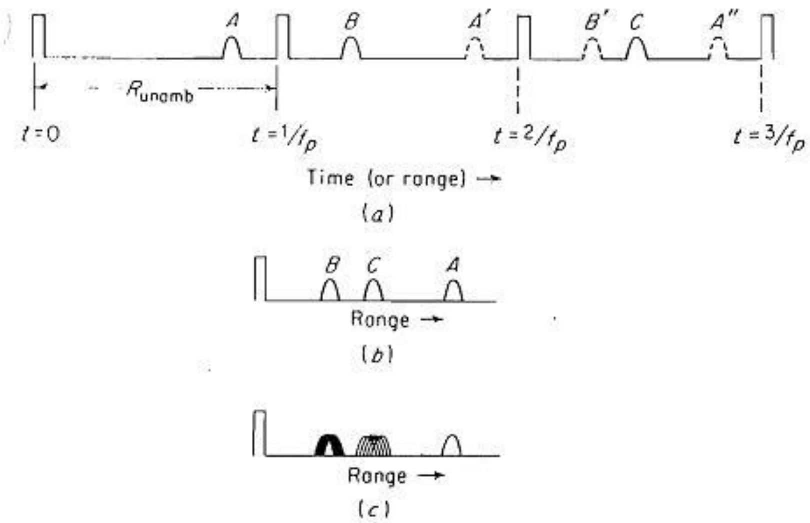


Fig.1.9 Multiple-time-around echoes that give rise to ambiguities in range (a) Three targets A, B and C, where A is within R_{unamb}, and B and C are multiple-time-around targets;(b) the appearance of the three targets on the A-scope; (c) appearance of the three targets on the A-scope with a changing prf.

Instead of modulating the prf, other schemes that might he employed to" mark" successive pulses so as to identify multiple-time-around echoes include changing the pulse: amplitude, pulse width, frequency, phase, or polarization of transmission from pulse to pulse. Generally, such schemes are not so successful in practice as one would like. One of the fundamental limitations is the fold over of nearby targets; that is, nearby strong ground targets (clutter) can be quite large and can mask weak multiple-time-around targets appearing at the same place on the display. Also, more time is required to process the data when resolving ambiguities. Ambiguities may theoretically be resolved by observing the variation of the echo signal with time (range). This is not always a practical technique; however, since the echo-signal amplitude can fluctuate strongly for reasons other than a change in range.Instead, the range ambiguities in a multiple prf radar can be conveniently decoded and the true range found by the use of the Chinese remainder theorem or other computational algorithms.

System Losses:

One of the important factors omitted from the simple radar equation was the losses that occur throughout the radar system. The losses reduce the signal-to-noise ratio at the receiver output. They may be of two kinds, depending upon whether or not they can be predicted with any degree of precision beforehand.

The antenna beam-shape loss, collapsing loss, and losses in the microwave plumbing are examples of losses which can be calculated if the system configuration is known. These losses are very real and cannot be ignored in any serious prediction of radar performance. Losses not readily subject to calculation and which are less predictable include those due to field degradation and to operator fatigue or lack of operator motivation. Estimates of the latter type of loss must be made on the basis of prior experience and experimental observations. They may be subject to considerable variation and uncertainty. Although the loss associated with any one factor may be small, there are many possible loss mechanisms in a complete radar system, and their sum total can be significant.

Plumbing loss: There is always some finite loss experienced in the transmission lines which connect the output of the transmitter to the antenna. The losses in decibels per 100 ft for radar transmission lines are shown in figure 1.10. At the lower radar frequencies the transmission line introduces little loss, unless its length is exceptionally long. At the higher radar frequencies, attenuation may not always be small and may have to be taken into account. In addition to the losses in the transmission line itself, an additional loss can occur at each connection or bend in the line and at the antenna rotary joint if used.

Connector losses are usually small, but if the connection is poorly made, it can contribute significant attenuation. Since the same transmission line is generally used for both receiving and transmission, the loss to be inserted in the radar equation is twice the one-way loss.

The signal suffers attenuation as it passes through the duplexer. Generally, the greater the isolation required from the duplexer on transmission, the larger will be the insertion loss. By insertion loss is meant the loss introduced when the component, in this case the duplexer, is inserted into the transmission line. The precise value of the insertion loss depends to a large extent on the particular design. For a typical duplexer it might be of the order of 1 dB. A gas-tube duplexer also introduces loss when in the fired condition (arc loss); approximately 1 dB is typical.

In S-band (3000 MHz) radar, for example, the plumbing losses might be as follows:

100 ft of RG-113/U A1 waveguide transmission line (two-way)	1.0 dB
Loss due to poor connections (estimate)	0.5 dB
Rotary-joint loss	0.4 dB
Duplexer loss	1.5 dB
Total plumbing loss	3.4 dB

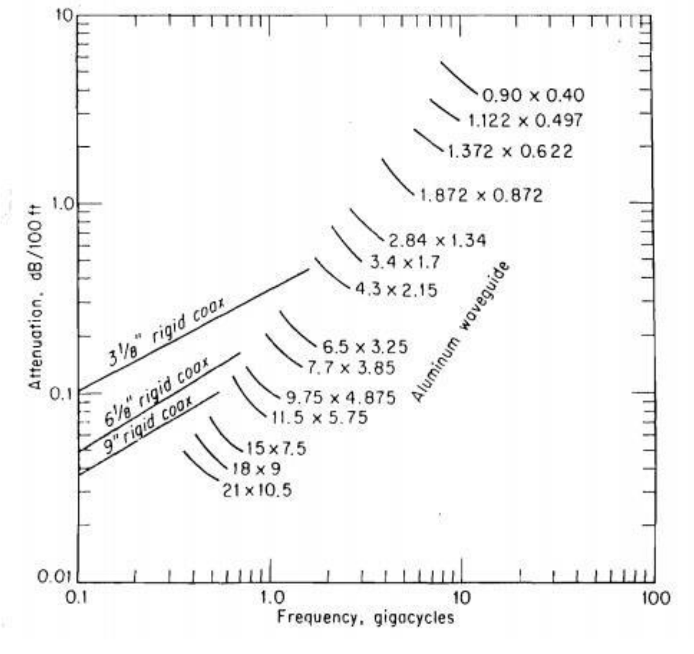


Fig.1.10 Theoretical (one-way) attenuation of RF transmission lines Waveguide sizes are inches and are the inside dimensions.

Beam-shape loss: The antenna gain that appears in the radar equation was assumed to be a constant equal to the maximum value. But in reality the train of pulses returned from a target with scanning radar is modulated in amplitude by the shape of the antenna beam. To properly take into account the pulse-train modulation caused by the beam shape, the computations of the probability of detection would have to be performed assuming a modulated train of pulses rather than constant-amplitude pulses. Some authors do indeed take account of the beam shape in this manner when computing the probability of detection. Therefore, when using published computations of probability of detection it should be noted whether the effect of the beam shape has been included. This approach is not used. Instead a beam-shape loss is added to the radar equation to account for the fact that the maximum gain is employed in the radar equation rather than a gain that changes pulse to pulse. This is a simpler, albeit less accurate, method. It is based on calculating the reduction in signal power and thus does not depend on the probability of detection. It applies for detection probabilities in the vicinity of 0.50, but it is used as an approximation with other values as a matter of convenience.

Let the one-way-power antenna pattern be approximated by the Gaussian expression exp $(-2.78\theta^2/\theta_B^2)$, where θ is the angle measured from the center of the beam and θ_B is the beamwidth measured between half-power points. If n_B is the number of pulses received within the half-power beamwidth θ_B, and n is the total number of pulses integrated, then the beam-shape loss (number greater than unity) relative to a radar that integrates all n pulses with an antenna gain corresponding to the maximum gain at the beam center is

$$\text{Beam-shape loss} = \frac{n}{1 + 2\sum_{k=1}^{(n-1)/2} \exp(-5.55k^2/n_B^2)}$$

Limiting loss: Limiting in the radar receiver can lower the probability of detection. Although a well-designed and engineered receiver will not limit the received signal under normal circumstances, intensity modulated CRT displays such as the PPI or the B-scope have limited dynamic range and may limit. Some receivers, however, might employ limiting for some special purpose, as for pulse compression processing for example. Limiting results in a loss of only a fraction of a decibel for a large number of pulses integrated provided the limiting ratio (ratio of video limit level to rms noise level) is as large as 2 or 3. Other analyses of band pass limiters show that for small signal-to-noise ratio, the reduction in the signal-to-noise ratio of a sine-wave imbedded in narrowband Gaussian noise is Π/4 (about 1 dB). However, by appropriately shaping the spectrum of the input noise, it has been suggested that the degradation can be made negligibly small.

Collapsing loss: If the radar were to integrate additional noise samples along with the wanted signal-to-noise pulses, the added noise results in degradation called the collapsing loss. It can occur in displays which collapse the range information, such as the C-scope which displays elevation vs. azimuth angle. The echo signal from a particular range interval must compete in a collapsed-range C-scope display, not only with the noise energy contained within that range interval but with the noise energy from all other range intervals at the same elevation and azimuth. In some 3D radars (range, azimuth, and elevation) that display the outputs at all elevations on a single PPI (range, azimuth) display, the collapsing of the 3D radar information onto a 2D display results in a loss. A collapsing loss can occur when the output of a high-resolution radar is displayed on a device whose resolution is coarser than that inherent in the radar. A collapsing loss also results if the outputs of two (or more) radar receivers are combined and only one contains signal while the other contains noise.

The mathematical derivation of the collapsing loss, assuming a square-law detector may be carried out as suggested by Marcum who has shown that the integration of m noise pulses, along with n signal-plus-noise pulses with signal-to-noise ratio per pulse $(S/N)_n$, is equivalent to the integration of m + n signal-to-noise pulses each with signal-to-noise ratio $n(S/N)_n/(m + n)$. The collapsing loss in this case is equal to the ratio of the integration loss L_i for m + n pulses to the integration loss for n pulses, or

$$L_i(m, n) = \frac{L_i(m + n)}{L_i(n)}$$

UNIT-II

UNIT-III

CW AND FREQUENCY MODULATED RADAR

Doppler Effect

A radar detects the presence of objects and locates their position in space by transmitting electromagnetic energy and observing the returned echo. A pulse radar transmits a relatively short burst of electromagnetic energy, after which the receiver is turned on to listen for the echo. The echo not only indicates that a target is present, but the time that elapses between the transmission of the pulse and the receipt of the echo is a measure of the distance to the target. Separation of the echo signal and the transmitted signal is made on the basis of differences in time.

The radar transmitter may be operated continuously rather than pulsed if the strong transmitted signal can be separated from the weak echo. The received-echo-signal power is considerably smaller than the transmitter power; it might be as little as 10^{-18} that of the transmitted power-sometimes even less. Separate antennas for transmission and reception help segregate the weak echo from the strong leakage signal, but the isolation is usually not sufficient. A feasible technique for separating the received signal from the transmitted signal when there is relative motion between radar and target is based on recognized the change in the echo-signal frequency caused by the Doppler effect.

It is well known in the fields of optics and acoustics that if either the source of oscillation or the observer of the oscillation is in motion, an apparent shift in frequency will result. This is the Doppler Effect and is the basis of CW radar. If R is the distance from the radar to target, tile total number of wavelengths λ contained in the two-way path between the radar and the target is 2R / λ. The distance R and the wavelength λ are assumed to be measured in the same units. Since one wavelength corresponds to an angular excursion of 2π radians, the total angular excursion φ made by the electromagnetic wave during its transit to and from the target is 4πR / λ radians. If the target is in motion, R and the phase φ are continually changing. A change in φ with respect to time is equal to a frequency. This is the Doppler angular frequency ω_d given by

$$\omega_d = 2\pi f_d = d\varphi / dt = (4\pi / \lambda)(dR/dt) = 4\pi v_r / \lambda$$

Where,

f_d = Doppler frequency shift and v_r = relative (or radial) velocity of target with to radar.

The Doppler frequency shift is

$$f_d = 2 * v_r / \lambda = 2 * v_r * f_o / c$$

where,

f_o = transmitted frequency

c = velocity of propagation = 3 x 10^8 m/s.

If f_d is in hertz v_r in knots, and λ in meters,

$$f_d = 1.03 * v_r / \lambda$$

A plot of this equation is shown in figure 3.1

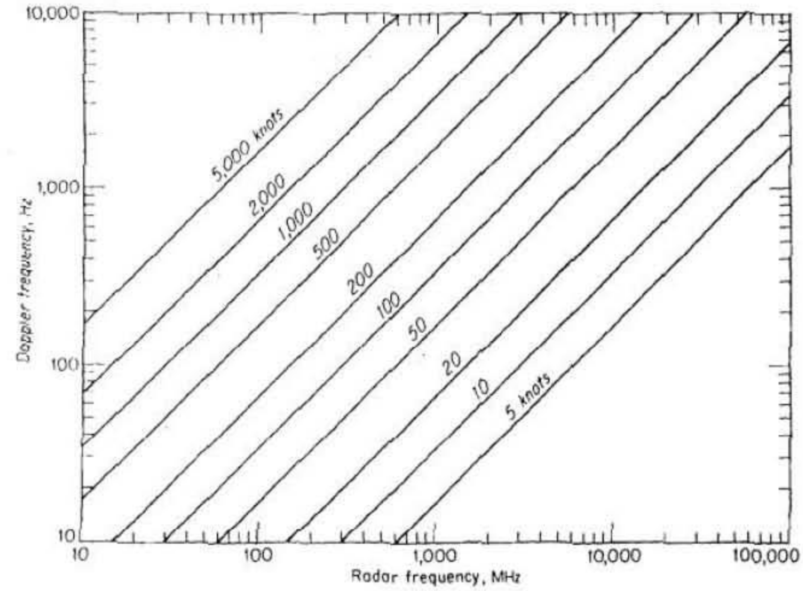


Fig.3.1 Doppler frequency as a function of radar frequency and target relative velocity

The relative velocity may be written $v_r = v \cos\theta$, where v is the target speed and θ is the angle made by the target trajectory and the line joining radar and target. When θ = 0. The Doppler frequency is maximum. The Doppler is zero when the trajectory is perpendicular to the radar line of sight (θ = 90°).

CW radar and its working.

Consider the simple CW radar as illustrated by the block diagram of figure3.2. The transmitter generates a continuous (unmodulated) oscillation of frequency f_0, which is radiated by the antenna. A portion of the radiated energy is intercepted by the target and is scattered, some of it in the direction of the radar, where it is collected by the receiving antenna. If the target is in motion with a velocity v_r, relative to the radar, the received signal will be shifted in frequency from the transmitted frequency f_0 by an amount ± fd as given $fd=2v_r/\lambda$

The plus sign associated with the Doppler frequency applies if the distance between target and radar is decreasing (closing target), that is, when the received signal frequency is greater than the transmitted signal frequency. The minus sign applies if the distance is increasing (receding target). The received echo signal at a frequency $f_0 \pm f_d$ enters the radar via the antenna and is heterodyned in the detector (mixer) with a portion of the transmitter signal f_0 to produce a doppler beat note of frequency f_d. The sign of f_d is lost in this process.

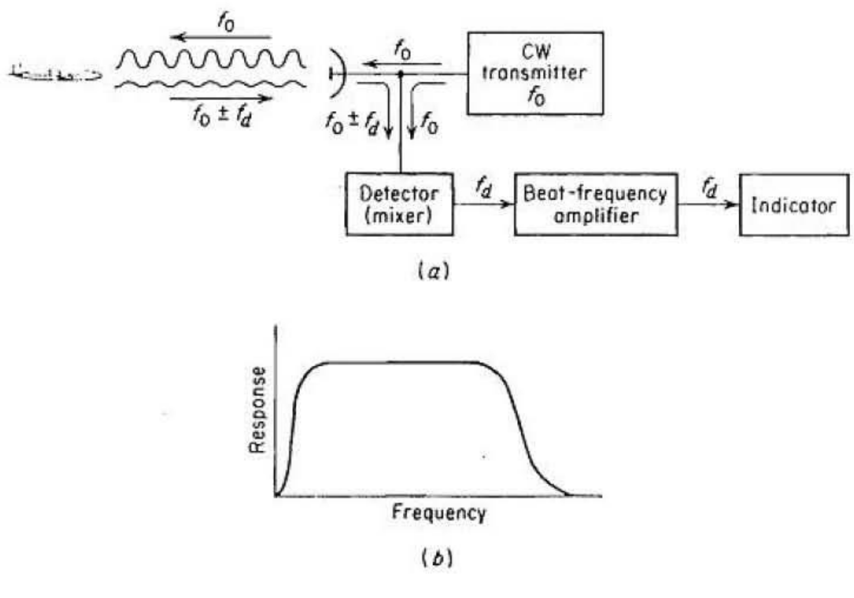


Fig.3.2 (a) Simple CW radar block diagram; (b) response characteristic of beat-frequency amplifier

The purpose of the doppler amplifier is to eliminate echoes from stationary targets and to amplify the doppler echo signal to a level where it can operate an indicating device. It might have a frequency-response characteristic similar to that of figure 3.2 (b). The low-frequency cutoff must be high enough to reject tile d-c component caused by stationary targets, but yet it might be low enough to pass the smallest Doppler frequency expected. Sometimes both conditions cannot be met simultaneously and a compromise is necessary. The upper cutoff frequency is selected to pass the highest Doppler frequency expected.

The indicator might be a pair of earphones or a frequency meter. If exact knowledge of the Doppler frequency is not necessary, earphones are especially attractive provided the Doppler frequencies lie within the audio-frequency response of the ar. Earphones are not only simple devices. But the ear acts as a selective bandpass filter with a passband of the order of 50Hz centered about the signal frequency. The narrow-bandpass characteristic of the ear results in an effective increase in the signal-to-noise ration of the echo signal. With subsonic aircraft targets and transmitter frequencies in the middle range of the microwave frequency region, the Doppler frequencies usually fall within the passband of the ear. If audio detection were desired for those combinations of target velocity and transmitter frequency which do not result in audible Doppler frequencies, the dopppler signal could be heterodyned to the audible range. The doppler frequency can also be detected and measured by conventional frequency meters, usually one that counts cycles.

CW Radar with nonzero intermediate-frequency receiver

The receiver of the simple CW radar is in some respects analogous to a superheterodyne receiver. Receivers of this type are called homodyne receivers, or superheterodyne receivers with zero IF. The function of the local oscillator is replaced by the leakage signal from the transmitter. Such a receiver is simpler than one with a more conventional intermediate frequency since no IF amplifier or local oscillator is required. However, the simpler receiver is not as sensitive because of increased noise at the lower intermediate frequencies caused by flicker effect. Flicker-effect noise occurs in semiconductor devices such as diode detectors and cathodes of vacuum tubes. The noise power produced by the flicker effect varies as $1/f^{\alpha}$, where α is approximately unity. This is in contrast to shot noise or thermal noise, which is independent of frequency. Thus, at the lower range of frequencies (audio or video region), where the doppler frequencies usually are found, the detector of the CW receiver can introduce a considerable amount of flicker noise, resulting in reduced receiver sensitivity.

For short-range, low-power, applications this decrease in sensitivity might be tolerated since it can be compensated by a modest increase in antenna aperture and/or additional transmitter power. But for 'maximum efficiency with CW radar, the reduction in sensitivity caused by the simple doppler receiver with zero IF, cannot be tolerated. The effects of flicker noise are overcome in the normal superheterodyne receiver by using an intermediate frequency high enough to render the flicker noise small compared with the normal receiver noise. This results from the inverse, frequency dependence of flicker noise.

Figure 3.3 shows a block diagram of the CW radar whose receiver operates with a nonzero IF. Separate antennas are shown for transmission and reception instead of the usual local oscillator found in the conventional susuperheterodyne receiver, the local oscillator (or reference signal) is derived in this receiver from a portion of the transmitted signal mixed with a locally generated signal of frequency equal to that of the receiver IF. Since the output of the mixer consists of two sidebands on either side of the carrier plus higher harmonics, a narrowband filter selects one of the sidebands as the reference signal. The improvement in receiver sensitivity with an intermediate - frequency superheterodyne might be as much as 30 dB over simple receiver.

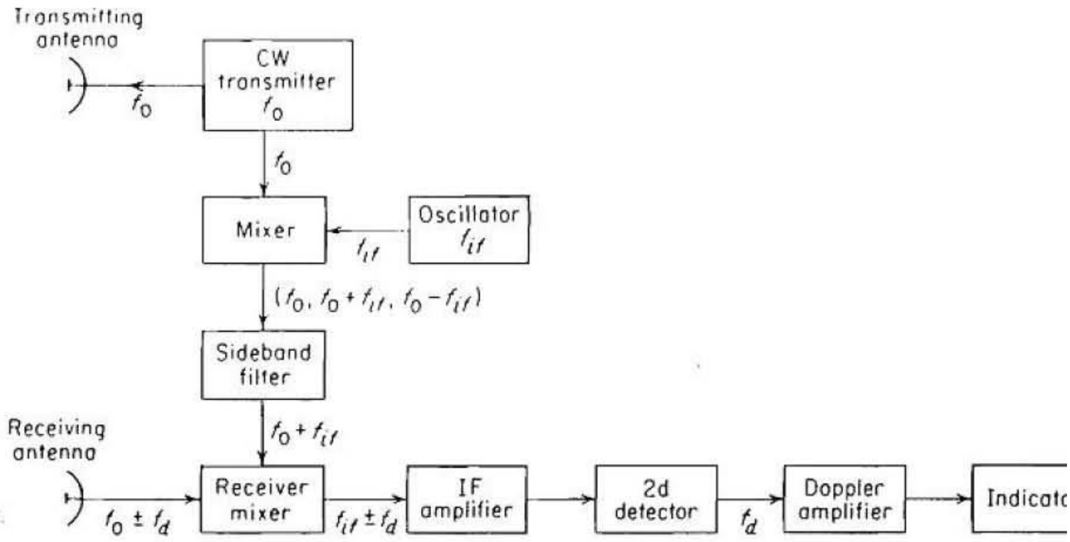


Fig. 3.3 Block diagram of CW doppler radar with nonzero IF receiver, sometimes called sideband superheterodyne.

