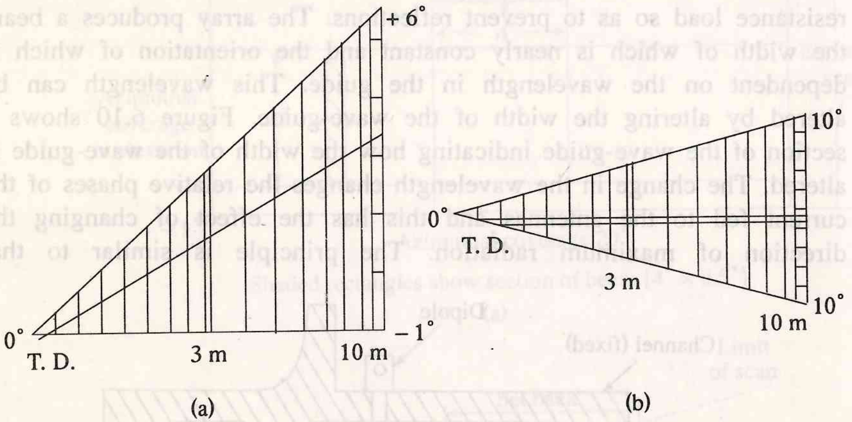
employed in phase array radars for beam steering.[22] This wave-guide has been called a 'squeezable wave-guide'. Altering the section of the wave-guide presents a much simpler mechanical problem than the rotation of any array, because of the lightness of the parts involved.

The data obtained by the PAR are displayed on two long-persistence cathode ray tubes, one displaying the range and elevation angle and the other the range and azimuth angle (Fig. 6.11). Note that the lines of constant range appear as straight lines instead of arcs of circles. Also, the angular sweeps appear expanded and the range sweep is progressively slowed down as the range increases. These deliberately introduced distortions help to increase the accuracy of angle measurements and also to increase range accuracy at the shorter ranges where it is most required. The distance markers appear at certain fixed ranges and angle markers at the ends of the display. The two display tubes are mounted on a single console and one controller uses both. The localizer and glide-slope courses may be permanently marked on the display so that the controller can know the position of the aircraft with respect to the desired path of descent.

While the radar itself is located near the runway end, the console may be in the control tower. The two points will be connected by cables which carry the control signals and video data.

Equipment of the type described is capable of resolving targets which are 60 m apart. The accuracy is such that at a distance of 1 mile, it is possible to detect deviations from glide-slope of as little as 8 m.



(a) (b)

Fig. 6.11 PAR display [(a) elevation, (b) azimuth]. (Range shown in miles)

6.3 MICROWAVE LANDING SYSTEM (MLS)

A later development in landing aids is the microwave landing system which operates in the range of 5031 to 5090 MHz. It was developed to

airports and aerodromes where ground conditions are unfavourable for the operation of the ILS. One of the disadvantages of the ILS is that it provides a single approach path along the extended centre line of the runway. Further, it is 'site-sensitive', and subject to distortion and bending of the approach path due to site irregularities. Relatively small distortions can be overcome by the capture-effect localiser, but this is not always possible and some sites can be so bad that ILS cannot be installed. The main reason for this is that the ILS operates in the VHF/UHF band where the surrounding terrain plays an important part in shaping the beam. It has also the drawback that the number of channels it can provide is limited to 40. It is also prone to interference from broadcasting stations.

The MLS, on the other hand, can accommodate 200 channels. Because of the small wavelength, the antennas are small and they can be designed to be relatively free from the effect of the surrounding area. The technique of scanning employed covers a larger area and permits approach by 'dog-legged' and curved paths.

The basic elements of the MLS are shown in Fig. 6.12.