The sign of the doppler frequency, and therefore the direction of target motion, may be found by splitting the received signal into two channels as shown in figure 3.4.

In channel A the signal is processed as in the simple CW radar. The received signal and a portion of the transmitter heterodyne in the detector (mixer) to yield a difference signal

$$E_A = K_2 \ E_0 \cos(\pm w_d t + \varphi)$$

E_A = amplitude of transmitter signal, K_2 = a constant determined from the radar equation

w_d = dopper angular frequency shift and φ = a constant phase shift, which depends upon range of initial detection

The other channel is similar, except for a 90° phase delay introduced in the reference signal. The output of the channel B mixer is

$$E_B = K_2 \; E_0 \cos(\pm w_d t + \varphi + \pi / 2)$$

If the target is approaching (positive Doppler), the outputs from the two channels are

$$E_A(+) = K_2 \; E_0 \cos(w_d t + \varphi)$$

$$E_B(+) = K_2 \; E_0 \cos(w_d t + \varphi + \pi / 2)$$

If the targets are receding (negative doppler), the outputs from the two channels are

$$E_A(-) = K_2 \; E_0 \cos(w_d t - \varphi)$$

$$E_B(-) = K_2 \; E_0 \cos(w_d t - \varphi - \pi / 2)$$

The sign of w_d and the direction of the target's motion may be determined according to whether the output of channel B leads or lags the output of channel A. One method of determining the relative phase relationship between the two channels is to apply the outputs to a synchronous two-phase motor. The direction of motor rotation is an indication of the direction of the target motion.Electronic methods may be used instead of a synchronous motor to sense the relative phase of the two channels.

One application of this technique has been described for a rate-of-climb meter for vertical take-off aircraft to determine the velocity of the aircraft with respect to the ground during take-off and landing. It has also been applied to the detection of moving targets in the presence of heavy foliage.

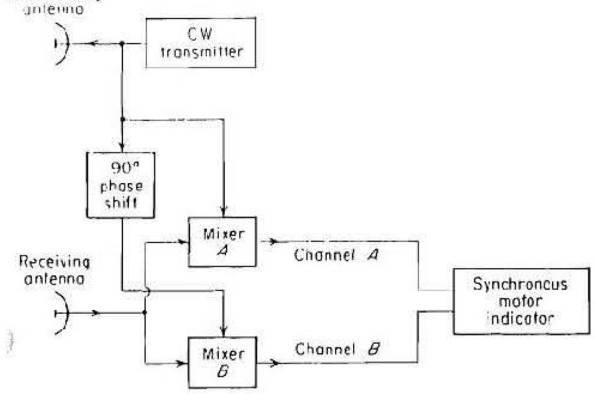


Fig. 3.4 Measurement of Doppler direction using synchronous, two-phase motor

Operation of CW tracking Illuminator

Some anti-air-warfare guided missile systems employ semi active homing guidance in which a receiver in the missile receives energy from the target, the energy having been transmitted from an illuminator external to the missile. The illuminator, for example, might be at the launch platform. CW illumination has been used in many successful systems. An example is the Hawk tracking illuminator. It is tracking radar as well as in illuminator since it must be able to follow the target as it travels through space.

The Doppler discrimination of a CW radar allows operation in the presence of clutter and has been well suited for low altitude missile defense systems. A block diagram of a CW tracking illuminator is shown in figure 3.5.

The operation in presence of clutter is possible, due to the Doppler discrimination of continuous wave radar. In this type of radar, the receiver in the missile receives the energy from the target and this energy has been transmitted to the missile by an illuminator. The illuminator may be at the launch platform. This CW illuminator has been used in many successful systems.

The wide-band Doppler amplifier is a speed gate, which is a narrow-band tracking filter that acquires the target's Doppler and tracks its changing Doppler frequency shift.

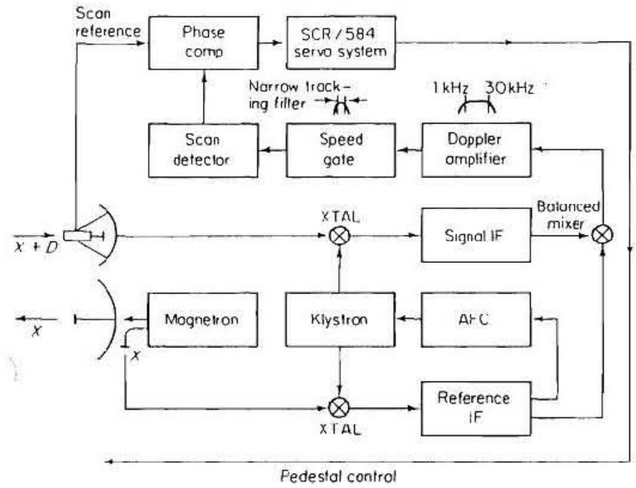


Fig. 3.5 Block diagram of a CW tracking- illuminator

Applications of CW radar:

1. The chief use of the simple, unmodulated CW radar is for the measurement of the relative velocity of a moving target, as in the police speed monitor or in the previously mentioned rate-of-climb meter for vertical-take-off aircraft.

2. In support of auto- mobile traffic, CW radar has been suggested for the control of traffic lights, regulation of tolbooths, vehicle counting, as a replacement for the fifth-wheel speedometer in vehicle testing as a sensor in antilock braking systems, and for collision avoidance.
3. For railways, CW radar can be used as a speedometer to replace the conventional axle-driven tachometer. In such an application it would be unaffected by errors caused by wheelslip on accelerating or wheelslide when braking.
4. It has been used for the measurement of railroad-freight-car velocity during humping operations in marshalling yards, and as a detection device to give track maintenance personnel advance warning of approaching trains.
5. CW radar is also employed for monitoring the docking speed of large ships.
6. It has also seen application for intruder alarms and for the measurement of the velocity of missiles, ammunition, and baseballs.
7. The principal advantage of a CW doppler radar over other (nonradar) methods of measuring speed is that there need not be any physical contact with the object whose speed is been measured.
8. In industry this has been applied to the measurement of turbine-blade vibration, the peripheral speed of grinding wheels, and the monitoring of vibrations in the cables of suspension bridges.
9. High-power CW radars for the detection of aircraft and other targets have been developed and have been used in such systems as the Hawk missile systems.

IF Doppler filter bank.

A relative wide band of frequencies called as bank of narrowband filters are used to measure the frequency of echo signal. When the doppler-shifted echo signal is known to lie somewhere within a relatively wide band of frequencies, a bank of narrowband filters (figure 3.6) spaced throughout the frequency range permits a measurement of frequency and improves the signal-to-noise ratio. The bandwidth of each individual filter is wide enough to accept the signal energy, but not so wide as to introduce more noise than need be. The centre frequencies of the filters are staggered to cover the entire range of Doppler frequencies. If the filters are spaced with their half-power points overlapped, the maximum reduction in signal-to-noise ratio of a signal lies midway between adjacent channels compared with the signal-to-noise ratio at band is 3 dB. The more filters used to cover the band, the less will be the maximum loss experienced, but the greater the probability of false alarm.

A bank of narrowband filters may be used after the detector in the video of the simple CW radar instead of in the IF. The improvement in signal-to-noise ratio with a video filter bank is not as good as can be obtained with an IF filter bank, but the ability to measure the magnitude of doppler frequency is still preserved. Because of fold over, a frequency which lies to one side of the IF carrier appears, after detection, at the same video frequency as one which lies an equal amount on the other side of the IF. Therefore the sign of the doppler shift is lost with a video filter bank, and it cannot be directly determined whether the Doppler frequency corresponds to an approaching or to a receding target. (The sign of the Doppler may he determined in the video by other means, as described later.) One advantage of the fold over in the video is that only half the number of filters are required than in the IF filter bank.

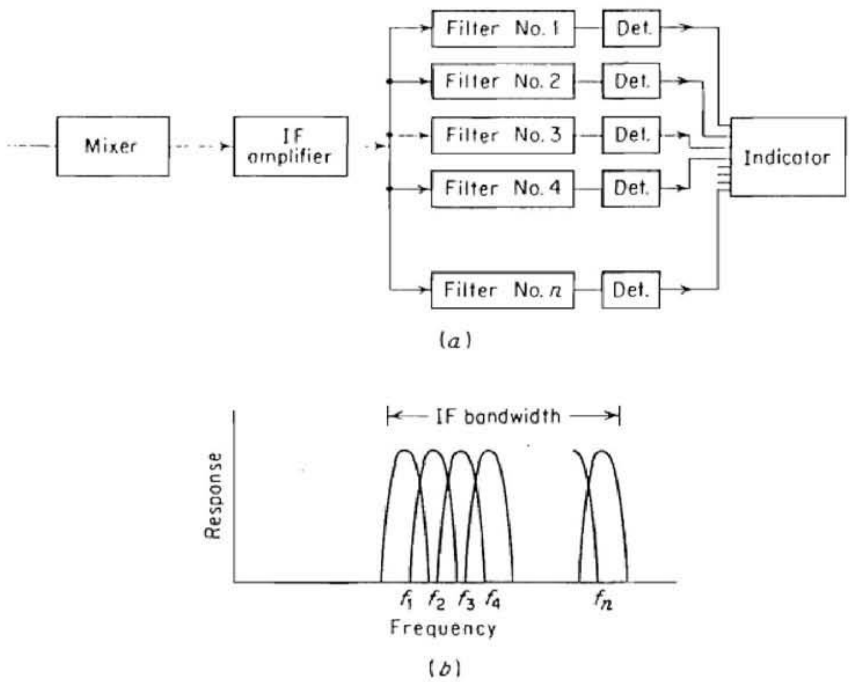


Fig.3.6 (a) Block diagram of IF doppler filter bank; (b) frequency-response characteristic of doppler filter bank.

Isolation between transmitter and receiver

A single antenna serves the purpose of transmission and reception in the simple CW radar. In principle, a single antenna may be employed since the necessary isolation between the transmitted and the received signals is achieved via separation in frequency as a result of the doppler effect. In practice, it is not possible to eliminate completely the transmitter leakage. However, transmitter leakage is not always undesirable. A moderate amount of leakage entering the receiver along with the echo signal supplies the reference necessary for the detection of the doppler frequency shift. If a leakage signal of sufficient magnitude were not present, a sample of the transmitted signal would have to be deliberately introduced into the receiver to provide the necessary reference frequency.

There are two practical effects which limit the amount of transmitter leakage power which can be tolerated at the receiver. These are (1) the maximum amount of power the receiver input circuitry can withstand before it is physically damaged or its sensitivity reduced (burnout) and (2) the amount of transmitter noise due to hum, microphonics, stray pick-up, and instability which enters the receiver from the transmitter. The additional noise introduced by the transmitter reduces the receiver sensitivity. Except where the CW radar operates with relatively low transmitter power and insensitive receivers, additional isolation is usually required between the transmitter and the receiver if the sensitivity is not to be degraded either by burnout or by excessive noise.

The amount of isolation required depends on the transmitter power and the accompanying transmitter noise as well as the ruggedness and the sensitivity of the receiver. For example, if the safe value of power which might be applied to a receiver were 10 mW and if the transmitter power were 1 kW, the isolation between transmitter and receiver must be at least 50 dB.

The amount of isolation needed in a long-range CW radar is more often determined by the noise that accompanies the transmitter leakage signal rather than by any damage cause by high power. For example, suppose the isolation between the transmitter and receivers were such that 10 mW of leakage signal appeared at the receiver. If the minimum detectable signal were 10^{-13} watt (100 dB below 1 mW), the transmitter noise must be at least 110 dB (preferably 120 or 130 dB) below the transmitted carrier.

The amount of isolation which can be readily achieved between the arms of practical hybrid junctions such as the magic-T, rat race, or short-slot coupler is of the order of 20 to 30 dB.An important factor which limits the use of isolation devices with a common antenna is the reflections produce in the transmission line by the antenna.

Doppler direction is identified with FMCW radar.

A block diagram shown in figure 3.7 illustrating the principle of the FM-CW .A Portion of the transmitter signal acts as the reference signal required to produce the beat frequency. It is introduced directly into the receiver via a cable or other direct connection. Ideally the isolation between transmitting and receiving antennas is made sufficiently large so as to reduce to a negligible level the transmitter leakage signal which arrives at the receiver via the coupling between antennas. The beat frequency is amplified and limited to remove any amplitude fluctuations. The frequency of the amplitude-limited beat note is measured with a cycle-counting frequency meter calibrated in distance.

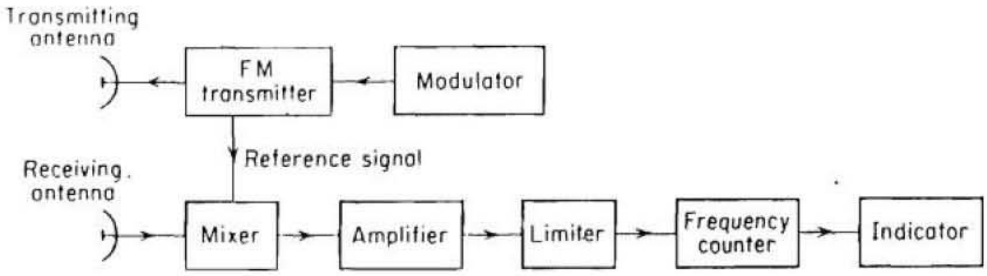


Fig.3.7 Block diagram of FM-CW radar

In the above, the target was assumed to be stationary. If this assumption is not applicable, a doppler frequency shift will be superimposed on the FM range beat note and an erroneous range measurement results. The doppler frequency shift causes the frequency-time plot of the echo signal to be shifted up or down (figure 3.8(a)). On one portion of the frequency-modulation cycle the heat frequency (figure 3.8 (b)) is increased by the doppler shift, while on the other portion it is decreased. If for example, the target is approaching the radar, the beat frequency f_b(up) produced during the increasing, or up, portion of the FM cycle will be the difference between the beat frequency due to the range f_r, and the doppler frequency shift f_d. Similarly, on the decreasing portion, the beat frequency, f_b(down) is the sum of the two.

$$f_b(up) = f_r - f_d$$

$$f_b(down) = f_r + f_d$$

The range frequency f_r, may be extracted by measuring the average beat frequency; that is,

$$f_r = 1/2[f_b(up) + f_b(down)].$$

If f_b(up) and f_b(down) are measured separately, for example, by switching a frequency counter every half modulation cycle, one-half the difference between the frequencies will yield the doppler frequency. This assumes $f_r > f_d$.

If, on the other hand, $f_r < f_d$ such as might occur with a high-speed target at short range, the roles of the averaging and the difference-frequency measurements are reversed; the averaging meter will measure Doppler velocity, and the difference meter, range. If it is not known that the roles of the meters are reversed because of a change in the inequality sign between f_r and f_d an incorrect interpretation of the measurements may result.

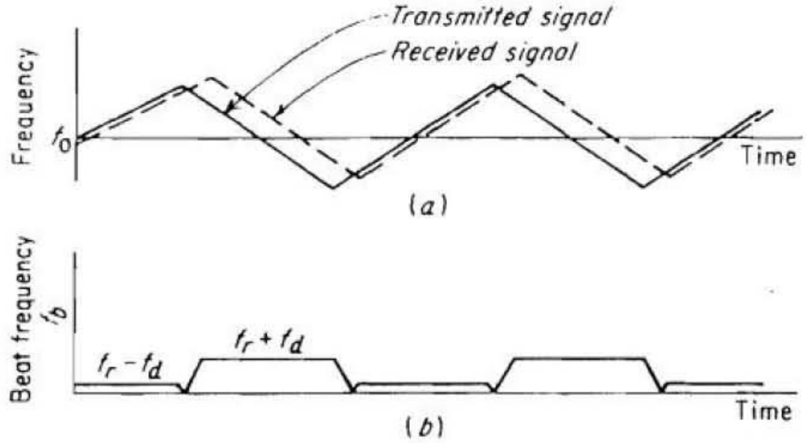


Fig.3.8 Frequency-time relation-ships in FM-CW radar when the $f_r + f_d$ received signal is shifted in frequency by the doppler effect (a) Transmitted (solid curve) and echo (dashed curve); (b) beat frequency

Range and Doppler measurement for an FMCW radar

In the frequency-modulated CW radar (abbreviated as FM-CW), the transmitter frequency is changed as a function of time in a known manner. Assume that the transmitter frequency increases linearly with time, as shown by the solid line in figure 3.9 (a). If there is a reflecting object at a distance R, an echo signal will return after a time T = 2R/c. The dashed line in the figure represents the echo signal. If the echo signal is heterodyned with a portion of the transmitter signal in a nonlinear element such as a diode, a beat note f_b will be produced. If there is no doppler frequency shift, the beat note (difference frequency) is a measure of the target's range and $f_b = f_r$ where f_r is the beat frequency due only to the target's range. If the rate of change of the carrier frequency is f_0, the beat frequency is $f_r = f_0 T = 2 R f_0 / c$

In any practical CW radar, the frequency cannot be continually changed in one direction only. Periodicity in the modulation is necessary, as in the triangular frequency-modulation waveform shown in figure 3.9 (b). The modulation need not necessarily be triangular; it can be saw tooth, sinusoidal, or some other shape. The resulting beat frequency as a function of time is shown in figure 3.9 (c) for triangular modulation. The beat note is of constant frequency except at the turn-around region. If the frequency is modulated at a rate f_m over a range Δf, the beat frequency is $f_r = 2 * 2 R f_m / c = 4 R f_m \Delta f / c$

Thus the measurement of the beat frequency determines the range R.

$R = c f_r / 4 f_m \Delta f$

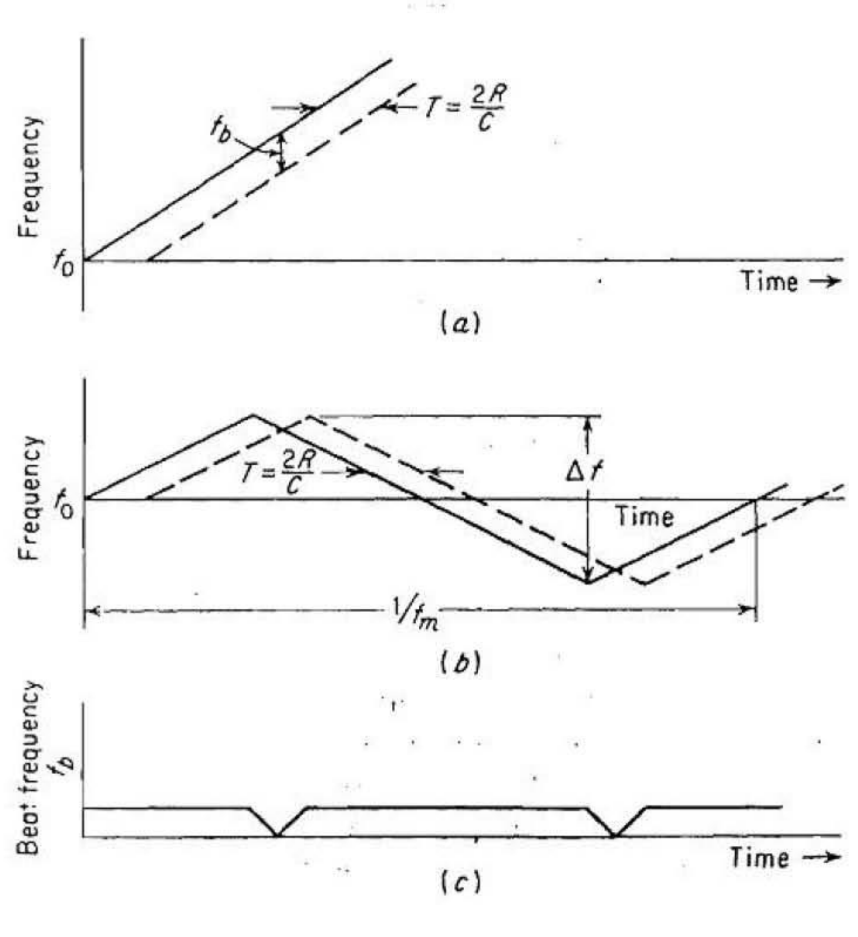


Fig.3.9 Frequency-time relationships in FM-CW radar. Solid curve represents transmitted signal, dashed curve represents echo. (a) Linear frequency modulation; (b) triangular frequency modulation; (c) beat note of (b).

Principle of operation of FMCW Altimeter.

The FM-CW radar principle is used in the aircraft radio altimeter to measure height above the surface of the earth. The large backscatter cross section and the relatively short ranges required of altimeters permit low transmitter power and low antenna gain. Since the relative motion between the aircraft and ground is small, the effect of the doppler frequency shift may usually be neglected. The band from 4.2 to 4.4 G Hz is reserved for radio altimeters, although they have in the past operated at UHF. The transmitter power is relatively low and can be obtained from a CW magnetron, a backward-wave oscillator, or a reflex klystron, but these have been replaced by the solid state transmitter.

The altimeter can employ a simple homodyne receiver, but for better sensitivity and stability the superheterodyne is to be preferred whenever its more complex construction can be tolerated. A block diagram of the FM-CW radar with a sideband superheterodyne receivers shown in figure 3.10. A portion of the frequency-modulated transmitted signal is applied to a mixer along with the oscillator signal. The selection of the local-oscillator frequency is a bit different from that in the usual superheterodyne receiver. The local-oscillator frequency f_{IF} should be the same as the intermediate frequency used in the receiver, whereas in the conventional superheterodyne the LO frequency is of the same order of magnitude as the RF signal. The output of the mixer consists of the varying transmitter frequency $f_o(t)$ plus two sideband frequencies, one on either side of $f_o(t)$ and separated from $f_o(t)$ by the local-oscillator frequency f_{IF} . The filter selects the lower sideband $f_o(t) - f_{IF}$ and rejects the carrier and the upper sideband. The sideband that is passed by the filter is modulated in the same fashion as the transmitted signal. The sideband filter must have sufficient bandwidth to pass the modulation, but not the carrier or other sideband. The filtered sideband serves the function of the local oscillator.

When an echo signal is present, the output of the receiver mixer is an IF signal of frequency $f_{IF} + f_b$ where f_b is composed of the range frequency f_r and the doppler velocity frequency f_d. The IF signal is amplified and applied to the balanced detector along with the local-oscillator signal f_{IF}. The output of the detector contains the beat frequency (range frequency and the doppler velocity frequency), which is amplified to a level where it can actuate the frequency-measuring circuits.

In figure 3.10, the output of the low-frequency amplifier is divided into two channels: one feeds an average-frequency counter to determine range, the other feeds a switched frequency counter to determine the doppler velocity (assuming fr > fd) Only the averaging frequency counter need be used in an altimeter application, since the rate of change of altitude is usually small.

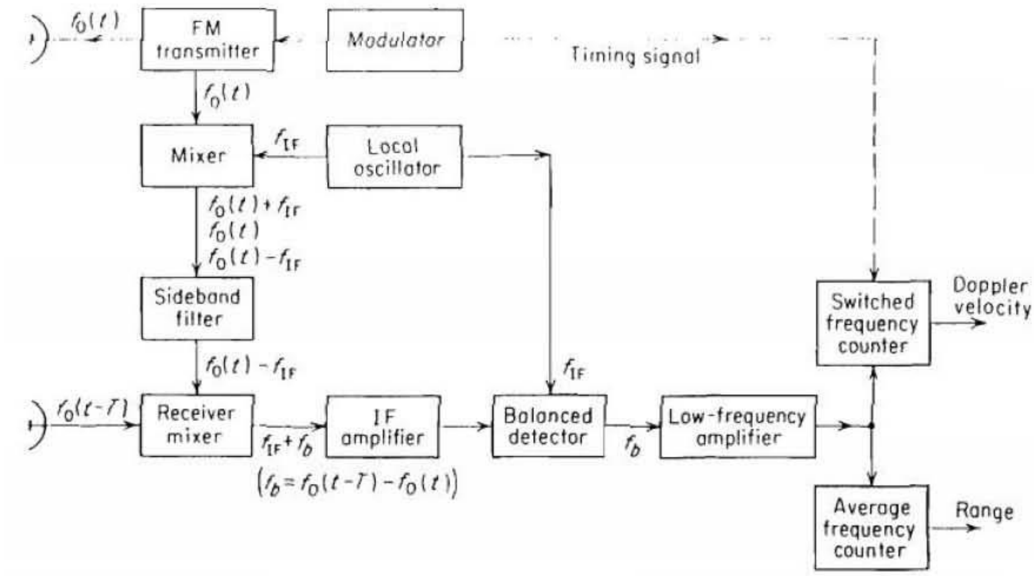


Fig.3.10 Block diagram of FM-CW radar using sideband superheterodyne receiver

Operation of sinusoidally modulated FM-CW radar extracting the third harmonic.

The block diagram for sinusoidally modulated FM-CW radar extracting the third harmonic is shown in figure 3.11

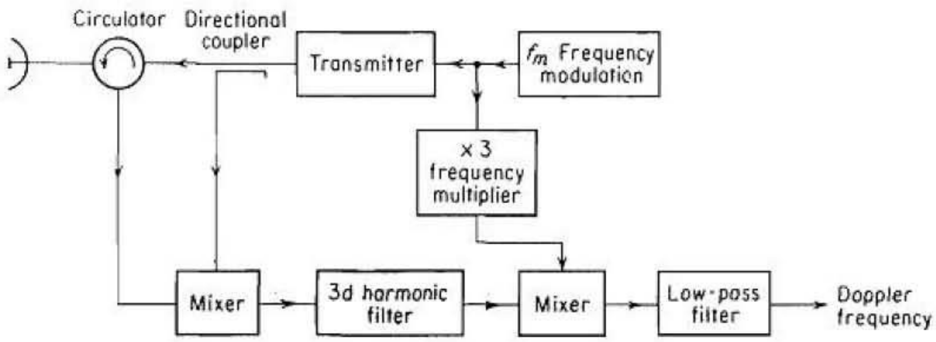


Fig.3.11 Sinusoidally modulated FM-CW radar extracting the third harmonic

The ability of the FM-CW radar to measure range provides an additional basis for obtaining isolation. Echoes from short-range targets-including the leakage signal-may be attenuated relative to the desired target echo from longer ranges by properly processing the difference-frequency signal obtained by heterodyning the transmitted and received signals.

If the CW carrier is frequency-modulated by a sine wave, the difference frequency obtained by heterodyning the returned signal with a portion of the transmitter signal may be expanded in a trigonometric series whose terms are the harmonics of the modulating frequency f_m

Assume the form of the transmitted signal to be

$$\sin\left(2\pi f_0 t + \frac{\Delta f}{2 f_m} \sin 2\pi f_m t\right)$$

where ,

f_o = carrier frequency

f_m = modulation frequency

Δf = frequency excursion (equal to twice the frequency derivation)

The difference frequency signal may be written as

$$\begin{aligned} v_D = {} & J_0(D) \cos(2\pi f_d t - \phi_0) + 2J_1(D) \sin(2\pi f_d t - \phi_0) \cos(2\pi f_m t - \phi_m) \\ & - 2J_2(D) \cos(2\pi f_d t - \phi_0) \cos 2(2\pi f_m t - \phi_m) \\ & - 2J_3(D) \sin(2\pi f_d t - \phi_0) \cos 3(2\pi f_m t - \phi_m) \\ & + 2J_4(D) \cos(2\pi f_d t - \phi_0) \cos 4(2\pi f_m t - \phi_m) + 2J_5(D) \cdots \end{aligned}$$

where J_0, J_1, J_2, etc = Bessel functions of first kind and order 0, 1, 2, etc., respectively

$D = (\Delta f / f_m) \sin 2\pi f_m R_0 / c$

R_0 = distance to target at time t = 0 (distance that would have been measured if target were stationary)

c = velocity of propagation f_d = doppler frequency shift

v_r = relative velocity of target with respect to radar

φ_0 = phase shift approximately equal to angular distance $2\pi f_0 R_0 / c$

φ_m = phase shift approximately equal to $2\pi f_m R_0 / c$

The difference-frequency signal consists of a doppler-frequency component of amplitude J o(D) and a series of cosine waves of frequency f_m, 2 f_m, etc. Each of these harmonics of f_m is modulated by a doppler-frequency component with amplitude proportional to J n(D). The product of the doppler-frequency factor times the nth harmonic factor is equivalent to a suppressed-carrier double-sideband modulation.

Multiple frequency CW radar

The multiple frequency CW radar is used to measure the accurate range. The transmitted waveform is assumed to consist of two continuous sine waves of frequency f_1 and f_2 separated by an amount Δf. Let the amplitudes of all signals are equal to unity. The voltage waveforms of the two components of the transmitted signal v_{1r} and v_{2r}, may be written as

$$v_{1r} = \sin(2\pi f_1 t + \varphi_1)$$

$$v_{2r} = \sin(2\pi f_2 t + \varphi_2)$$

where φ_1 and φ_2 are arbitrary (constant) phase angles.

The echo signal is shifted in frequency by the doppler effect. The form of the doppler-shifted signals at each of the two frequencies f_1 and f_2 may be written as

$$v_{1R} = \sin\left[2\pi(f_1 \pm f_{d1})t - \frac{4\pi f_1 R_0}{c} + \phi_1\right]$$

$$v_{2R} = \sin\left[2\pi(f_2 \pm f_{d2})t - \frac{4\pi f_2 R_0}{c} + \phi_2\right]$$

Where,

Ro = range to target at a particular time $t = t_0$ (range that would be measured if target were not moving)

f_{d1} = doppler frequency shift associated with frequency f_1

f_{d2} = doppler frequency shift associated with frequency f_2

Since the two RF frequencies f_1 , and f_2 are approximately the same the doppler frequency shifts f_{d1} and f_{d2} are approximately equal to one another. Therefore

$$f_{d1} = f_{d2} = f_d$$

The receiver separates the two components of the echo signal and heterodynes each received signal component with the corresponding transmitted waveform and extracts the two doppler-frequency components given below:

$$v_{1D} = \sin\left(\pm 2\pi f_d t - \frac{4\pi f_1 R_0}{c}\right)$$

$$v_{2D} = \sin\left(\pm 2\pi f_d t - \frac{4\pi f_2 R_0}{c}\right)$$

The phase difference between these two components is

$$\Delta\phi = \frac{4\pi(f_2 - f_1)R_0}{c} = \frac{4\pi\ \Delta f R_0}{c}$$

Hence

$$R_0 = \frac{c\ \Delta\phi}{4\pi\ \Delta f}$$

A large difference in frequency between the two transmitted signals improves the accuracy of the range measurement since large Δf means a proportionately large change in $\Delta\varphi$ for a given range. However, there is a limit to the value of Δf, since $\Delta\varphi$ cannot be greater than 2π radians if the range is to remain unambiguous. The maximum unambiguous range $R_{unamb} = c / 2\Delta f$

The two-frequency CW radar is essentially a single target radar since only one phase difference can be measured at a time. If more than one target is present, the echo signal becomes complicated and the meaning of the phase measurement is doubtful.

Errors introduced in radar are measured.

The absolute accuracy of radar altimeters is usually of more importance at low altitudes than at high altitudes. Errors of a few meters might not be of significance when cruising at altitudes of 10 km, but are important if the altimeter is part of a blind landing system.

The theoretical accuracy with which distance can be measured depends upon the band-width of the transmitted signal and the ratio of signal energy to noise energy. In addition, measurement accuracy might be limited by such practical restrictions as the accuracy of the frequency-measuring device, the residual path-length error caused by the circuits and transmission lines, errors caused by multiple reflections and transmitter leakage, and the frequency error due to the turn-around of the frequency modulation.

A common form of frequency- measuring device is the cycle counter, which measures the number of cycles or half cycles of the beat during the modulation period. The total cycle count is a discrete- number since the counter is unable to measure fractions of a cycle. The discreteness of the frequency measurement gives rise to an error called the fixed error, or step error. It has also been called the quantization error, a more descriptive name. The average number of cycles N of the beat frequency f_b in one period of the modulation cycle f_m is $\overline{f_b} / f_m$, where the bar over , denotes time average.

$$R = c N / 4 \Delta f$$

Where,R = range (altitude), c = velocity of propagation. Δf = frequency excursion.

Since the output of the frequency counter N is an integer, the range will be an integral multiple of c / 4 Δ*f* and will give rise to a quantization error equal to

$$\delta R = c / 4 \Delta f$$

$$\delta R (m) = 75 / \Delta f (MHz)$$

Since the fixed error is due to the discrete nature of the frequency counter, its effects can be reduced by wobbling the modulation frequency or the phase of the transmitter output. Wobbling the transmitter phase results in a wobbling of the phase of the beat signal so that an average reading of the cycle counter somewhere between N and N + 1 will be obtained on a normal meter movement. In one altimeter, the modulation frequency was varied at a 10-Hz rate, causing the phase shift of the beat signal to vary cyclically with time.

The indicating system was designed so that it did not respond to the 10-Hz modulation directly, but it caused the fixed error to be averaged. Normal fluctuations in aircraft altitude due to uneven terrain, waves on the water, or turbulent air can also average out the fixed error provided the time constant of the indicating device is large compared with the time between fluctuations. Over smooth terrain, such as an airport runway, the fixed error might not be averaged out.

Target motion can cause an error in range equal to $v_r T_0$, where v_r, is the relative velocity and To is the observation time. The residual path error is the error caused by delays in the circuitry and transmission lines. Multipath signals also produce error. Reflections from the landing gear can also cause errors.

Signals in FM altimeter

Unwanted signals in FM altimeter are shown in figure 3.12

The unwanted signals include

1. The reflection of the transmitted signals at the antenna caused by impedance mismatch.
2. The standing-wave pattern on the cable feeding the reference signal to the receiver, due to poor mixer match.

3. The leakage signal entering the receiver via coupling between transmitter and receiver antennas. This can limit the ultimate receiver sensitivity, especially at high altitudes.
4. The interference due to power being reflected back to the transmitter, causing a change in the impedance seen by the transmitter. This is usually important only at low altitudes. It can be reduced by an attenuator introduced in the transmission line at low altitude or by a directional coupler or an isolator.
5. The double bounce signal.

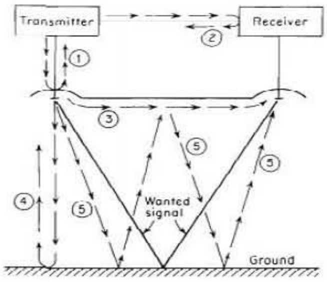


Fig.3.12 Unwanted signals in FM altimeter

UNIT-III

MTI AND PULSE DOPPLER RADAR

Principle of MTI radar

The radar which uses the concept of Doppler frequency shift for distinguishing desired moving targets from stationary objects i.e., clutter is called as MTI radar (Moving Target Indicator).The block diagram of MTI radar employing a power amplifier is shown in figure 3.1. The significant difference between this MTI configuration and that of Pulse Doppler radar is the manner in which the reference signal is generated. In figure 3.1, the coherent reference is supplied by an oscillator called the coho, which stands for coherent oscillator. The coho is a stable oscillator whose frequency is the same as the intermediate frequency used in the receiver.

In addition to providing the reference signal, the output of the coho f_C is also mixed with the local-oscillator frequency f_l. The local oscillator must also be a stable oscillator and is called stalo, for stable local oscillator. The RF echo signal is heterodyned with the stalo signal to produce the IF signal, just as in the conventional superheterodyne receiver. The stalo, coho, and the mixer in which they are combined plus any low-level amplification are called the receiver-exciter because of the dual role they serve in both the receiver and the transmitter.

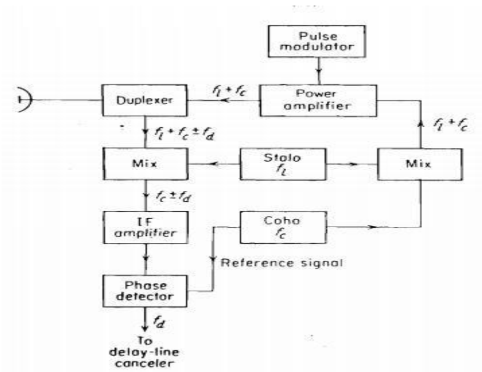


Fig. 3.1 Block diagram of MTI radar with power-amplifier transmitter

The characteristic feature of coherent MTI radar is that the transmitted signal must be coherent (in phase) with the reference signal in the receiver. This is accomplished in the radar system diagramed in figure 4.1 by generating the transmitted signal from the coho reference signal. The function of the stalo is to provide the necessary frequency translation from the IF to the transmitted (R F) frequency. Although the phase of the stalo influences the phase of the transmitted signal, any stalo phase shift is canceled on reception because the stalo that generates the transmitted signal also acts as the local oscillator in the receiver. The reference signal from the coho and the IF echo signal are both fed into a mixer called the pulse detector The phase detector differs from the normal amplitude detector since its output is proportional to the phase difference between the two input signals.

Any one of a number of transmitting-tube types might be used as the power amplifier. These include the triode, tetrode, klystron, traveling-wave tube, and the crossed-field amplifier.

Butterfly effect that is produced by MTI

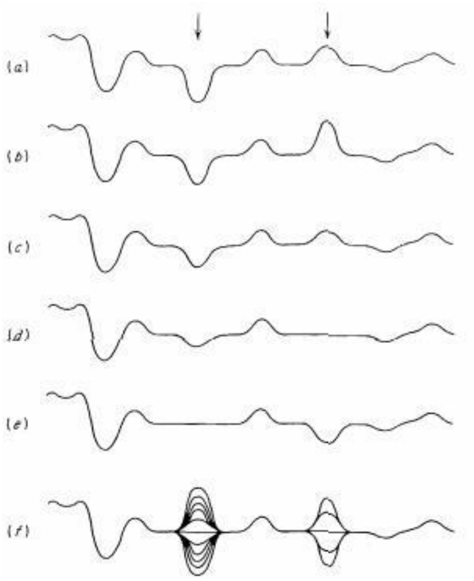


Fig 3.2 (a-e) Successive sweeps of an MTI radar A-scope display (echo amplitude as a function of time); (f) superposition of many sweeps; arrows indicate position of moving targets.

Moving targets may be distinguished from stationary targets by observing the video output on an A-scope (amplitude vs. range). A single sweep on an A-scope might appear as in figure 3.2(a). This sweep shows several fixed targets and two moving targets indicated by the two arrows. On the basis of a single sweep, moving targets cannot be distinguished from fixed targets. Successive A-scope sweeps (pulse-repetition intervals) are shown in figure 3.2(b) to (e). Echoes from fixed targets remain constant throughout but echoes from moving targets vary in amplitude from sweep to sweep at a rate corresponding to the doppler frequency. The superposition of the successive A-scope sweeps is shown in figure 3.2(f). The moving targets produce, with time, a butterfly effect on the A-scope.

Delay line canceller

The simple MTI delay-line canceler shown in figure 3.3 is an example of a time-domain filter. The capability of this device depends on the quality of the medium used is the delay line. The Pulse modulator delay line must introduce a time delay equal to the pulse repetition interval. For typical ground-based air-surveillance radars this might be several milliseconds. Delay times of this magnitude cannot be achieved with practical electromagnetic transmission lines. By converting the electromagnetic signal to an acoustic signal it is possible to utilize delay lines of a reasonable physical length since the velocity of propagation of acoustic waves is about 10^{-5} that of electromagnetic waves. After the necessary delay is introduced by the acoustic line, the signal is converted back to an electromagnetic signal for further processing.

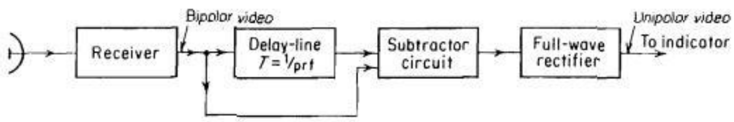


Fig.3.3 MTI receiver with delay-line canceler

The early acoustic delay lines developed during World War II used liquid delay lines filled with either water or mercury. Liquid delay lines were large and inconvenient to use. They were replaced in the mid-1950s by the solid fused-quartz delay line that used multiple internal reflections to obtain a compact device. These analog acoustic delay lines were, in turn supplanted in the early 1970's by storage devices based on digital computer technology. The use of digital delay lines requires that the output of the MTI receiver phase-detector be quantized into a sequence of digital words. The compactness and convenience of digital processing allows the implementation of more complex delay-line cancellers with filter characteristics not practical with analog methods.

One of the advantages of a time-domain delay-line canceller as compared to the more conventional frequency-domain filter is that a single network operates at all ranges and does not require a separate filter for each range resolution cell. Frequency-domain doppler filter- banks are of interest in some forms of MTI and pulse-doppler radar.

Filter characteristics of the delay-line canceller

Filter characteristics of the delay-line canceller. The delay-line canceller acts as a filter which rejects the d-c component of clutter. Because of its periodic nature, the filter also rejects energy in the vicinity of the pulse repetition frequency and its harmonics.

The video signal received from a particular target at a range R_0 is

$$V_1 = k \sin (2\pi f_d t - \varphi_0)$$

Where, φ_0 = phase shift k = amplitude of video signal.

The signal from the previous transmission, which is delayed by a time T = pulse repetition interval, is

$$V_2 = k \sin (2\pi f_d(t - T) - \varphi_0)$$

Everything else is assumed to remain essentially constant over the interval T so that k is the same for both pulses. The output from the subtractor is

$$V = V_1 - V_2 = 2*k \sin \pi f_d T \cos [2\pi f_d(t - T / 2) - \varphi_0]$$

It is assumed that the gain through the delay-line canceller is unity. The output from the canceller V consists of a cosine wave at the doppler frequency f_d with an amplitude $2*k \sin \pi f_d T$. Thus the amplitude of the canceled video output is a function of the Doppler frequency shift and the pulse-repetition interval, or prf. The magnitude of the relative frequency-response of the delay-line canceller [ratio of the amplitude of the output from the delay-line canceller, $2*k \sin \pi f_d T$, to the amplitude of the normal radar video kj is shown in figure 3.4.

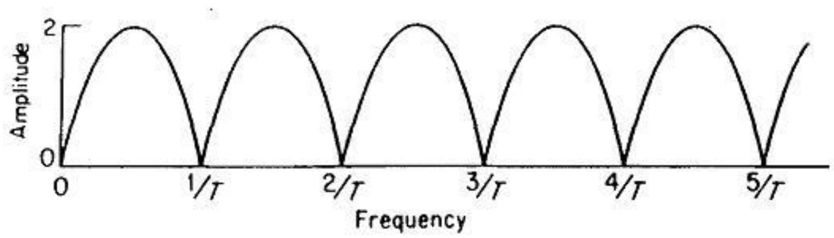


Fig. 3.4 Frequency response of the single delay-line canceler; T = delay time = 1 / f_p

1. Blind speed is defined as the radial velocity of the target at which the MTI response is zero.
2. It is also defined as the radial velocity of the target which results in a phase difference of exactly 2π radians between successive pulses.
3. Blind speed is defined as the radial velocity of the target at which no shift appears

making the target appearing stationary and echows from the targets are cancelled.

The speeds at which these occur are called the blind speed of the radar. These are

$$v_n = \frac{n\lambda}{2T};\ n=0,1,2 \qquad \text{or } v_n = \frac{n\lambda f_p}{2}$$

For λ in metres, f_p in Hz and vn in knots we have the following: $v_n = n\lambda f_p/(1.02) \approx n\lambda f_p$

$$v_n = n\lambda f_p$$

The first blind speeds in knots is given by $v_{b1} = 0.97\lambda f_p = \lambda f_p$

The blind speeds are one of the limitations of pulse MTI radar which do not occur with CW radar. They are present in pulse radar because doppler is measured by discrete samples (pulses) at the prf rather than continuously. If the first blind speed is to be greater than the maximum radial velocity expected from the target, the product λf_p must be large. Thus the MTI radar must operate at long wavelengths (low frequencies) or with high pulse repetition frequencies.

In figure 3.5, the first blind speed v_1 is plotted as a function of the maximum unambiguous range ($R_{unamb} = cT / 2$), with radar frequency as the parameter. If the first blind speed were 600 knots, the maximum unambiguous range would be 130 nautical miles at a frequency of 300 MHz (UHF).

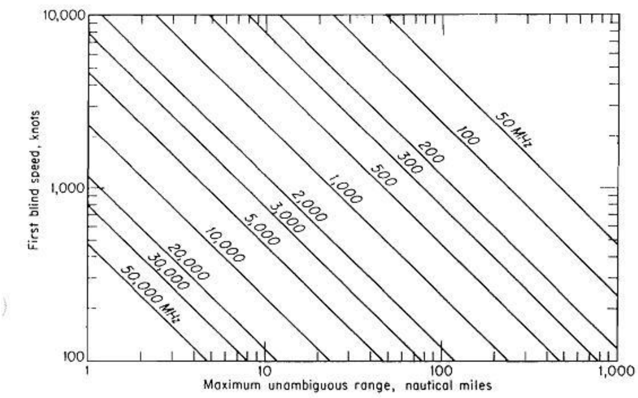


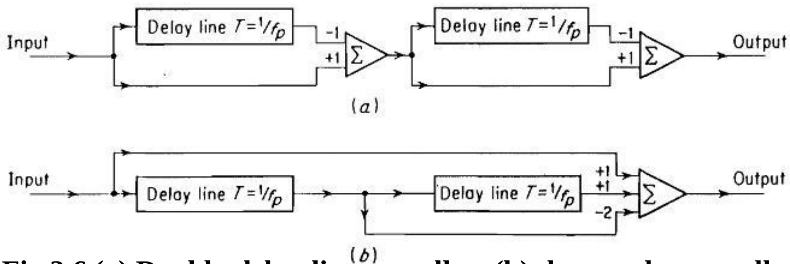
Fig. 3.5 Plot of MTI radar first blind speed as a function of maximum unambiguous range

Double delay line canceller and three pulse canceller

The frequency response of a single-delay-line canceller does not always have as broad a clutter-rejection null as might be desired in the vicinity of d-c. The clutter-rejection notches may be widened by passing the output of the delay-line canceller through a second delay-line canceller as shown in figure 3.6. The output of the two single-delay- line cancellers in cascade is the square of that from a single canceller. Thus the frequency response is 4 $\sin^2 \pi f_d T$. The configuration of figure 4.6. is called a double-delay-line canceller, or simply a double canceller. The relative response of the double canceller compared with that of a single-delay-line canceller is shown in figure 4.6. The finite width of the clutter spectrum is also shown in this figure so as to illustrate the additional cancellation of clutter offered by the double canceler.

The two-delay-line configuration of figure 4.6. has the same frequency-response characteristic as the double-delay-line canceller. The operation of the device is as follows. A signal f (t) is inserted into the adder along with the signal from the preceding pulse period, with its amplitude weighted by the factor - 2, plus the signal from two pulse periods previous. The output of the adder is therefore

$$f(t) – 2f(t + T) + f(t + 2T$$



.Fig.3.6 (a) Double-delay-line canceller; (b) three-pulse canceller

which is the same as the output from the double-delay-line canceller

$$f(t) – f(t + T) - f(t + T) + f(t + 2T)$$

This configuration is commonly called the three-pulse canceller.

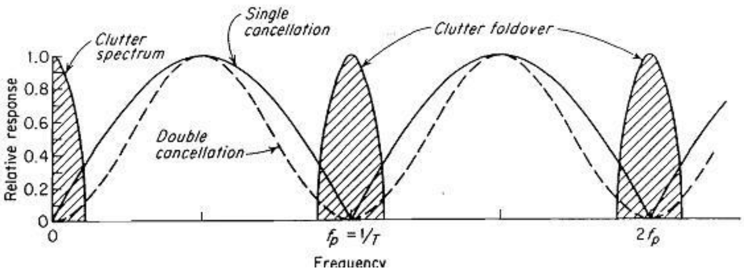


Fig.3.7. Relative frequency response of the single-delay-line canceller (solid curve) and the double- delay-line canceller (dashed curve)

Transversal filters

The three-pulse canceller is an example of a transversal filter. Its general form with N pulses and N - 1 delay lines is shown in figure 3.8. It is also sometimes known as a feed forward filter, a non recursive filter, a finite memory filter or a tapped delay-line filter. The weights w_i for a three-pulse canceler utilizing two delay lines arranged as a transversal filter are 1. -2, 1. The frequency response function is proportional to $\sin^2 \pi f_d T$. A transversal filter with three delay lines whose weights are 1, - 3, 3, - 1 gives a $\sin^3 \pi f_d T$ response. This is a four-pulse canceler. Its response is equivalent to a triple canceler consisting of a cascade of three single delay-line cancellers. Note the potentially confusing nomenclature.

A cascade configuration of three delay line's, each connected as a single canceler, is called a triple canceler, but when connected as a transversal filter it is called a four-pulse canceler. The weights for a transversal filter with n delay lines that gives a response $\sin^n \pi f_d T$ are the coefficients of the expansion of (1 - x)", which are the binomial coefficients with alternating signs

$$w_i = (-1)^{i-1} \frac{n!}{(n-i+1)!\,(i-1)!}, \qquad i = 1, 2, \ldots, n+1$$

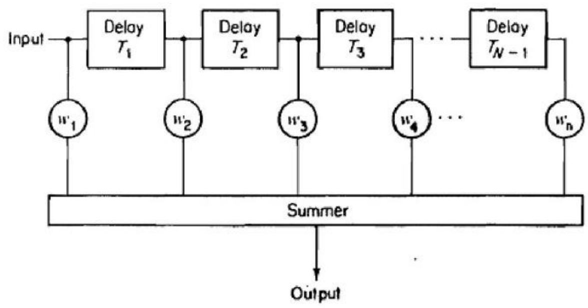


Fig 3.8 General form of a transversal (or non recursive) filter for MTI signal processing

MTI radar using range gates and filters

The block diagram of the video of an MTI radar with multiple range gates followed by clutter-rejection filters is shown in figure 3.9. The output of the phase detector is sampled sequentially by the range gates. Each range gate opens in sequence just long enough to sample the voltage of the video waveform corresponding to a different range interval in space. The range gate acts as a switch or a gate which opens and closes at the proper time. The range gates are activated once each pulse-repetition interval.

The output for a stationary target is a series of pulses of constant amplitude. An echo from a moving target produces a series of pulses which vary in amplitude according to the doppler frequency. The output of the range gates is stretched in a circuit called the boxcar generator, or sample-and-hold circuit, whose purpose is to aid in the filtering and detection process by emphasizing the fundamental of the modulation frequency and eliminating harmonics of the pulse repetition frequency. The clutter- rejection filter is a bandpass filter whose bandwidth depends upon the extent of the expected clutter spectrum.

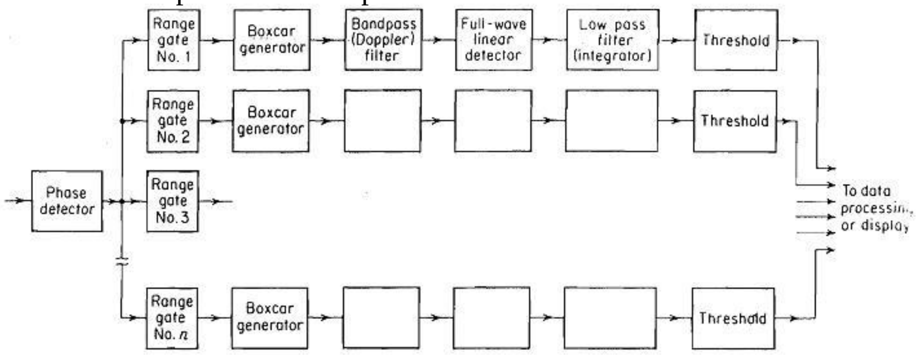


Fig. 3.9 Block diagram of MTI radar using range gates and filters

Following the doppler filter is a full-wave linear detector and an integrator (a low-pass filter). The purpose of the detector is to convert the bipolar video to unipolar video. The output of the integrator is applied to a threshold-detection circuit. Only those signals which cross the threshold are reported as targets. Following the threshold detector, the outputs from each of the range channels must be properly combined for display on the PPI or A-scope or for any other appropriate indicating or data-processing device. The CRT display from this type of MTI radar appears" cleaner" than the display from a normal MTI radar, not only because of better clutter rejection, but also because the threshold device eliminates many of the unwanted false alarms due to noise. The frequency-response characteristic of the range-gated MTl might appear as in figure 4.10. The shape of the rejection band is determined primarily by the shape of the bandpass filter of figure 4.10.

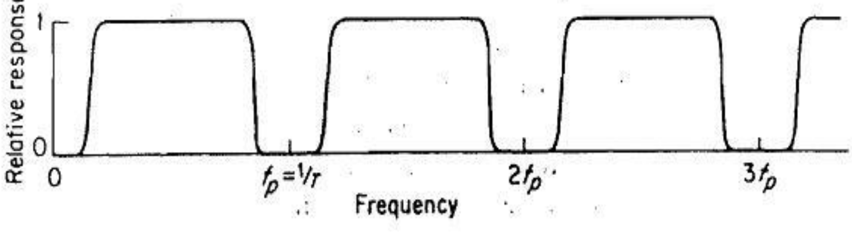


Fig.3.10 Frequency-response characteristic of an MTI using range gates and filters

Non coherent MTI radar

The composite echo signal from a moving target and clutter fluctuates in both phase and amplitude. The coherent MTI and the pulse-doppler radar make use of the phase fluctuations in the echo signal to recognize the doppler component produced by a moving target. In these systems, amplitude fluctuations are removed by the phase detector. The operation of this type of radar, which may be called coherent MTI, depends upon a reference signal at the radar receiver that is coherent with the transmitter signal.

It is also possible to use the amplitude fluctuations to recognize the doppler component produced by a moving target. MTI radar which uses amplitude instead of phase fluctuations is called noncoherent. It has also been called externally coherent, which is a more descriptive name. The noncoherent MTI radar does not require an internal coherent reference signal or a phase detector as does the coherent form of MTI. Amplitude limiting cannot be employed in the noncoherent MTI receiver, else the desired amplitude fluctuations would be lost. Therefore the IF amplifier must be linear, or if a large dynamic range is required, it can be logarithmic.

The detector following the IF amplifier is a conventional amplitude detector. The phase detector is not used since phase information is of no interest to the noncoherent radar. The local oscillator of the noncoherent radar does not have to be as frequency-stable as in the coherent MTI.

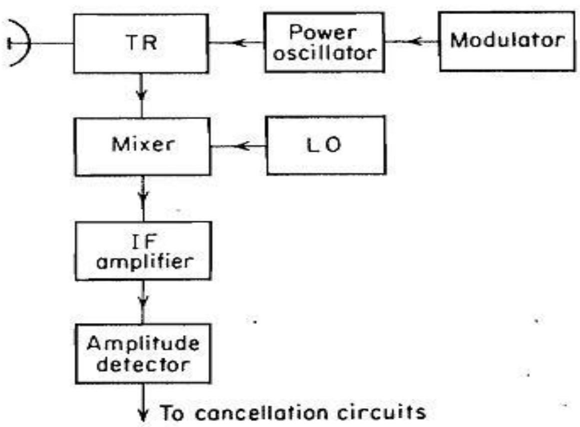


Fig.3.11 Block diagram of a non coherent MTI radar

The output of the amplitude detector is followed by an MTI processor such as a delay-line canceler. The doppler component contained in the amplitude fluctuations may also be detected by applying the output of the amplitude detector to an A-scope.

The advantage of the noncoherent MTI is its simplicity; hence it is attractive for those applications where space and weight are limited. Its chief limitation is that the target must be in the presence of relatively large clutter signals if moving-target detection is to take place. Clutter echoes may not always be present over the range at which detection is desired.

Limitations of MTI radar.

1. Equipment instabilities.
2. Scanning modulation.
3. Internal fluctuations of clutter

1. Equipment instabilities:

Pulse-to-pulse changes in the amplitude, frequency, or phase of the transmitter signal, changes in the stalo or coho oscillators in the receiver, jitter in the timing of the pulse transmission, variations in the time delay through the delay lines, and changes in the pulse width can cause the apparent frequency spectrum from perfectly stationary clutter to broaden and thereby lower the improvement factor of an MTI radar. The stability of the equipment in a MTI radar must be considerably better than that of an ordinary radar. It can limit the performance of the MTI radar if sufficient care is not taken in design, construction, and maintenance.

2. Scanning modulation:

As the antenna scans by a target, it observes the target for a finite time equal to

$$t_0 = n_B / f_p = \theta_B / \theta_S$$

where,

n_B = number of hits received. f_p = pulse repetition frequency,

θ_B = antenna beamwidth and θ_S = antenna scanning rate.

The received pulse train of finite duration to has a frequency spectrum (which can be found by taking the Fourier transform of the waveform) whose width is proportional to l/t_0. Therefore, even if the clutter were perfectly stationary, there will still be a finite width to the clutter spectrum because of the finite time on target. If the clutter spectrum is too wide because the observation time is too short, it will affect the improvement factor. This limitation has sometimes been called scanning fluctuations or scanning modulation.

3. Internal Fluctuation of Clutter*:* Absolutely stationary clutter - buildings, water towers, mountains, bare hills, dynamic clutter - trees, vegetation, sea, rain, chaff can limit the performance of MTI. Most fluctuating clutter targets can be represented by a model consisting of many independent scatterers located in the resolution cell.

Staggered pulse repetition frequencies

The pulse repetition frequency in which the switching is pulse to pulse is known as staggered PRF.The use of more than one pulse repetition frequency offers additional flexibility in the design of MTI doppler filters. It not only reduces the effect of the blind speeds but it also allows a sharper low-frequency cutoff in the frequency response than might be obtained with a cascade of single-delay-line cancelers with $\sin^n \pi f_d T$ response.

The blind speeds of two independent radars operating at the same frequency will be different if their pulse repetition frequencies are different. Therefore, if one radar were blind to moving targets, it would be unlikely that the other radar would be blind also. Instead of using two separate radars, the same result can be obtained with one radar which time-shares its pulse repetition frequency between two or more different values (multiple prf's). The pulse repetition frequency might be switched every other scan or every time the antenna is scanned a half beamwidth, or the period might be alternated on every other pulse. When the switching is pulseto pulse, it is known as a staggered prf.

An example of the composite (average) response of an MTI radar operating with two eparate pulse repetition frequencies on a time-shsred basis is shown in figure 3.12. The pulse repetition frequencies are in the ratio of 5 : 4. Note that the first blind speed of the composite response is increased several times over what it would be for a radar operating on only a single pulse repetition frequency. Zero response occurs only when the blind speeds of each prf coincide. In the example of figure.4.12, the blind speeds are coincident for $4/T_1 = 5/T_2$.

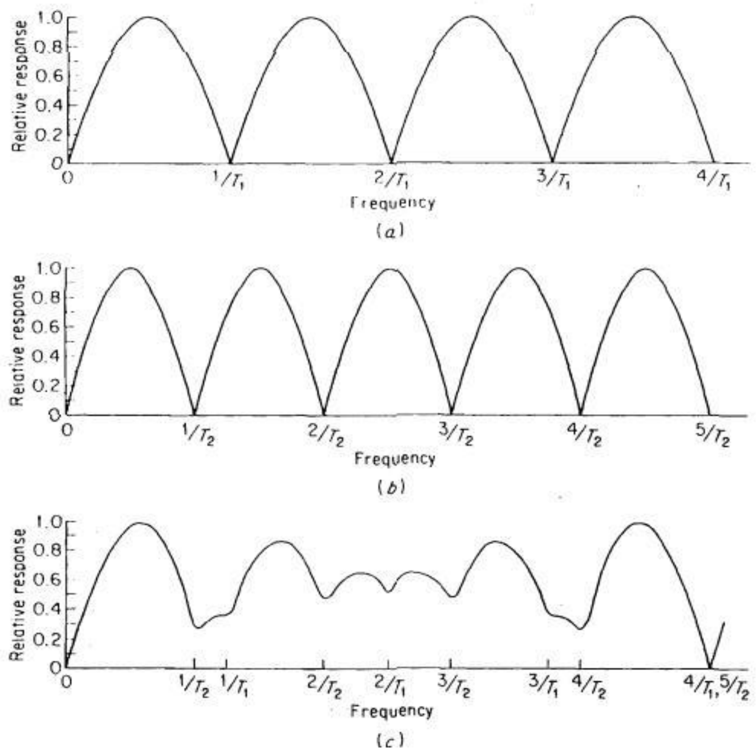


Fig. 3.12 (a) Frequency-response of a single-delay-line canceler for $f_p = 1/T_1$; (b) same for $f_p = 1/T_2$ (c) composite response with $4/T_1 = 5/T_2$.

The closer the ratio T_1:T_2 approaches unity, the greater will be the value of the first blind speed. However, the first null in the vicinity of $f_d = 1/T_1$ becomes deeper. Thus the choice of T_1/T_2 is a compromise between the value of the first blind speed and the depth of the nulls within the filter pass band. The depth of the nulls can be reduced and the first blind speed increased by operating with more than two interpulse periods.

A disadvantage of the staggered prf is its inability to cancel second-time-around clutter echoes. Such clutter does not appear at the same range from pulse to pulse and thus produces uncanceled residue. Second-time-around clutter echoes can be removed by use of a constant prf. providing there is pulse-to-pulse coherence as in the power amplifier form of MTI.

Difference between Pulse Dopler Radar and MTI Radar

Pulse doppler radar, in this the radar send the pulse train to detect the position of target and MTI (moving target indicator) in which it detect the target which is moving but pulse radar can detect the moving target but there is a disadvantage that the problem of blind speed arises and pulse radar dosn't continuosly transmit the pulse after transmitting wait for receiving in this time it doesn't transmit any pulse.

Pulse doppler radar and MTI radar both are used to find the target range by using doppler
effect(doppler shift) .but MTI radar uses low PRF whereas pulse doppler uses high PRF.

The different between MTI radar and PD radar is a unique even though they all relay on Doppler principle , but the MTI radar determine moving targets by detecting the phase and amplitude of the received wave and compare it with saved replica of the original transmitted wave but at opposite phase , so if the target are not moving then the phase and amplitude of the 2 signals will match but at different value will result of canceling each other, but if the 2 signals are not matched they will give positive or negative value and that is indication of moving target.

Pulse Doppler radar has anther interest, it is interested in the changes happen to the transmitted wave (DOPPLER SHIFT) either it will be compressed if the target moving toward the radar.

The PD radar are not interested in the transmitted frequency any more after it has been transmitted but it does set filters around it at the expected reflected frequency.

MTI Radar with Power Oscillator

A block diagram of an MTI radar (with a power oscillator) is shown in figure 3.13. A portion of the transmitted signal is mixed with the stalo output to produce an IF beat signal whose phase is directly related to the phase of the transmitter. This IF pulse is applied to the coho and causes the phase of the coho CW oscillation to "lock" in step with the phase of the IF reference pulse. The phase of the coho is then related to the phase of the transmitted pulse and may be used as the reference signal for echoes received from that particular transmitted pulse. Upon the next transmission another IF locking pulse is generated to relock the phase of the CW coho until the next locking pulse comes along.

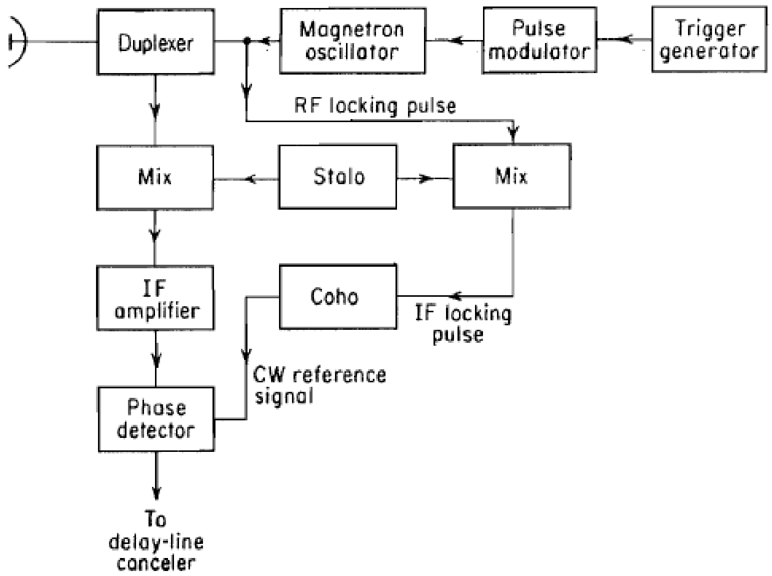


Fig. 3.13 Block diagram of MTI Radar with power oscillator

UNIT-IV

TRACKING RADAR

Tracking principles.

A tracking-radar system

(1) measures the coordinates of a target and

(2) provides data which may be used to determine the target path and to predict its future position.

All or only part of the available radar data-range, elevation angle, azimuth angle, and doppler frequency shift may be used in predicting future position; that is, a radar might track in range, in angle, in doppler, or with any combination. Almost any radar can be considered a tracking radar provided its output information is processed properly. But, in general, it is the method by which angle tracking is accomplished that distinguishes what is normal normally considered a tracking radar from any other radar. It is also necessary to distinguish between a continuous tracking radar and a track-while-scan (TWS) radar.The continuous tracking radar supplies continuous tracking data on a particular target, while the track-while-scan supplies sampled data on one or more targets. In general, the continuous tracking radar and the TWS radar employ different types of equipment.

The antenna beam in the continuous tracking radar is positioned in angle by a servomechanism actuated by an error signal. The various methods for generating the error signal may be classified as sequential lobing, conical scan, and simultaneous lobing or monopulse. The range and doppler frequency shift can also be continuously tracked, if desired, by a servo-control loop actuated by an error signal generated in the radar receiver.

Sequential lobing.

The antenna pattern commonly employed with tracking radars is the symmetrical pencil beam in which the elevation and azimuth beam widths are approximately equal. However, a simple pencil-beam antenna is not suitable for tracking radars unless means are provided for determining the magnitude and direction of the target's angular position with respect to some reference direction, usually the axis of the antenna. The difference between the target position and the reference direction is the angular error. The tracking radar attempts to position the antenna to make the angular error zero. When the angular error is zero, the target is located along the reference direction.

One method of obtaining the direction and the magnitude of the angular error in one coordinate is by alternately switching the antenna beam between two positions . This is called lobe switching, sequential switching, or sequential lobing. Figure 4.1 (a) is a polar representation of the antenna beam (minus the side lobes) in the two switched positions. A plot in rectangular coordinates is shown in figure 4.1 (b), and the error signal obtained from a target not on the switching axis (reference direction) is shown in figure 4.1 (c). The difference in amplitude between the voltages obtained in the two switched positions is a measure of the angular displacement of the target from the switching axis.

The sign of the difference determines the direction the antenna must be moved in order to align the switching axis with the direction of the target. When the voltages in the two switched positions are equal, the target is on axis and, its position may be determined from the axis direction.

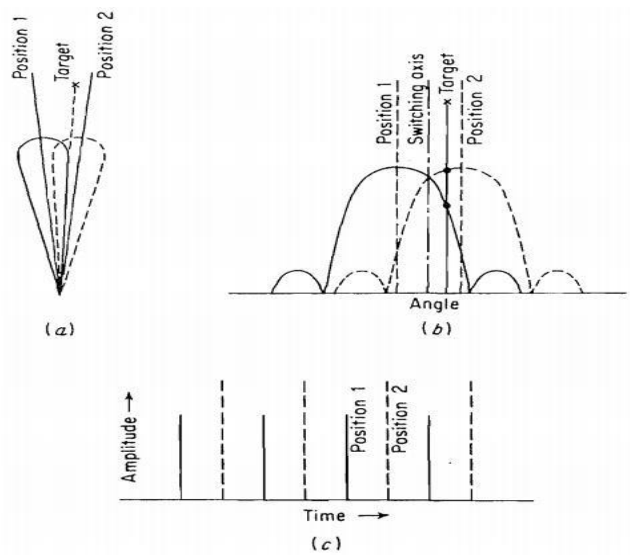


Fig. 4.1 Lobe-switching antenna patterns and error signal (one dimension). (a) Polar representation of switched antenna patterns (b) rectangular representation (c) error signal.

Two additional switching positions are needed to obtain the angular error in the orthogonal coordinate. Thus a two-dimensional sequentially lobing radar might consist of a cluster of four feed horns illuminating a single antenna, arranged so that the right-left, up-down sectors are covered by successive antenna positions. Both transmission and reception are accomplished at each position. A cluster of five feeds might also be employed, with the central feed used for transmission while the outer four feeds are used for receiving. High-power RF switches are not needed since only the receiving beams, and not the transmitting beam, are stepped in this five-feed arrangement.

One of the limitations of a simple unswitched non scanning pencil-beam antenna is that the angle accuracy can be no better than the size of the antenna beam width. An important feature of sequential lobing (as well as the other tracking techniques to be discussed) is that the target-position accuracy can be far better than that given by the antenna beam width. The accuracy depends on how well equality of the signals in the switched positions can be determined. The fundamental limitation to accuracy is system noise caused either by mechanical or electrical fluctuations.

Sequential lobing, or lobe switching, was one of the first tracking-radar techniques to be employed. Early applications were in airborne-interception radar, where it provided directional information for homing on a target, and in ground-based antiaircraft fire-control radars. It is not used as often in modern tracking-radar applications.

Conical scanning method.

The logical extension of the sequential lobing technique is to rotate continuously an offset antenna beam rather than discontinuously step the beam between four discrete positions. This is known as conical scanning (figure 4.2). The angle between the axis of rotation (which is usually, but not always, the axis of the antenna reflector) and the axis of the antenna beam is called the squint angle.

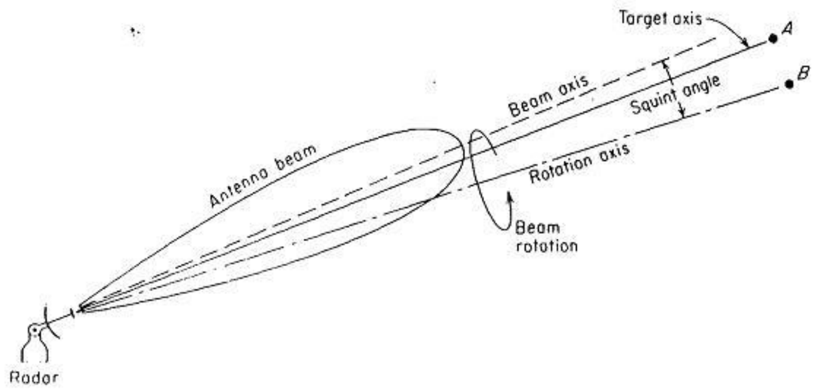


Fig.4.2 Conical-scan tracking

Consider a target at position A. The echo signal will be modulated at a frequency equal to the rotation frequency of the beam. The amplitude of the echo-signal modulation will depend upon the shape of the antenna pattern, the squint angle and the angle between the target line of sight and the rotation axis. The phase of the modulation depends on the angle between the target and the rotation axis. The conical scan modulation is extracted from the echo signal and applied to a servo-control system which continually positions the antenna on the target. When the antenna is on target, as in B of figure 4.2, the line of sight to the target and the rotation axis coincide, and the conical-scan modulation is zero.

A block diagram of the angle-tracking portion of a typical conical-scan tracking radar is shown in figure 4.3. The antenna is mounted so that it can be positioned in both azimuth and elevation by separate motors, which might be either electric- or hydraulic-driven. The antenna beam is offset by tilting either the feed or the reflector with respect to one another.

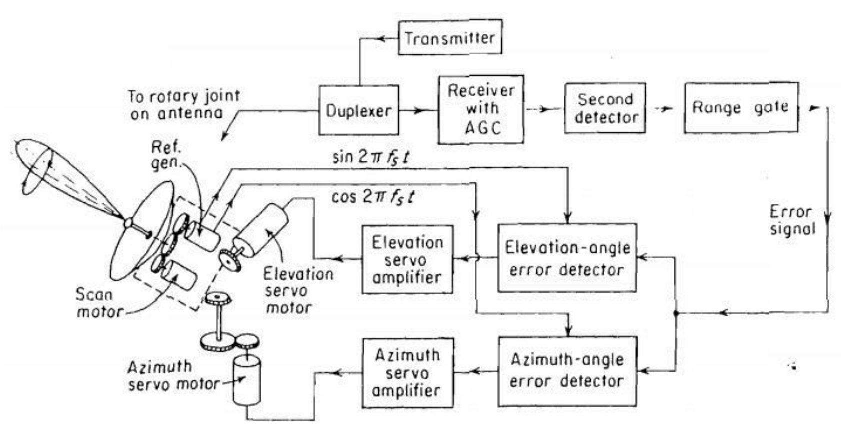


Fig.4.3 Block diagram of conical-scan tracking radar

One of the simplest conical-scan antennas is a parabola with an offset rear feed rotated about the axis of the reflector. If the feed maintains the plane of polarization fixed as it rotates, it is called a nutating feed. A rotating feed causes the polarization to rotate. The latter type of feed requires a rotary joint. The nutating feed requires a flexible joint. If the antenna is small, it may be easier to rotate the dish, which is offset, rather than the feed, thus avoiding the problem of a rotary or flexible RF joint in the feed. A typical conical-scan rotation speed might be 30 r/s. The same motor that provides the conical-scan rotation of the antenna beam also drives a two phase reference generator with two outputs 90° apart in phase. These two outputs serve as a reference

to extract the elevation and azimuth errors. The received echo signal is fed to the receiver from the

antenna via two rotary joints (not shown in the block diagram). One rotary joint permits motion in azimuth, the other, in elevation.

The receiver is a conventional super heterodyne except for features peculiar to the conical scan tracking radar. One feature not found in other radar receivers is a means of extracting the conical-scan modulation, or error signal. This is accomplished after the second detector in the video portion of the receiver. The error signal is compared with the elevation and azimuth reference signals in the angle-error detectors, which are phase-sensitive detectors. A phase-sensitive detector is a nonlinear device in which the input signal (in this case the angle-error signal) is mixed with the reference signal. The input and reference signals are of the same frequency. The output d-c voltage reverses polarity as the phase of the input signal changes through 180°. The magnitude of the d-c output from the angle-error detector is proportional to the error, and the sign (polarity) is an indication of the direction of the error. The angle-error-detector outputs are amplified and drive the antenna elevation and azimuth servo motors.

The angular position of the target may be determined from the elevation and azimuth o the antenna axis. The position can be read out by means of standard angle transducers such as potentiometers, or analog-to-digital-data converters.

Block diagram of the AGC portion of tracking radar receiver.

The echo-signal amplitude at the tracking-radar receiver will not be constant but will vary with time. The three major causes of variation in amplitude are (1) the inverse-fourth-power relationship between the echo signal and range, (2) the conical- scan modulation (angle-error signal), and (3) amplitude fluctuations in the target cross section. The function of the automatic gain control (AGC) is to maintain the d-c level of the receiver output constant and to smooth or eliminate as much of the noise like amplitude fluctuations as possible without disturbing the extraction of the desired error signal at the conical-scan frequency.

One of the purposes of AGC in any receiver is to prevent saturation by large signals. The scanning modulation and the error signal would he lost if the receiver were to saturate. In the conical-scan tracking radar an AGC that maintains the d-c level constant results in an error signal that is a true indication of the angular pointing error. The d-c level of the receiver must be maintained constant if the angular error is to be linearly related to the angle-error signal voltage.

An example of the AGC portion of a tracking-radar receiver is shown in figure 4.4. A portion of the video-amplifier output is passed through a low-pass or smoothing filter and fed back to control the gain of the IF amplifier. The larger the video output, the larger will be the feedback signal and the greater will be the gain reduction. The filter in the AGC loop should pass all frequencies from direct current to just below the conical-scan-modulation frequency. The

loop gain of the AGC filter measured at the conical-scan frequency should be low so that the error signal will not be affected by AGC action.

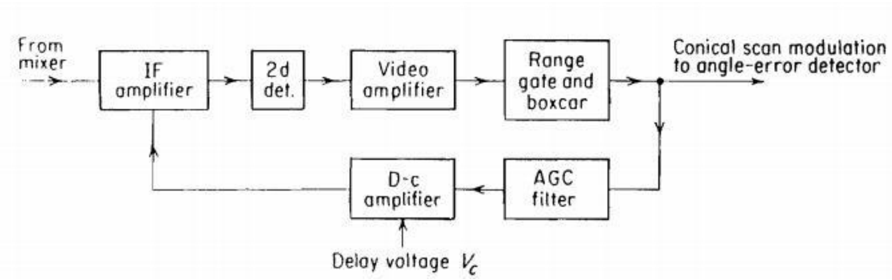


Fig.4.4 Block diagram of the AGC portion of a tracking-radar receiver

The output of the feedback loop will be zero unless the feedback voltage exceeds a prespecified minimum value V_C. In the block diagram the feedback voltage and the voltage V_C are compared in the d-c amplifier. If the feedback voltage exceeds V_C, the AGC is operative, while if it is less, there is no AGC action. The voltage V_C is called the delay voltage. The purpose of the delay voltage is to provide a reference for the constant output signal and permit receiver gain for weak signals. If the delay voltage were zero, any output which might appear from the receiver would be due to the failure of the AGC circuit to regulate completely.

In many applications of AGC the delay voltage is actually zero. This is called undelayed AGC. In such cases the AGC can still perform satisfactorily since the loop gain is usually low for small signals. Thus the AGC will not regulate weak signals. The effect is similar to having a delay voltage, but the performance will not be as good.

Amplitude-comparison monopulse radar for single angular coordinate

The amplitude-comparison monopulse employs two overlapping antenna patterns (figure 4.5 (a)) to obtain the angular error in one coordinate. The two overlapping antenna beams may be generated with a single reflector or with a lens antenna illuminated by two adjacent feeds. (A cluster of four feeds may be used if both elevation- and azimuth-error signals are wanted.) The sum of the two antenna patterns of figure 4.5(a) is shown in figure 4.5(b), and the difference in figure 4.5(c). The sum patterns are used for transmission, while both the sum pattern and the difference pattern are used on reception. The signal received with the difference pattern provides the magnitude of the angle error. The sum signal provides the range measurement and is also used as a reference to extract the sign of the error signal. Signals received from the sum and the difference patterns are amplified separately and combined in a phase-sensitive detector to produce the error-signal characteristic shown in figure 4.5(d).

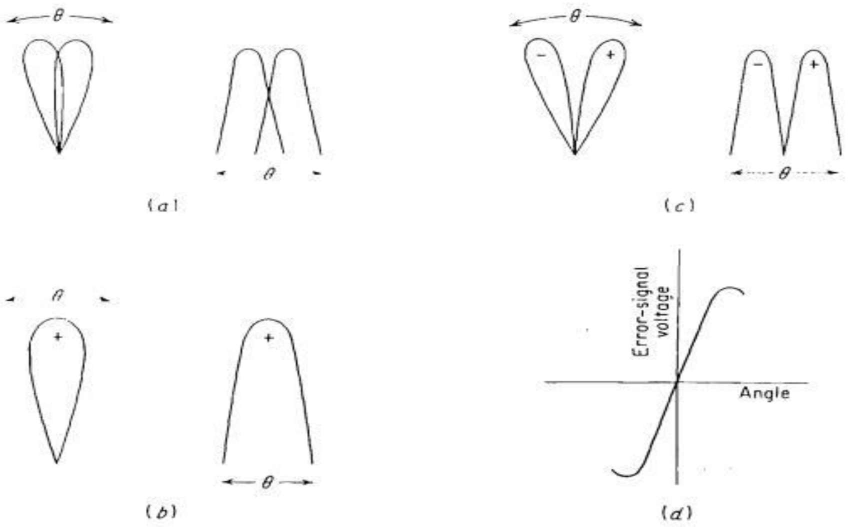


Fig. 4.5. Monopulse antenna patterns and error signal

A block diagram of the amplitude-comparison-monopulse tracking radar for a single angular coordinate is shown in figure 4.6. The two adjacent antenna feeds are connected to the two arms of a hybrid junction such as a magic T, a rat race, or a short-slot coupler. The sum and difference signals appear at the two other arms of the hybrid. On reception, the outputs of the sum arm and the difference arm are each heterodyned to an intermediate frequency and amplified as in any superheterodyne receiver. The transmitter is connected to the sum arm. Range information is also extracted from the sum channel. A duplexer is included in the sum arm for the protection of the receiver. The output of the phase-sensitive detector is an error signal whose magnitude is proportional to the angular error and whose sign is proportional to the direction.

The output of the monopulse radar is used to perform automatic tracking. The angular error signal actuates a servo-control system to position the antenna, and the range output from the sum channel feeds into an automatic-range-tracking unit.

The sign of the difference signal (and the direction of the angular error) is determined by comparing the phase of the difference signal with the phase of the sum signal. If the sum signal in the IF portion of the receiver were $A_s \cos \omega_{IF} t$, the difference signal would be either $A_d \cos \omega_{IF} t$ or $-A_d \cos \omega_{IF} t$ ($A_s > 0$, $A_d > 0$), depending on which side of center is the target. Since $-A_d$

$\cos \omega_{IF} t = -A_d \cos \omega_{IF} (t + \pi)$, the sign of the difference signal may be measured by determining whether the difference signal is in phase with the sum or 180^0 out of phase.

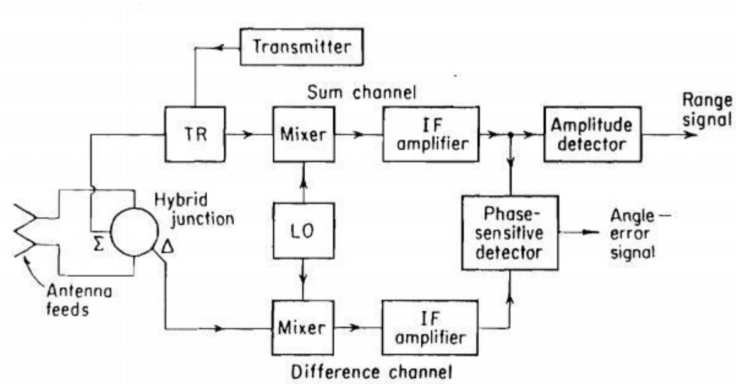


Fig. 4.6 Block diagram of amplitude-comparison monopulse radar

Amplitude-comparison monopulse radar for extracting error signals in both elevation and azimuth.

A block diagram of a monopulse radar with provision for extracting error signals in both elevation and azimuth is shown in figure 4.7. The cluster of four feeds generates four partially overlapping antenna beams. The feeds might be used with a parabolic reflector, Cassegrain antenna, or a lens. All four feeds generate the sum pattern. The difference pattern in one plane is formed by taking the sum of two adjacent feeds and subtracting this from the sum of the other two adjacent feeds. The difference pattern in the orthogonal plane is obtained by adding the differences of the orthogonal adjacent pairs. A total of four hybrid junctions generate the sum channel, the azimuth difference channel, and the elevation difference channel. Three separate mixers and IF amplifiers are shown, one for each channel. All three mixers operate from a single local oscillator in order to maintain the phase relationships between the three channels. Two phase-sensitive detectors extract the angle-error information, one for azimuth, the other for

elevation. Range information is extracted from the output of the sum channel after amplitude detection. Since a phase comparison is made between the output of the sum channel and each of the difference channels, it is important that the phase shifts introduced by each of the channels be almost identical. The phase difference between channels must be maintained to within 25° or
better for reasonably proper performance. The gains of the channels also must not differ by more than specified amounts.

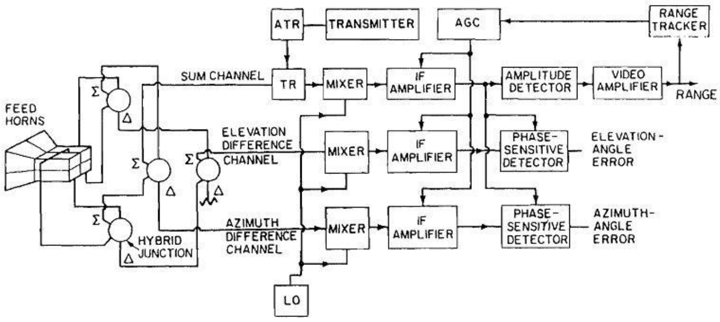


Fig.4.7 Block diagram of two-coordinate (azimuth and elevation) amplitude-comparison monopulse tracking radar.

The purpose in using one- or two-channel monopulse receivers is to ease the problem associated with maintaining identical phase and amplitude balance among the three channels of the conventional receiver. Two-channel monopulse receivers have also been used by combining the sum and the two difference signals in a manner such that they can be again resolved into three components after amplification.

The approximately "ideal" feed-illuminations for a monopulse radar is shown in figure 4.8. This has been approximated in some precision tracking radars by a five-horn feed consisting of one horn generating the sum pattern surrounded by four horns generating the difference patterns.

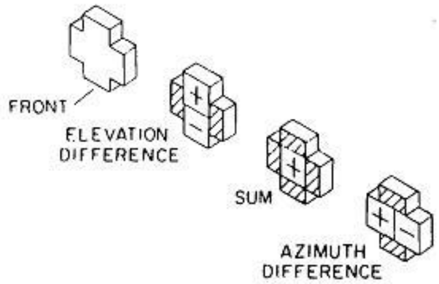


Fig.4.8 approximately ideal feed-aperture illumination for monopulse sum and differenc channels

Phase-comparison monopulse tracking radar technique.

The measurement of angle of arrival by comparison of the phase relationships in the signals from the separated antennas of a radio interferometer has been widely used by the radio astronomers for precise measurements of the positions of radio stars. The interferometer as used by the radio astronomer is a passive instrument, the source of energy being radiated by the target itself. A tracking radar which operates with phase information is similar to an active interferometer and might be called an interferometer radar. It has also been called Simultaneous phase comparison radar, or phase-comparison monopulse.

In figure 4.9 two antennas are shown separated by a distance d. The distance to the target is R and is assumed large compared with the antenna separation d. The line of sight to the target makes an angle θ to the perpendicular bisector of the line joining the two antennas. The distance from antenna 1 to the target is

$R_1 = R + (d \sin \theta) / 2$

and the distance from antenna 2 to the target is

$R_2 = R – (d \sin \theta)/ 2$

The phase difference between the echo signals in the two antennas is approximately

$\Delta\varphi = 2\pi\, d \sin\theta / \lambda$

For small angles where sin θ = 0, the phase difference is a linear function of the angular error and may be used to position the antenna via a servo-control loop.

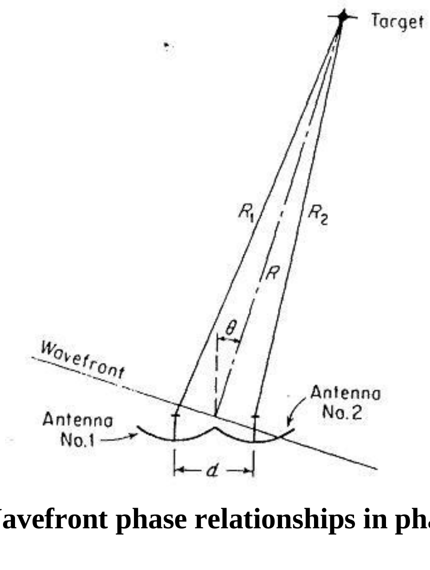


Fig. 4.9 Wavefront phase relationships in phase- comparison monopulse radar

In the early versions of the phase-comparison monopulse radar, the angular error was determined by measuring the phase difference between the outputs of receivers connected to each antenna. The output from one of the antennas was used for transmission and for providing the range information. With such an arrangement it was difficult to obtain the desired aperture illuminations and to maintain a stable boresight. A more satisfactory method of operation is to form the sum and difference patterns in the RF and to process the signals as in a conventional amplitude-comparison monopulse radar.

Split range gates.

The technique for automatically tracking in range is based on the split range gate. Two range gates are generated as shown in figure 4.10. One is the early gate, and the other is the late gate. The echo pulse is shown in figure 4.10 (a), the relative position of the gates at a particular instant in figure 4.10 (b), and the error signal in figure 4.10 (c). The portion of the signal energy contained in the early gate is less than that in the late gate. If the outputs of the two gates are subtracted, an error signal (figure 4.10(c)) will result which may be used to reposition the center of the gates. The magnitude of the error signal is a measure of the difference between the center of the pulse and the center of the gates. The sign of the error signal determines the direction in which the gates must be repositioned by a feedback-control system. When the error signal is zero the range gates are centered on the pulse.

The range gating necessary to perform automatic tracking offers several advantages as by products. It isolates one target excluding targets at other ranges. This permits the boxcar generator to be employed. Also range gating improves the signal-to-noise ratio since it eliminates the noise from the other range intervals. Hence the width of the gate should be sufficiently narrow to minimize extraneous noise. On the other hand, it must not be so narrow that an appreciable fraction of the signal energy is excluded. A reasonable compromise is to make the gate width of the order of the pulse width.

A target of finite length can cause noise in range-tracking circuits in an analogous manner to angle-fluctuation noise (glint) in the angle-tracking circuits. Range-tracking noise depends or the length of the target and its shape. It has been reported that the rms value of the range noise is approximately 0.8 of the target length when tracking is accomplished with a video split-range-gate error detector.

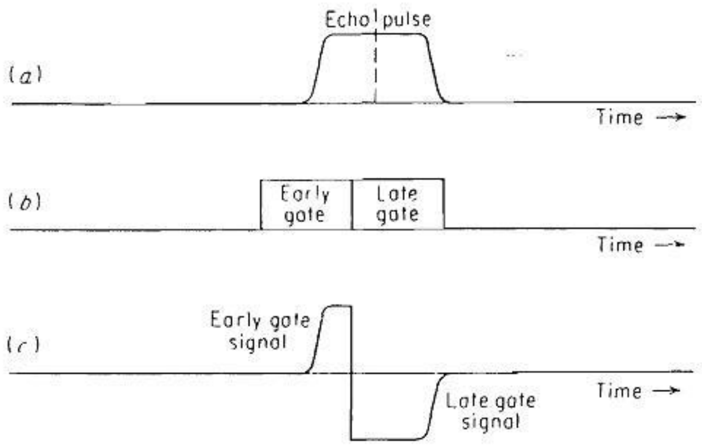


Fig.4.10 Split-range-gate tracking (a) Echo pulse; (h) early-late range gates; (c) difference signal between early and late range gates.

Various methods of acquisition before tracking a target with a radar

A tracking radar must first find and acquire its target before it can operate as a tracker. Therefore it is usually necessary for the radar to scan an angular sector in which the presence of the target is suspected. Most tracking radars employ a narrow pencil-beam antenna. Examples of the common types of scanning patterns employed with pencil-beam antennas are illustrated in figure 4.11

.

In the helical scan, the antenna is continuously rotated in azimuth while it is simultaneously raised or lowered in elevation. It traces a helix in space. Helical scanning was employed for the search mode of the SCR-584 fire-control radar, developed during World War II for the aiming of antiaircraft-gun batteries. The SCR-584 antenna was .rotated at the rate of 6 rpm and covered a 20" elevation angle in 1 min.

The Palmer scan derives its name from the familiar penmanship exercises of grammar school days. It consists of a rapid circular scan (conical scan) about the axis of the antenna, combined with a linear movement of the axis of rotation. When the axis of rotation is held stationary, the Palmer scan reduces to the conical scan. Because of this property, the Palmer scan is sometimes used with conical-scan tracking radars which must operate with a search as well as a track mode since the same mechanisms used to produce conical scanning can also be used for Palmer scanning.Some conical-scan tracking radars increase the squint angle during search in order to reduce the time required to scan a given volume. The conical scan of the SCR-584 was operated during the search mode and was actually a Palmer scan in a helix. In general, conical scan is performed during the search mode of most tracking radars.

The Palmer scan is suited to a search area which is larger in one dimension than another. The spiral scan covers an angular search volume with circular symmetry. Both the spiral scan and the Palmer scan suffer from the disadvantage that all parts of the scan volume do not receive the same energy unless the scanning speed is varied during the scan cycle. As a consequence, the number of hits returned from a target when searching with a constant scanning rate depends upon the position of the target within the search area.

The raster, or TV, scan, unlike the Palmer or the spiral scan, paints the search area in a uniform manner. The raster scan is a simple and convenient means for searching a limited sector, rectangular in shape. Similar to the raster scan is the nodding scan produced by oscillating the antenna beam rapidly in elevation and slowly in azimuth. Although it may be employed to cover a limited sector-as does the raster scan-nodding scan may also be used to obtain hemispherical coverage, that is, elevation angle extending to 90^0 and the azimuth scan angle to 360^0.

The helical scan and the nodding scan can both be used to obtain hemispheric coverage with a pencil beam. The nodding scan is also used with height-finding radars. The Palmer, spiral, and raster scans are employed in fire-control tracking radars to assist in the acquisition of the target when the search sector is of limited extent.

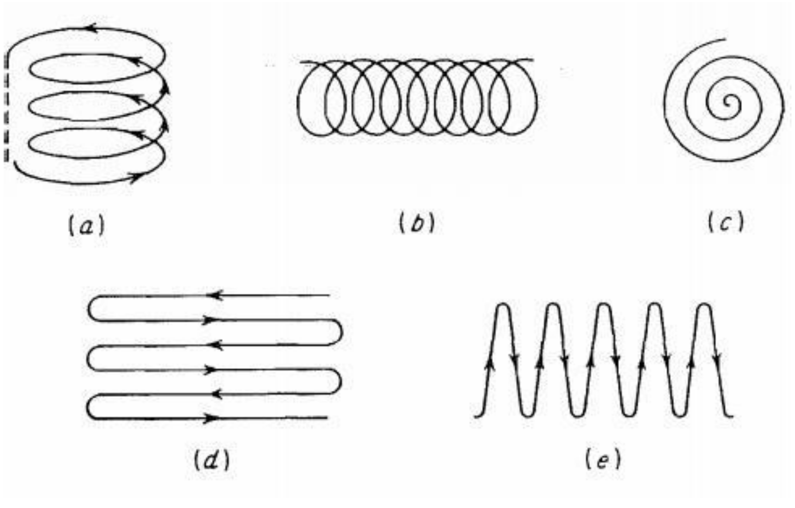


Fig.4.11 Examples of. (a) Trace of helical scanning beam; (b) Palmer scan; (c) spiral scan; (d) raster, or TV, scan; (e) nodding scan.

Main limitations to tracking accuracy of radar are,

1. Amplitude fluctuations**.**

2. Angle fluctuations.

3. Receiver and servo noise

Amplitude fluctuations:

A complex target such as an aircraft or a ship may be considered as a number of independent scattering elements. The echo signal can be represented as the vector addition of the contributions from the individual scatterers. If the target aspect changes with respect to the radar-as might occur because of motion of the target, or turbulence in the case of aircraft targets-the relative phase and amplitude relationships of the contributions from the individual scatterers also change. Consequently, the vector sum, and therefore the amplitude change with changing target aspect.

Amplitude fluctuations of the echo signal are important in the design of the lobe-switching radar and the conical-scan radar but are of little consequence to the monopulse tracker. Both the conical-scan tracker and the lobe-switching tracker require a finite time to obtain a measurement of the angle error.

To reduce the effect of amplitude noise on tracking, the conical-scan frequency should be chosen to correspond to a low value of amplitude noise. If considerable amplitude fluctuation noise were to appear at the conical-scan or lobe-switching frequencies, it could not be readily eliminated with filters or AGC. A typical scan frequency might be of the order of 30 Hz. Higher frequencies might also be used since target amplitude noise generally decreases with increasing frequency.It has been found experimentally that the tracking accuracy of radars operating with pulse repetition frequencies from 1000 to 4000 Hz and a lobing or scan rate one-quarter of the prf are not limited by echo amplitude fluctuations.

Angle fluctuations:

Changes in the target aspect with respect to the radar can cause the apparent center of radar reflections to wander from one point to another. (The apparent center of radar reflection is the direction of the antenna when the error signal is zero.) In general, the apparent center of reflection might not correspond to the target center. .

The random wandering of the apparent radar reflecting center gives rise to noisy or jittered angle tracking. This form of tracking noise is called angle noise, angle scintillations, angle fluctuations, or target glint. The angular fluctuations produced by small targets at long range may be of little consequence in most instances. However, at short range or with relatively large targets (as might be seen by a radar seeker on a homing missile), angular fluctuations may be the chief factor limiting tracking accuracy. Angle fluctuations affect all tracking radars whether conical-scan, sequential lobing, or monopulse.

Receiver and servo noise:

Another limitation on tracking accuracy is the receiver noise power. The accuracy of the angle measurement is inversely proportional to the square root of the signal-to-noise power ratio Since the signal-to-noise ratio is proportional to $1/R^4$ (from the radar equation), the angular error due to receiver noise is proportional to the square of the target distance. Servo noise is the hunting action of the tracking servomechanism which results from backlash and compliance in the gears, shafts, and structures of the mount. The magnitude of servo noise is essentially independent of the target echo and will therefore be independent of range.

Low angle tracking.

A radar that tracks a target at a low elevation angle, near the surface of the earth, can receive two echo signals from the target, figure 5.12. One signal is reflected directly from the target, and the other arrives via the earth's surface. The direct and the surface-reflected signals combine at the radar to yield angle measurement that differs from the true measurement that would have been made with a single target in the absence of surface reflections. The result is an error in the measurement of elevation. The surface-reflected signal may be thought of as originating from the image of the target mirrored by the earth's surface. Thus, the effect on tracking is similar to the two-target model used to describe glint. The surface-reflected signal is sometimes called a multipath signal.

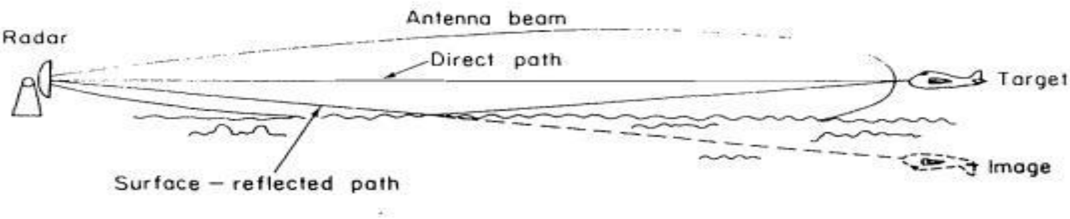


Fig 4.12 Low-angle tracking

The surface-reflected signal travels a longer path than the direct signal so that it may be possible in some cases to separate the two in time (range). Tracking on the direct signal avoids the angle errors introduced by the multipath. The range-resolution required to separate the direct from the ground-reflected signal is

$$\Delta R = \frac{2h_a h_t}{R}$$

Where,

h_a = radar antenna

height, h_t = target height,

R = range to the target.

For a radar height of 30 m, a target height of 100 m and a range of 10 km, the range-resolution must be 0.6 m, corresponding to a pulse width of 4 ns. This is a much shorter pulse than is commonly employed in radar. Although the required range-resolutions for a ground-based radar are achievable in principle, it is usually not applicable in practice. The use of frequency diversity can also reduce the multipath tracking error.

UNIT-V

DETECTION OF RADAR SIGNALS IN NOISE

Matched-Filter Receiver:

A network whose frequency-response function maximizes the output peak-signal-to-mean-noise (power) ratio is called a matched filter. This criterion, or its equivalent, is used for the design of almost all radar receivers.

The frequency-response function, denoted H (f), expresses the relative amplitude and phase of the output of a network with respect to the input when the input is a pure sinusoid. The magnitude |H (f)| of the frequency-response function is the receiver amplitude pass band characteristic. If the bandwidth of the receiver pass band is wide compared with that occupied by the signal energy, extraneous noise is introduced by the excess bandwidth which lowers the output signal-to-noise ratio. On the other hand, if the receiver bandwidth is narrower than the bandwidth occupied by the signal, the noise energy is reduced along with a considerable part of the signal energy. The net result is again a lowered signal-to-noise ratio. Thus there is an optimum bandwidth at which the signal-to-noise ratio is a maximum. This is well known to the radar receiver designer. The rule of thumb quoted in pulse radar practice is that the receiver bandwidth B should be approximately equal to the reciprocal of the pulse width τ. This is a reasonable approximation for pulse radars with conventional superheterodyne receivers. It is not generally valid for other waveforms, however, and is mentioned to illustrate in a qualitative manner the effect of the receiver characteristic on signal-to-noise ratio. The exact specification of the optimum receiver characteristic involves the frequency-response function and the shape of the received waveform.

The receiver frequency-response function, is assumed to apply from the antenna terminals to the output of the IF amplifier. (The second detector and video portion of the well-designed radar superheterodyne receiver will have negligible effect on the output signal-to-noise ratio if the receiver is designed as a matched filter.) Narrow banding is most conveniently accomplished in the IF. The bandwidths of the RF and mixer stages of the normal superheterodyne receiver are usually large compared with the IF bandwidth. Therefore the frequency-response function of the portion of the receiver included between the antenna terminals to the output of the IF amplifier is taken to be that of the IF amplifier alone. Thus we need only obtain the frequency-response function that maximizes the signal-to-noise ratio at the output of the IF. The IF amplifier may be considered as a filter with gain. The response of this filter as a function of frequency is the property of interest. For a received waveform s(t) with a given ratio of signal energy E to noise energy N_o (or noise power per hertz of bandwidth), North showed that the frequency-response function of the linear, time-invariant filter which maximizes the output peak to mean signal to noise ratio for a fixed input signal to noise ratio is

$$H(f) = G_a S^*(f) \exp(-j2\pi f t_1)$$

where $S(f) = \int_{-\infty}^{\infty} s(t) \exp(-j2\pi f t)\, dt$ = voltage spectrum (Fourier transform) of input signal

$S^*(f)$ = complex conjugate of $S(f)$

t_1 = fixed value of time at which signal is observed to be maximum

G_a = constant equal to maximum filter gain (generally taken to be unity)

The noise that accompanies the signal is assumed to be stationary and to have a uniform spectrum (white noise). It need not be gaussian. The filter whose frequency-response function is given by Eq. above has been called the North filter, the conjugate filter, or more usually the matched filter. It has also been called the Fourier transform criterion. It should not be confused with the circuit-theory concept of impedance matching, which maximizes the power transfer rather than the signal-to-noise ratio.

The frequency-response function of the matched filter is the conjugate of the spectrum of the received waveform except for the phase shift exp (- j2Πft_1). This phase shift varies uniformly with frequency. Its effect is to cause a constant time delay. A time delay is necessary in the specification of the filter for reasons of physical realizability since there can be no output from the filter until the signal is applied. The frequency spectrum of the received signal may be written as an amplitude spectrum |S(f)| (and a phase spectrum exp [-jφ_s (f)]. The matched- filter frequency-response function may similarly be written in terms of its amplitude and phase spectra |H (f)| and exp [-jφ_m (f)]. Ignoring the constant G_a, Eq. above for the matched filter may then be written as

$$|H(f)| \exp[-j\phi_m(f)] = |S(f)| \exp\{j[\phi_s(f) - 2\pi f t_1]\}$$

or

$$|H(f)| = |S(f)|$$

and

$$\phi_m(f) = -\phi_s(f) + 2\pi f t_1$$

Thus the amplitude spectrum of the matched filter is the same as the amplitude spectrum of the signal, but the phase spectrum of the matched filter is the negative of the phase spectrum of the signal plus a phase shift proportional to frequency.

The matched filter may also be specified by its impulse response h (t), which is the inverse Fourier transform of the frequency-response function.

$$h(t) = \int_{-\infty}^{\infty} H(f) \exp(j2\pi ft)\, df$$

Physically, the impulse response is the output of the filter as a function of time when the input is an impulse (delta function).

$$h(t) = G_a \int_{-\infty}^{\infty} S^*(f) \exp[-j2\pi f(t_1 - t)]\, df$$

Since $S^*(f) = S(-f)$, we have

$$h(t) = G_a \int_{-\infty}^{\infty} S(f) \exp[j2\pi f(t_1 - t)]\, df = G_a s(t_1 - t)$$

A rather interesting result is that the impulse response of the matched filter is the image of the received waveform; that is, it is the same as the received signal run backward in time starting from the fixed time t_1. Figure 5.1 shows a received waveform s (t) and the impulse response h (t) of its matched filter. The impulse response of the filter, if it is to be realizable, is not defined for t< 0. (One cannot have any response before the impulse is applied.) Therefore we must always have $t < t_1$. This is equivalent to the condition placed on the transfer function H (f) that there be a phase shift exp ($-j2\Pi ft_1$). However, for the sake of convenience, the impulse response of the matched filter is sometimes written simply as s (- t).

Derivation of the matched-filter characteristic: The frequency-response function of the matched filter has been derived by a number of authors using either the calculus of variations or the Schwartz inequality. We shall derive the matched-filter frequency-response function using the Schwartz inequality.

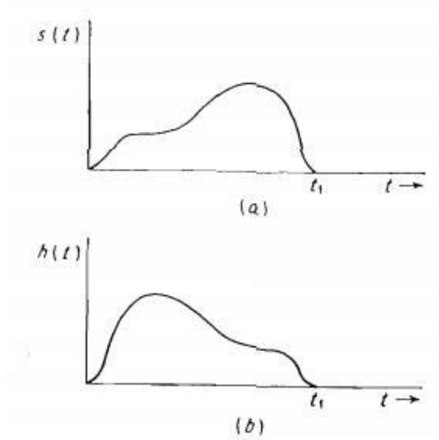


Fig.5.1 (a) Received waveform s(t); (b) impulse response h(t) of the matched filter.

We wish to show that the frequency-response function of the linear, time-invariant filter which maximizes the output peak-signal-to-mean-noise ratio is

$$H(f) = G_a S^*(f) \exp(-j2\pi f t_1)$$

when the input noise is stationary and white (uniform spectral density). The ratio we wish to maximize is

$$R_f = \frac{|s_o(t)|^2_{max}}{N}$$

where $|s_o(t)|_{max}$ = maximum value of output signal voltage and N = mean noise power at receiver output. The ratio R_f is not quite the same as the signal-to-noise ratio which has been considered in the radar equation. The output voltage of a filter with frequency-response function H(f) is

$$|s_o(t)| = \left| \int_{-\infty}^{\infty} S(f)H(f) \exp(j2\pi ft)\, df \right|$$

where S(f) is the Fourier transform of the input (received) signal. The mean output noise power is

$$N = \frac{N_o}{2} \int_{-\infty}^{\infty} |H(f)|^2\, df$$

where N_o is the input noise power per unit bandwidth. The factor appears before the integral because the limits extend from - ∞ to +∞, whereas N_0 is defined as the noise power per cycle of bandwidth over positive values only. Assuming that the maximum value of $|s_o(t)|^2$ occurs at time $t = t_1$, the ratio R_f becomes

Schwartz's inequality states that if P and Q are two complex functions, then

$$\int P^*P\,dx \int Q^*Q\,dx \geq \left| \int P^*Q\,dx \right|^2$$

The equality sign applies when P = kQ, where k is a constant.

Letting

$$P^* = S(f)\exp(j2\pi f t_1) \quad \text{and} \quad Q = H(f)$$

and recalling that

$$\int P^*P\,dx = \int |P|^2\,dx$$

we get, on applying the Schwartz inequality to the numerator of Eq. earlier, we get

$$R_f \leq \frac{\int_{-\infty}^{\infty} |H(f)|^2\,df \int_{-\infty}^{\infty} |S(f)|^2\,df}{\frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2\,df} = \frac{\int_{-\infty}^{\infty} |S(f)|^2\,df}{\frac{N_0}{2}}$$

From Parseval's theorem,

$$\int_{-\infty}^{\infty} |S(f)|^2\,df = \int_{-\infty}^{\infty} s^2(t)\,dt = \text{signal energy} = E$$

Therefore we have

$$R_f \leq \frac{2E}{N_0}$$

The frequency-response function which maximizes the peak-signal-to-mean-noise ratio R_f may be obtained by noting that the equality sign in Eq. applies when P = kQ, or

$$H(f) = G_a S^*(f)\exp(-j2\pi f t_1)$$

where the constant k has been set equal to $1/G_a$.

Relation between the matched filter characteristics and correlation function.

The matched filter and the correlation function. The output of the matched filter is not a replica of the input signal. However, from the point of view of detecting signals in noise, preserving the shape of the signal is of no importance. If it is necessary to preserve the shape of the input pulse rather than maximize the output signal-to-noise ratio, some other criterion must be employed.

The output of the matched filter may be shown to be proportional to the input signal cross-correlated with a replica of the transmitted signal, except for the time delay t_1. The cross-correlation function R (t) of two signals y(λ) and s(λ), each of finite duration, is defined as

$$R(t) = \int^{\infty} y(\lambda)s(\lambda - t)\, d\lambda$$

$$y_0(t) = \int_{-\infty}^{\infty} y_{in}(\lambda)h(t - \lambda)\, d\lambda$$

If the filter is a matched filter, then h(λ) = s(t_1 - λ) and Eq. above becomes

$$y_0(t) = \int_{-\infty}^{\infty} y_{in}(\lambda)s(t_1 - t + \lambda)\, d\lambda = R(t - t_1)$$

Thus the matched filter forms the cross correlation between the received signal corrupted by noise and a replica of the transmitted signal. The replica of the transmitted signal is "built in" to the matched filter via the frequency-response function. If the input signal y_{in} (t) were the same as the signal s(t) for which the matched filter was designed (that is, the noise is assumed negligible), the output would be the autocorrelation function. The autocorrelation function of a rectangular pulse of width τ is a triangle whose base is of width 2τ.

Efficiency of non-matched filters.

Efficiency of non matched filters: In practice the matched filter cannot always be obtained exactly. It is appropriate, therefore, to examine the efficiency of non matched filters compared with the ideal matched filter. The measure of efficiency is taken as the peak signal-to-noise ratio from the non matched filter divided by the peak signal-to-noise ratio ($2E/N_0$) from the matched filter. Figure 5.2 plots the efficiency for a single-tuned (RLC) resonant filter and a rectangular-shaped filter of half-power bandwidth B_T when the input is a rectangular pulse of width τ. The maximum efficiency of the single-tuned filter occurs for $B_T \approx 0.4$. The corresponding loss in signal-to-noise ratio is 0.88 dB as compared with a matched filter. Table 5.1 lists the values of B_T which maximize the signal-to-noise ratio (SNR) for various combinations of filters and pulse shapes. It can be seen that the loss in SNR incurred by use of these non-matched filters is small.

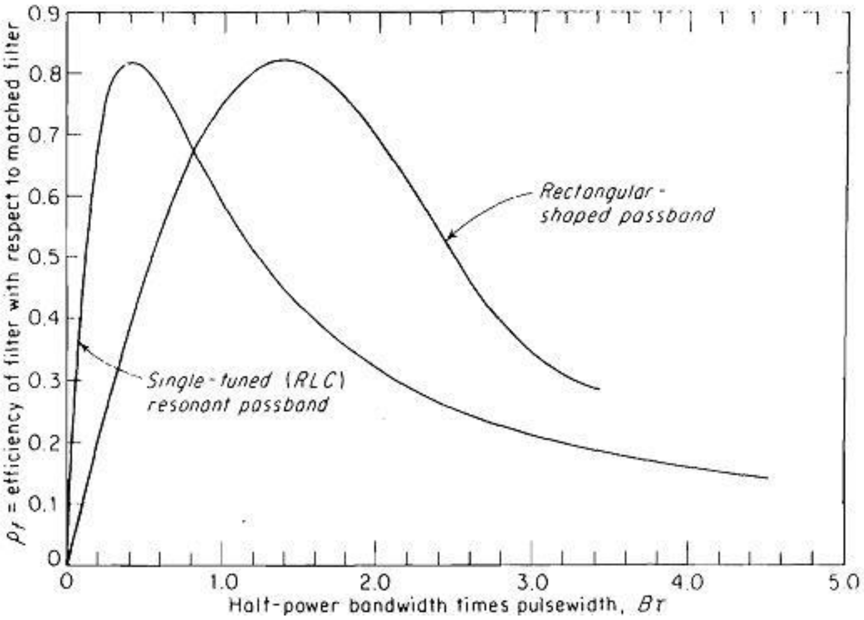


Fig.5.2 Efficiency, relative to a matched filter, of a single-tuned resonant filter and a rectangular shaped filter, when the input signal is a rectangular pulse of width τ. B = filter bandwidth.

Efficiency of nonmatched filters compared with the matched filter is shown in table 5.1

Table 5.1 Efficiency of nonmatched filters compared with the matched filter

Input signal	Filter	Optimum $B\tau$	Loss in SNR compared with matched filter, dB
Rectangular pulse	Rectangular	1.37	0.85
Rectangular pulse	Gaussian	0.72	0.49
Gaussian pulse	Rectangular	0.72	0.49
Gaussian pulse	Gaussian	0.44	0 (matched)
Rectangular pulse	One-stage, single-tuned circuit	0.4	0.88
Rectangular pulse	2 cascaded single-tuned stages	0.613	0.56
Rectangular pulse	5 cascaded single-tuned stages	0.672	0.5

Matched filter with nonwhite noise: In the derivation of the matched-filter characteristic, the spectrum of the noise accompanying the signal was assumed to be white; that is, it was independent of frequency. If this assumption were not true, the filter which maximizes the output signal-to-noise ratio would not be the same as the matched filter. It has been shown that if the input power spectrum of the interfering noise is given by $[N_i\ (f)]^2$, the frequency-response function of the filter which maximizes the output signal-to-noise ratio is

$$H(f) = \frac{G_a S^*(f) \exp(-j2\pi f t_1)}{[N_i(f)]^2}$$

When the noise is nonwhite, the filter which maximizes the output signal-to-noise ratio is called the NWN (nonwhite noise) matched filter. For white noise $[N_i(f)]^2$ = constant and the NWN matched-filler frequency-response function of Eq. above reduces to that of Eq. discussed earlier in white noise. Equation above can be written as

$$H(f) = \frac{1}{N_i(f)} \times G_a \left(\frac{S(f)}{N_i(f)} \right)^* \exp(-j2\pi f t_1)$$

This indicates that the NWN matched filter can be considered as the cascade of two filters. The first filter, with frequency-response function $l/N_i\ (f)$, acts to make the noise spectrum uniform, or white. It is sometimes called the whitening filter. The second is the matched filter when the input is white noise and a signal whose spectrum is $S(f)/N_i(f)$.

Correlation Detection:

$$y_0(t) = \int_{-\infty}^{\infty} y_{in}(\lambda) s(t_1 - t + \lambda)\, d\lambda = R(t - t_1)$$

Equation above describes the output of the matched filter as the cross correlation between the input signal and a delayed replica of the transmitted signal. This implies that the matched-filter receiver can be replaced by a cross-correlation receiver that performs the same mathematical operation as shown in figure 5.3. The input signal y (t) is multiplied by a delayed replica of the transmitted signal $s(t - T_r)$, and the product is passed through a low-pass filter to perform the integration. The cross-correlation receiver of Fig.5 tests for the presence of a target at only a single time delay T_r. Targets at other time delays, or ranges, might be found by varying T_r.

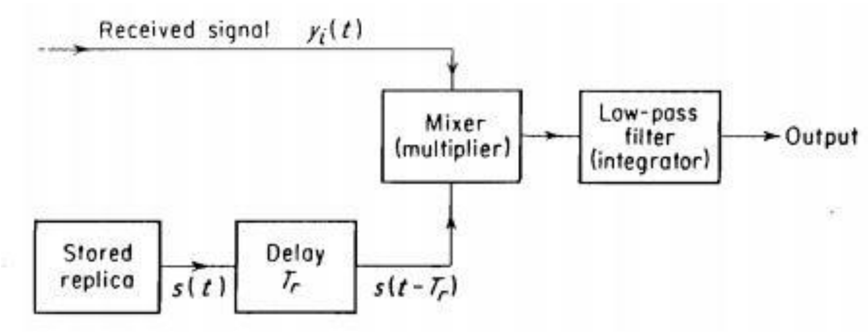


Fig.5.3 Block diagram of a cross-correlation receiver.

However, this requires a longer search time. The search time can be reduced by adding parallel channels, each containing a delay line corresponding to a particular value of T_r, as well as a multiplier and low-pass filter. In some applications it may be possible to record the signal on some storage medium, and at a higher playback speed perform the search sequentially with different values of T_r. That is, the playback speed is increased in proportion to the number of time-delay intervals T_r that are to be tested.Since the cross-correlation receiver and the matched-filter receiver are equivalent mathematically, the choice as to which one to use in a particular radar application is determined by which is more practical to implement. The matched-filter receiver, or an approximation, has been generally preferred in the vast majority of applications.

Constant False Alarm Rate Receiver

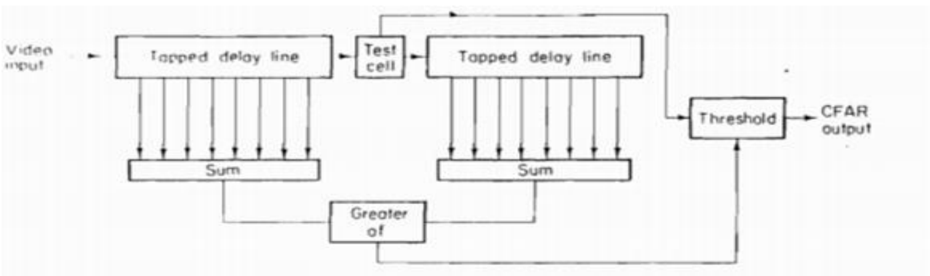


Fig.5.4 Cell averaging CFAR

The false-alarrn rate is quite sensitive to the threshold level. For example, a 1 dB change in the threshold can result in three orders of magnitude change in the false alarm probability. It does not take much of a drift in the receiver gain, a change in receiver noise, or the presence of external noise or clutter echoes to inundate the radar display with extraneous responses.

If changes in the false-alarm rate are gradual, an operator viewing a display can compensate with a manual gain adjustment. It has been said that the maximum increase in noise level that can be tolerated with a manual system using displays and operators is from 5 to 10 dB. But with an automatic detection and tracking (ADT) system, the tolerable increase is less than 1 dB. Excessive false alarms in an ADT system cause the computer to overload as it attempts to associate false alarms with established tracks or to generate new, but false, tracks. Manual control is too slow and imprecise for automatic systems. Some automatic, instantaneous means is required to maintain a constant false-alarm rate. Devices that accomplish this purpose are called CFAR.

A CFAR may be obtained by observing the noise or clutter background in the vicinity of the target and adjusting the threshold in accordance with the measured background. Figure 5.4 illustrates the cell-averaging CFAR which utilizes a tapped delay-line to sample the range cells to either side of the range cell of interest, or test cell. The output of the test cell is the radar output. The spacing between the taps is equal to the range resolution. The outputs from the delay line taps are summed. This sum, when multiplied by the appropriate constant, determines the threshold level for achieving the desired probability of false alarm. Thus the threshold varies continuously according to the noise or the clutter environment found within a range interval surrounding the range cell under observation. This form of CFAR has sometimes been called Adaptive Video Threshold, or AVT.

UNIT-VI
RADAR RECEIVERS

The Radar Receiver:

The function of the radar receiver is to detect desired echo signals in the presence of noise, interference, or clutter. It must separate wanted from unwanted signals, and amplify the wanted signals to a level where target information can be displayed to an operator or used in an automatic data processor. The design of the radar receiver will depend not only on the type of waveform to be detected, but on the nature of the noise, interference, and clutter echoes with which the desired echo signals must compete.

Noise can enter the receiver via the antenna terminals along with the desired signals, or it might be generated within the receiver itself. At the microwave frequencies usually used for radar, the external noise which enters via the antenna is generally quite low so that the receiver sensitivity is usually set by the internal noise generated within the receiver. The measure of receiver internal noise is the noise-figure. Good receiver design is based on maximizing the output signal-to-noise ratio. To maximize the output signal-to-noise ratio, the receiver must be designed as a matched filter, or its equivalent. The matched filter specifies the frequency response function of the IF part of the radar receiver. Obviously, the receiver should be designed to generate as little internal noise as possible, especially in the input stages where the desired signals are weakest. Although special attention must be paid to minimize the noise of the input stages, the lowest noise receivers are not always desired in many radar applications if other important receiver properties must be sacrificed.

Receiver design also must be concerned with achieving sufficient gain, phase, and amplitude stability, dynamic range, tuning, ruggedness, and simplicity. Protection must be provided against overload or saturation, and burnout from nearby interfering transmitters. Timing and reference signals are needed to properly extract target information. Specific applications such as MTI radar, tracking radar, or radars designed to minimize clutter place special demands on the receiver. Receivers that must operate with a transmitter whose frequency can drift need some means of automatic frequency control (AFC). Radars that encounter hostile counter-measures need receivers that can minimize the effects of such interference. Thus there can be many demands placed upon the receiver designer in meeting the requirements of modern high-quality radar systems. The receiver engineer has responded well to the challenge, and there exists a highly refined state of technology available for radar applications.

Radar receiver design and implementation may not always be an easy task; but in tribute to the receiver designer, it has seldom been an obstacle preventing the radar systems engineer from eventually accomplishing the desired objectives.Although the super regenerative, crystal video and tuned radio frequency (TRF) receivers have been employed in radar systems, the superheterodyne has seen almost exclusive application because of its good sensitivity, high gain, selectivity, and reliability. No other receiver type has been competitive to the superheterodyne.

Noise Figure:

The noise figure of a receiver was described as a measure of the noise produced by a practical receiver as compared with the noise of an ideal receiver. The noise figure F_n of a linear network may be defined as

$$F_n = \frac{S_{in}/N_{in}}{S_{out}/N_{out}} = \frac{N_{out}}{kT_0 B_n G}$$

where S_{in} = available input signal power
N_{in} = available input noise power (equal to $kT_0 B_n$)
S_{out} = available output signal power
N_{out} = available output noise power

"Available power" refers to the power which would be delivered to a matched load. The available gain G is equal to S_{out}/S_{in}, k = Boltzmann's constant = 1.38×10^{-23} J/deg, T_o = standard temperature of 290 K and B, is the noise bandwidth. The product $kT_o \approx 4 \times 10^{-21}$ W/Hz. The purpose for defining a standard temperature is to refer any measurements to a common basis of comparison. Equation above permits two different but equivalent interpretations of noise figure. It may be considered as the degradation of the signal-to-noise ratio caused by the network (receiver), or it may be interpreted as the ratio of the actual available output noise power to the noise power which would be available if the network merely amplified the thermal noise. The noise figure may also be written

$$F_n = \frac{kT_0 B_n G + \Delta N}{kT_0 B_n G} = 1 + \frac{\Delta N}{kT_0 B_n G}$$

where N is the additional noise introduced by the network itself. The noise figure is commonly expressed in decibels, that is, $10\log F_n$. The term noise factor is also used at times instead of noise figure. The two terms are synonymous.

The definition of noise figure assumes the input and output of the network are matched. In some devices less noise is obtained under mismatched, rather than matched, conditions. Inspite of definitions, such networks would be operated so as to achieve the maximum output signal-to-noise ratio.

Noise temperature: The noise introduced by a network may also be expressed as an effective noise temperature, T_e, defined as that (fictional) temperature at the input of the network which would account for the noise N at the output. Therefore $N = k T_e B_n G$ and

$$F_n = 1 + \frac{T_e}{T_0}$$

$$T_e = (F_n - 1)T_0$$

The system noise temperature T_s is defined as the effective noise temperature of the receiver system including the effects of antenna temperature T_a. (It is also sometimes called the system operating noise temperature) If the receiver effective noise temperature is T_e, then

$$T_s = T_a + T_e = T_0 F_s$$

where F_s is the system noise-figure including the effect of antenna temperature. The effective noise temperature of a receiver consisting of a number of networks in cascade is

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} + \cdots$$

where T_i and G_i , are the effective noise temperature and gain of the ith network. The effective noise temperature and the noise figure both describe the same characteristic of a network. In general, the effective noise temperature has been preferred for describing low-noise devices, and the noise figure is preferred for conventional receiver.

Noise figure of networks in cascade: Consider two networks in cascade, each with the same noise bandwidth B_n but with different noise figures and available gain (figure 6.1). Let F_1, G_1 be the noise figure and available gain, respectively, of the first network, and F_2, G_2 be similar parameters for the second network. The problem is to find F_o, the overall noise-figure of the two circuits in cascade. From the definition of noise figure the output noise N_o of the two circuits in cascade is

$$N_o = F_o G_1 G_2 k T_0 B_n = \text{noise from network 1 at output of network 2} + \text{noise } \Delta N_2 \text{ introduced by network 2}$$

$$N_o = k T_0 B_n F_1 G_1 G_2 + \Delta N_2 = k T_0 B_n F_1 G_1 G_2 + (F_2 - 1) k T_0 B_n G_2$$

or

$$F_o = F_1 + \frac{F_2 - 1}{G_1} ,$$

Fig.6.1 Two networks in cascade

The contribution of the second network to the overall noise-figure may be made negligible if the gain of the first network is large. This is of importance in the design of multistage receivers. It is not sufficient that only the first stage of a low-noise receiver have a small noise figure. The succeeding stage must also have a small noise figure, or else the gain of the first stage must be high enough to swamp the noise of the succeeding stage if the first network is not an amplifier but is a network with loss (as in a crystal mixer), the gain G_I should be interpreted as a number less than unity.

The noise figure of N networks in cascade may be shown to be

$$F_o = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \cdots + \frac{F_N - 1}{G_1 G_2 \cdots G_{N-1}}$$

Displays:

The purpose of the display is to visually present in a form suitable for operator interpretation and action the information contained in the radar echo signal. When the display is connected directly to the video output of the receiver, the information displayed is called raw video. This is the "traditional" type of radar presentation. When the receiver video output is first processed by an automatic detector or automatic detection and tracking processor (ADT), the output displayed is sometimes called synthetic video.

The cathode-ray tube (CRT) has been almost universally used as the radar display. There are two basic cathode-ray tube displays. One is the dejection-modulated CRT, such as the A-scope, in which a target is indicated by the deflection of the electron beam.

The other is the intensity modulated CRT, such as the PPI, in which a target is indicated by intensifying the electron beam and presenting a luminous spot on the face of the CRT. In general, deflection-modulated displays have the advantage of simpler circuits than those of intensity- modulated displays, and targets may be more readily discerned in the presence of noise or interference. On the other hand, intensity-modulated displays have the advantage of presenting data in a convenient and easily interpreted form. The deflection of the beam or the appearance of an intensity-modulated spot on a radar display caused by the presence of a target is commonly referred to as a blip. .

Types of display presentations:

The various types of CRT displays which might be used for surveillance and tracking radars are defined as follows:

A-scope: A deflection-modulated display in which the vertical deflection is proportional to target echo strength and the horizontal coordinate is proportional to range.

B-scope: An intensity-modulated rectangular display with azimuth angle indicated by the horizontal coordinate and range by the vertical coordinate.

C-scope: An intensity-modulated rectangular display with azimuth angle indicated by the horizontal coordinate and elevation angle by the vertical coordinate.

D-scope: A C-scope in which the blips extend vertically to give a rough estimate of distance.

E-scope: An intensity-modulated rectangular display with distance indicated by the horizontal coordinate and elevation angle by the vertical coordinate. Similar to the RHI in which target height or altitude is the vertical coordinate.

F-Scope: A rectangular display in which a target appears as a centralized blip when the radar antenna is aimed at it. Horizontal and vertical aiming errors are respectively indicated by the horizontal and vertical displacement of the blip.

G-Scope: A rectangular display in which a target appears as a laterally centralized blip when the radar antenna is aimed at it in azimuth, and wings appear to grow on the pip as the distance to the target is diminished; horizontal and vertical aiming errors are respectively indicated by horizontal and vertical displacement of the blip.

H-scope: A B-scope modified to include indication of angle of elevation. The target appears as two closely spaced blips which approximate a short bright line, the slope of which is in proportion to the sine of the angle of target elevation.

I-scope: A display in which a target appears as a complete circle when the radar antenna is pointed at it and in which the radius of the circle is proportional to target distance; incorrect aiming of the antenna changes the circle to a segment whose arc length is inversely proportional to the magnitude of the pointing error, and the position of the segment indicates the reciprocal of the pointing direction of the antenna.

J-scope: A modified A-scope in which the time base is a circle and targets appear as radial deflections from the time base.

K-scope: A modified A-scope in which a target appears as a pair of vertical deflections. When the radar antenna is correctly pointed at the target, the two deflections are of equal height, and when not so pointed, the difference in deflection amplitude is an indication of the direction and magnitude of the pointing error.

L-scope: A display in which a target appears as two horizontal blips, one extending to the right from a central vertical time base and the other to the left; both blips are of equal amplitude when the radar is pointed directly at the target, any inequality representing relative pointing error, and distance upward along the baseline representing target distance.

M-scope: A type of A-scope in which the target distance is determined by moving an adjustable pedestal signal along the baseline until it coincides with the horizontal position of the target signal deflections; the control which moves the pedestal is calibrated in distance.

N-scope: A K-scope having an adjustable pedestal signal, as in the M-scope, for the measurement of distance.

O-scope: An A-scope modified by the inclusion of an adjustable notch for measuring distance.

PPI or Plan Position Indicator (also called P-scope): An intensity-modulated circular display on which echo signals produced from reflecting objects are shown in plan position with range and azimuth angle displayed in polar (rho-theta) coordinates, forming a map-like display. An offset, or off center, PPI has the zero position of the time base at a position other than at the center of the display to provide the equivalent of a larger display for a selected portion of the service area. A delayed PPI is one in which the initiation of the time base is delayed.

R-scope: An A-scope with a segment of the time base expanded near the blip for greater accuracy in distance measurement.

RHI or Range-Height Indicator: An intensity modulated display with height (altitude) as the vertical axis and range as the horizontal axis.

Duplexers: Branch type.

The branch-type duplexer, diagrammed in figure 6.2 was one of the earliest duplexer configurations employed. It consists of a TR (transmit-receive) switch and an ATR (anti-transmit receive) switch, both of which are gas-discharge tubes. When the transmitter is turned on, the TR and the ATR tubes ionize; that is, they break down, or fire. The TR in the fired condition acts as a short circuit to prevent transmitter power from entering the receiver. Since the TR is located a quarter wave length from the main transmission line, it appears as a short circuit at the receiver but as an open circuit at the transmission line so that it does not impede the flow of transmitter power.

Since the ATR is displaced a quarter wave length from the main transmission line, the short circuit it produces during the fired condition appears as an open circuit on the transmission line and thus has no effect on transmission. During reception, the transmitter is off and neither the TR nor the ATR is fired. The open circuit of the ATR, being a quarter wave from the transmission line, appears as a short circuit across the line. Since this short circuit is located a quarter wave from the receiver branch-line, the transmitter is effectively disconnected from the line and the echo signal power is directed to the receiver. The diagram of figure 8.2 is a parallel configuration. Series or series-parallel configurations are possible. The branch-type duplexer is of limited bandwidth and power-handling capability, and has generally been replaced by the balanced duplexer and other protection devices. It is used, in spite of these limitations, in some low-cost radars.

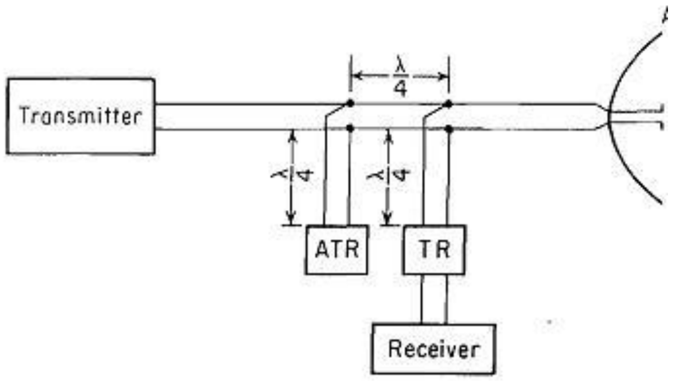


Fig.6.2 Branch type duplexer

(ii) Balanced duplexer: The balanced duplexer, figure 6.3, is based on the short-slot hybrid junction which consists of two sections of waveguides joined along one of their narrow walls with a slot cut in the common narrow wall to provide coupling between the two. The short-slot hybrid may be considered as a broadband directional coupler with a coupling ratio of 3 dB. In the transmit condition power is divided equally into each waveguide by the first short slot hybrid junction. Both TR tubes break down and reflect the incident power out the antenna arm as shown

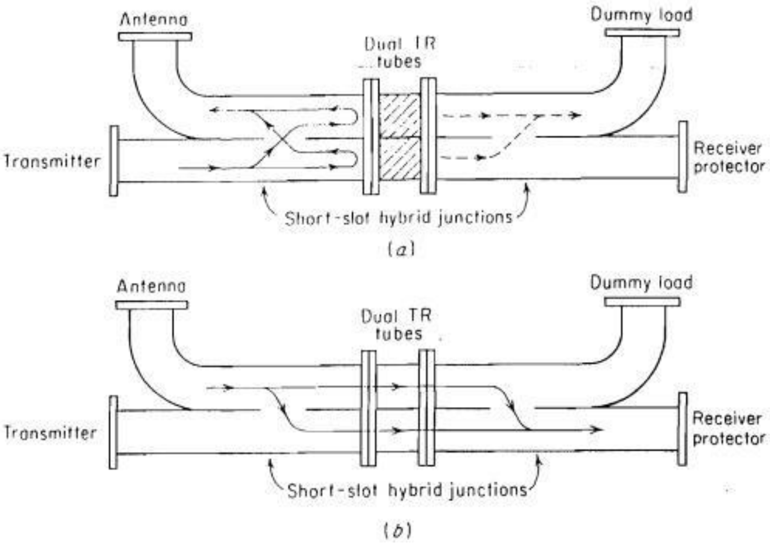


Fig.6.3. Balanced duplexer using dual TR tubes and two short-slot hybrid junctions (a) Transmit condition; (b) receive condition.

The short-slot hybrid has the property that each time the energy passes through the slot in either direction, its phase is advanced 90˚. Therefore, the energy must travel as indicated by the solid lines. Any energy which leaks through the TR tubes (shown by the dashed lines) is directed to the arm with the matched dummy load and not to the receiver, In addition to the attenuation provided by the TR tubes, the hybrid junctions provide an additional 20 to 30 dB of isolation.

On reception the TR tubes are unfired and the echo signals pass through the duplexer and into the receiver as shown in figure 8.3(b). The power splits equally at the first junction and because of the 90˚ phase advance on passing through the slot, the energy recombines in the receiving arm and not in the dummy-load arm. The power-handling capability of the balanced duplexer is inherently greater than that of the branch-type duplexer and it has wide bandwidth, over ten percent with proper design. A receiver protector is usually inserted between the duplexer and the receiver for added protection.

Circulator and receiver protector:

The ferrite circulator is a three-or four-port device that can, in principle, offer separation of the transmitter and receiver without the need for the conventional duplexer configurations. The circulator does not provide sufficient protection by itself and requires a receiver protector as in Fig. 6.4. The isolation between the transmitter and receiver ports of a circulator is seldom sufficient to protect the receiver from damage. However, it is not the isolation between transmitter and receiver ports that usually determines the amount of transmitter power at the receiver, but the impedance mismatch at the antenna which reflects transmitter power back into the receiver. The VSWR is a measure of the amount of power reflected by the antenna. For example, a VSWR of 1.5 means that about 4 percent of the transmitter power will be reflected by the antenna mismatch in the direction of the receiver, which corresponds to an isolation of only 14 dB. About 11 percent of the power is reflected when the VSWR is 2.0, corresponding to less than 10 dB of isolation. Thus, a receiver protector is almost always required. It also reduces to safe level radiations from nearby transmitters. The receiver protector might use solid-state diodes for an all solid-state configuration, or it might be a passive TR-limiter consisting of a radioactive primed TR-tube followed by a diode limiter. The ferrite circulator with receiver protector is attractive for radar applications because of its long life, wide bandwidth, and compact design.

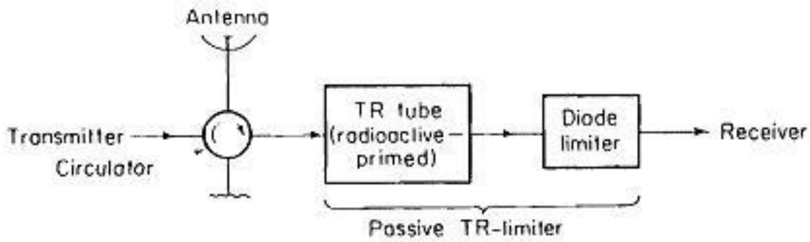


Fig.6.4 Circulator and receiver protector; A four-port circulator is shown with the fourth port terminated in a matched load to provide greater isolation between the transmitter and the receiver than provided by a three-port circulator.

Phased Array Antenna:

The phased array is a directive antenna made up of individual radiating antennas, or elements, which generate a radiation pattern whose shape and direction is determined by the relative phases and amplitudes of the currents at the individual elements. By properly varying the relative phases it is possible to steer the direction of the radiation. The radiating elements might be dipoles open-ended waveguides, slots cut in waveguide, or any other type of antenna. The inherent flexibility offered by the phased-array antenna in steering the beam by means of electronic control is what has made it of interest for radar. It has been considered in those radar applications where it is necessary to shift the beam rapidly from one position in space to another, or where it is required to obtain information about many targets at a flexible, rapid data rate. The full potential of a phased-array antenna requires the use of a computer that can determine in real time, on the basis of the actual operational situation, how best to use the capabilities offered by the array.

Radiation pattern for Phased array Antenna: Consider a linear array made up of N elements equally spaced a distance d apart (figure 6.5). The elements are assumed to be isotropic point sources radiating uniformly in all directions with equal amplitude and phase. Although isotropic elements are not realizable in practice, they are a useful concept in array theory, especially for the computation of radiation patterns. The array is shown as a receiving antenna for convenience, but because of the reciprocity principle, the results obtained apply equally well to a transmitting antenna. The outputs of all the elements are summed via lines of equal length to give a sum output voltage E_a. Element 1 will be taken as the reference signal with zero phase.

The difference in the phase of the signals in adjacent elements is $\Psi = 2\pi\,(d/\lambda)\sin\theta$, where θ is the direction of the incoming radiation. It is further assumed that the amplitudes and phases of the signals at each element are weighted uniformly. Therefore the amplitudes of the voltages in each element are the same and, for convenience, will be taken to be unity. The sum of all the voltages from the individual elements, when the phase difference between adjacent elements is Ψ, can be written

$$E_a = \sin \omega t + \sin(\omega t + \psi) + \sin(\omega t + 2\psi) + \cdots + \sin[\omega t + (N-1)\psi]$$

where ω is the angular frequency of the signal. The sum can be written

$$E_a = \sin\left[\omega t + (N-1)\frac{\psi}{2}\right]\frac{\sin(N\psi/2)}{\sin(\psi/2)}$$

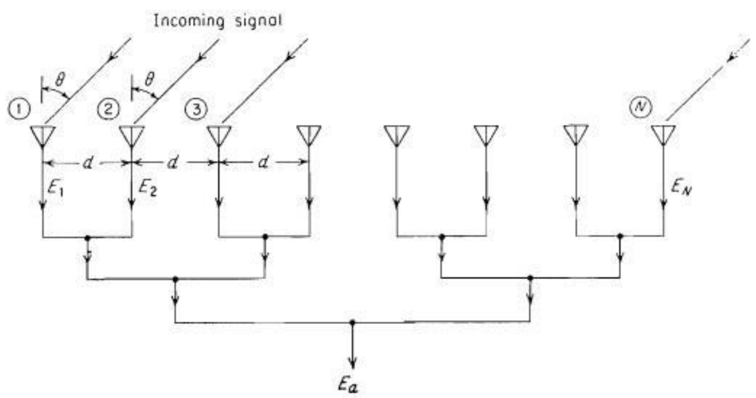


Fig.6.5 N-element linear array

The first factor is a sine wave of frequency ω with a phase shift (N - l)Ψ/2 (if the phase reference were taken at the center of the array, the phase shift would be zero), while the

second term represents an amplitude factor of the form sin (NΨ/2)/sin(Ψ/2). The field intensity pattern is the magnitude of the Eq. above or

$$|E_a(\theta)| = \left| \frac{\sin\,[N\pi(d/\lambda)\sin\theta]}{\sin\,[\pi(d/\lambda)\sin\theta]} \right|$$

The pattern has nulls (zeros) when the numerator is zero. The latter occurs when NΠ (d/λ) sinθ = 0, ± Π, ± 2Π..., ± nΠ, where n = integer. The denominator, however, is zero when Π (d/λ) sinθ, = 0, ± Π, ± 2Π..., ± nΠ. Note that when the denominator is zero, the numerator is also zero. The value of the field intensity pattern is indeterminate when both the denominator and numerator are zero. However, by applying L'Hopital's rule (differentiating numerator and denominator separately) it is found that $|E_a|$ is a maximum whenever sinθ = ± nλ/d. These maxima all have the same value and are equal to N. The maximum at sinθ = 0 defines the main beam. The other maxima are called grating lobes.

They are generally undesirable and are to be avoided. If the spacing between elements is a half-wavelength (d/λ = 0.5), the first grating lobe (n = ± 1) does not appear in real space since sinθ > 1, which cannot be.

Grating lobes appear at ± 90° when d = λ. For a non scanning array (which is what is considered here) this condition (d = λ) is usually satisfactory for the prevention of grating lobes. Equation discussed above applies to isotropic radiating elements, but practical antenna elements that are designed to maximize the radiation at θ = 0°, generally have negligible radiation in the direction θ = ± 90°. Thus the effect of a realistic element pattern is to suppress the grating lobes at ±90°. It is for this reason that an element spacing equal to one wavelength can be tolerated for a non scanning array. E_a (θ) = E_a (Π - θ). Therefore an antenna of isotropic elements has a similar pattern in the rear of the antenna as in the front. The same would be true for an array of dipoles. To avoid ambiguities, the backward radiation is usually eliminated by placing a reflecting screen behind the array. Thus only the radiation over the forward half of the antenna (-90˚ ≤ θ ≤ 90°) need be considered.

The radiation pattern is equal to the normalized square of the amplitude, or

$$G_a(\theta) = \frac{|E_a|^2}{N^2} = \frac{\sin^2\,[N\pi(d/\lambda)\sin\theta]}{N^2\sin^2\,[\pi(d/\lambda)\sin\theta]}$$

If the spacing between antenna elements is λ/2 and if the sine in the denominator of Eq. is replaced by its argument, the half-power beamwidth is approximately equal to

$$\theta_B = \frac{102}{N}$$

The first sidelobe, for N sufficiently large, is 13.2 dB be1ow the main beam. The pattern of a uniformly illuminated array with elements spaced λ/2 apart is similar to the pattern produced by a continuously illuminated uniform aperture. When directive elements are used, the resultant array antenna radiation pattern is

$$G(\theta) = G_e(\theta) \frac{\sin^2 [N\pi(d/\lambda) \sin \theta]}{N^2 \sin^2 [\pi(d/\lambda) \sin \theta]} = G_e(\theta)G_a(\theta)$$

where G_e (θ) is the radiation pattern of an individual element.

The resultant radiation pattern is the product of the element factor Ge(θ) and the array factor $G_a(\theta)$, the latter being the pattern of an array composed of isotropic elements. The array factor has also been called the space factor Grating lobes caused by a widely spaced array may therefore be eliminated with directive elements which radiate little or no energy in the directions of the undesired lobes. For example, when the element spacing d = 2λ, grating lobes occur at θ = ±30˚ and ±90˚ in addition to the main beam at θ = 0˚. If the individual elements have a beamwidth somewhat less than 60°, the grating lobes of the array factor will be suppressed.

Beam steering phased array antennas.

The beam of an array antenna may be steered rapidly in space without moving large mechanical masses by properly varying the phase of the signals applied to each element. Consider an array of equally spaced elements. The spacing between adjacent elements is d, and the signals at each element are assumed of equal amplitude. If the same phase is applied to all elements, the relative phase difference between adjacent elements is zero and the position of the main beam will be broadside to the array at an angle θ = 0. The main beam will point in a direction other than broadside if the relative phase difference between elements is other than zero. The direction of the main beam is at an angle θ_0**,** when the phase difference is Ø = 2π(d/λ) sin θ_0**.** The phase at each element is therefore $Ø_c$ = mØ, where m = 0, 1, 2, . . ,(N - 1), and $Ø_c$ **,** is any constant phase applied to all elements. The normalized radiation pattern of the array when the phase difference between adjacent elements is Ø is given by

$$G(\theta) = \frac{\sin^2 [N\pi(d/\lambda)(\sin \theta - \sin \theta_0)]}{N^2 \sin^2 [\pi(d/\lambda)(\sin \theta - \sin \theta_0)]}$$

The maximum of the radiation pattern occurs when sin θ = sin θ_o. The above equation states that the main beam of the antenna pattern may be positioned to an angle θ_0 by the insertion of the proper phase shift Ø at each element of the array. If variable rather than fixed, phase shifters are used, the beam may be steered as the relative phase between elements is changed.

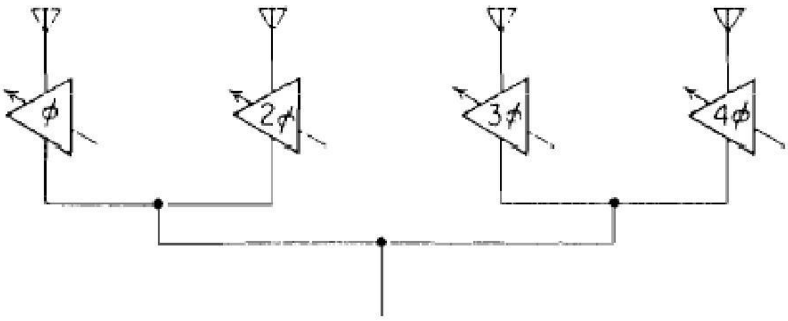


Fig.6.6 Steering of an antenna beam with variable phase shifters (parallel-fed array).

Using an argument similar to the nonscanning array described previously, grating lobes appear at an angle θ_o whenever the denominator is zero, when

$$\pi \frac{d}{\lambda}(\sin \theta_g - \sin \theta_0) = \pm n\pi$$

$$|\sin \theta_g - \sin \theta_0| = \pm n \frac{\lambda}{d}$$

If a grating lobe is permitted to appear at - 90° when the main beam is steered to +90°, it is found from the above that d = λ / 2. Thus the element spacing must not be larger than a half wavelength the beam is to be steered over a wide angle without having undesirable grating lobes appear. Practical array antennas do not scan ±90°. If the scan is limited to ±60°, the element spacing should be less than 0.54λ.

Change of beamwidth with steering angle: The half-power beamwidth in the plane of scan increases as the beam is scanned off the broadside direction. The beamwidth is approximately inversely proportional to $\cos\theta_o$, where θ_o is the angle measured from the normal to the antenna. This may be proved by assuming that the sine in the denominator of Eq. discussed earlier can be replaced by its argument, so that the radiation pattern is of the form $(\sin^2 u)/u^2$, where u = NΠ (d/λ)(sinθ - $\sin\theta_o$). The $(\sin^2 u)/u^2$ antenna pattern is reduced to half its maximum value when u = ± 0.443Π. Denote by θ_+ the angle corresponding to the half-power point when $\theta > \theta_o$, and θ_- , the angle corresponding to the half-power point when $\theta < \theta_o$; that is, θ_+ corresponds to u =+0.443Π and θ_- to u = -0.443Π. The sinθ - sinθ_o, term in the expression for u can be written

$$\sin \theta - \sin \theta_0 = \sin (\theta - \theta_0) \cos \theta_0 - [1 - \cos (\theta - \theta_0)] \sin \theta_0$$

The second term on the right-hand side of Eq. above can be neglected when θ_o is small (beam is near broadside), so that

$$\sin \theta - \sin \theta_0 \approx \sin (\theta - \theta_0) \cos \theta_0$$

Using the above approximation, the two angles corresponding to the 3-dB points of the antenna pattern are

$$\theta_+ - \theta_0 = \sin^{-1} \frac{0.443\lambda}{Nd \cos \theta_0} \approx \frac{0.443\lambda}{Nd \cos \theta_0}$$

$$\theta_- - \theta_0 = \sin^{-1} \frac{-0.443\lambda}{Nd \cos \theta_0} \approx \frac{-0.443\lambda}{Nd \cos \theta_0}$$

The half-power beamwidth is

$$\theta_B = \theta_+ - \theta_- \approx \frac{0.886\lambda}{Nd \cos \theta_0}$$

Therefore, when the beam is positioned an angle θ_o off broadside, the beamwidth in the plane of scan increases as $(\cos \theta_o)^{-1}$. The change in beamwidth with angle θ_o, as derived above is not valid when the antenna beam is too far removed from broadside. It certainly does not apply when the energy is radiated in the end fire direction.

Equation above applies for a uniform aperture illumination. With a cosine-on-a-pedestal aperture illumination of the form $A_n = a_o + 2a_1 \cos 2\Pi n/N$, the beamwidth is

$$\theta_B \simeq \frac{0.886\lambda}{Nd \cos \theta_0} [1 + 0.636(2a_1/a_0)^2]$$

The parameter n in the aperture illumination represents the position of the element. Since the illumination is assumed symmetrical about the center element, the parameter n takes on values of n = 0, ±1, ± 2 . . . ± (N - 1)/2. The range of interest is $0 \leq 2a_1 \leq a_o$ which covers the span fro

uniform illuminations to a taper so severe that the illumination drops to zero at the ends of the array. (The array is assumed to extend a distance d/2 beyond each end element.)

Applications of phased array antennas.

The phased-array antenna has been of considerable interest to the radar systems engineer because its properties are different from those of other microwave antennas. The array antenna takes several forms:

Mechanically scanned array*:*

The array antenna in this configuration is used to form a fixed beam that is scanned by mechanical motion of the entire antenna. No electronic beam steering is employed. This is an economical approach to air-surveillance radars at the lower radar frequencies, such as VHF. It is also employed at higher frequencies when a precise aperture illumination is required, as to obtain extremely low sidelobes. At the lower frequencies, the array might be a collection of dipoles or Yagis, and at the higher frequencies the array might consist of slotted waveguides.

Linear array with frequency scan:

The frequency-scanned, linear array feeding a parabolic cylinder or a planar array of slotted waveguides has seen wide application as a 3D air-surveillance radar. In this application, a pencil beam is scanned in elevation by use of frequency and scanned in azimuth by mechanical rotation of the entire antenna.

Linear array with phase scan:

Electronic phase steering, instead of frequency scanning, in the 3D air-surveillance radar is generally more expensive, but allows the use of the frequency domain for purposes other than beam steering. The linear array configuration is also used to generate multiple, contiguous fixed beams (stacked beams) for 3D radar. Another application is to use either phase- or frequency-steering in a stationary linear array to steer the beam in one angular coordinate, as for the GCA radar.

Phase-frequency planar array:

A two-dimensional (planar) phased array can utilize frequency scanning to steer the beam in one angular coordinate and phase shifters to steer in the orthogonal coordinate. This approach is generally easier than using phase shifters to scan in both coordinates, but as with any frequency-scanned array the use of the frequency domain for other purposes is limited when frequency is employed for beam-steering.

Phase-phase planar array:

The planar array which utilizes phase shifting to steer the beam in two orthogonal coordinates is the type of array that is of major interest for radar application because of its inherent versatility. Its application, however, has been limited by its relatively high cost. The phase-phase array is what is generally implied when the term electronically steered phased array is used.

Advantages and limitations of phased array antennas.

The array antenna has the following desirable characteristics not generally enjoyed by other antenna types:

Inertialess rapid beam-steering:

The beam from an array can be scanned, or switched from one position to another, in a time limited only by the switching speed of the phase shifters. Typically, the beam can be switched in several microseconds, but it can be considerably shorter if desired.

Multiple, independent beams:

A single aperture can generate many simultaneous independent beams. Alternatively, the same effect can be obtained by rapidly switching a single beam through a sequence of positions.

Potential for large peak and / or average power:

If necessary, each element of the array can be fed by a separate high-power transmitter with the combining of the outputs made in space to obtain a total power greater than can be obtained from a single transmitter.

Control of the radiation pattern:

A particular radiation pattern may be more readily obtained with the array than with other microwave antennas since the amplitude and phase of each array element may be individually controlled. Thus, radiation patterns with extremely low sidelobes or with a shaped main beam may be achieved. Separate monopulse sum and difference patterns, each with its own optimum shape, are also possible.

Graceful degradation:

The distributed nature of the array means that it can fail gradually rather than all at once (catastrophically).

Convenient aperture shape:

The shape of the array permits flush mounting and it can be hardened to resist blast.

Electronic beam stabilization:

The ability to steer the beam electronically can be used to stabilize the beam direction when the radar is on a platform, such as a ship or aircraft, that is subject to roll, pitch, or yaw.

Limitations:

1. The major limitation that has limited the widespread use of the conventional phased array in radar is its high cost, which is due in large part to its complexity.
2. When graceful degradation has gone too far a separate maintenance is needed.
3. When a planar array is electronically scanned, the change of mutual coupling that accompanies a change in beam position makes the maintenance of low side lobes more difficult.
4. Although the array has the potential for radiating large power, it is seldom that an array is required to radiate more power than can be radiated by other antenna types or to utilize a total power which cannot possibly be generated by current high-power microwave tube technology that feeds a single transmission line.