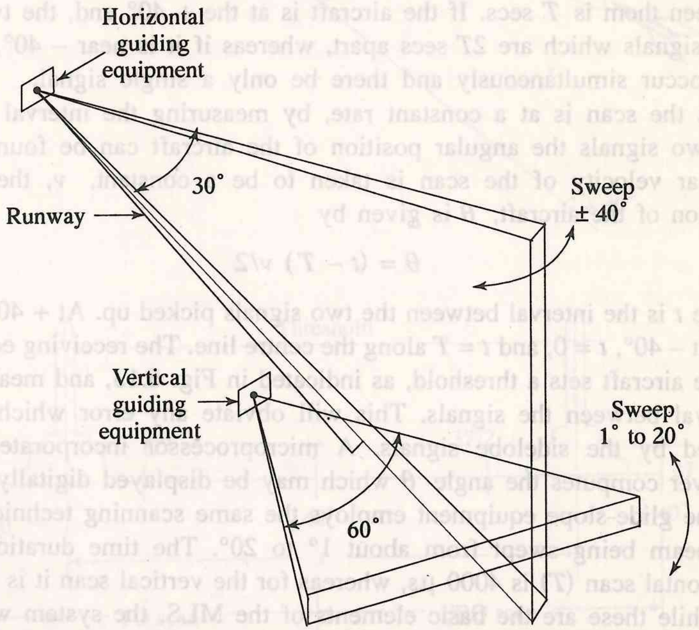


Fig. 6.12 Basic elements of the Microwave Landing System

(a) the azimuth beam equipment which is located at the far end of the runway. This produces a beam narrow in the horizontal plane and wide in the vertical plane which is swept rapidly about a vertical axis from side to side.

(b) the elevation beam equipment which produces a beam narrow in the vertical plane and wide in the horizontal plane is similarly rapidly scanned about a horizontal axis. This is located at the end of the runway near the touch-down point like the glide-slope of the ILS.

(c) In addition, a distance-measuring equipment is provided. This is located near the horizontal guiding equipment at the far end of the runway. This facility, when interrogated by the aircraft, gives the distance from the touch-down point. It is similar to the DME mentioned in Sec. 5.2.3 but is capable of higher precision. It is called 'precision DME' or PDME.

The manner in which the scanning beam gives the horizontal position to the aircraft is illustrated in Fig. 6.13.

The beam scans at a uniform rate from + 40° with respect to the centre line of the runway to – 40°, and back to + 40°. The first is called "TO-scan" and the other "FRO-scan". The aircraft picks up the signals as the beam scans past it. If each scan takes T secs, an aircraft on the centre line of the runway picks up the beam twice, at $T/2$ and $3T/2$, so that the interval between them is T secs. If the aircraft is at the + 40° end, the two scans give signals which are $2T$ secs apart, whereas if it is near – 40°, the two will occur simultaneously and there be only a single signal.

As the scan is at a constant rate, by measuring the interval between the two signals the angular position of the aircraft can be found. It the angular velocity of the scan is taken to be a constant, v, the angular position of the aircraft, θ is given by

$$\theta = (t - T)\, v/2$$

where t is the interval between the two signals picked up. At + 40°, $t = 2T$ and at – 40°, $t = 0$, and $t = T$ along the centre line. The receiving equipment in the aircraft sets a threshold, as indicated in Fig. 6.13, and measures the interval between the signals. This will obviate any error which may be caused by the sidelobe signals. A microprocessor incorporated in the receiver computes the angle θ which may be displayed digitally.

The glide-slope equipment employs the same scanning technique, with the beam being swept from about 1° to 20°. The time duration of the horizontal scan (T) is 4000 µs, whereas for the vertical scan it is 3350 µs.

While these are the basic elements of the MLS, the system which has been accepted as a standard by the International Civil Aviation Organisation (ICAO) is called Time-reference Scanning-beam (TRSB) Microwave Landing System.[34] This employs a time-sharing technique to include a number of other functions, such as Back-azimuth and Flare-guidance, interleaved with digital data to provide information regarding airport and runway identification, site data, wind vectors, system status, and other operational details pertaining to the airport. All the facilities operate on the same frequency, which is possible because of the time-sharing technique. The location of the additional facilities (Back-azimuth and Flare-guidance) is shown in Fig. 6.14 as also the time-sharing between the various

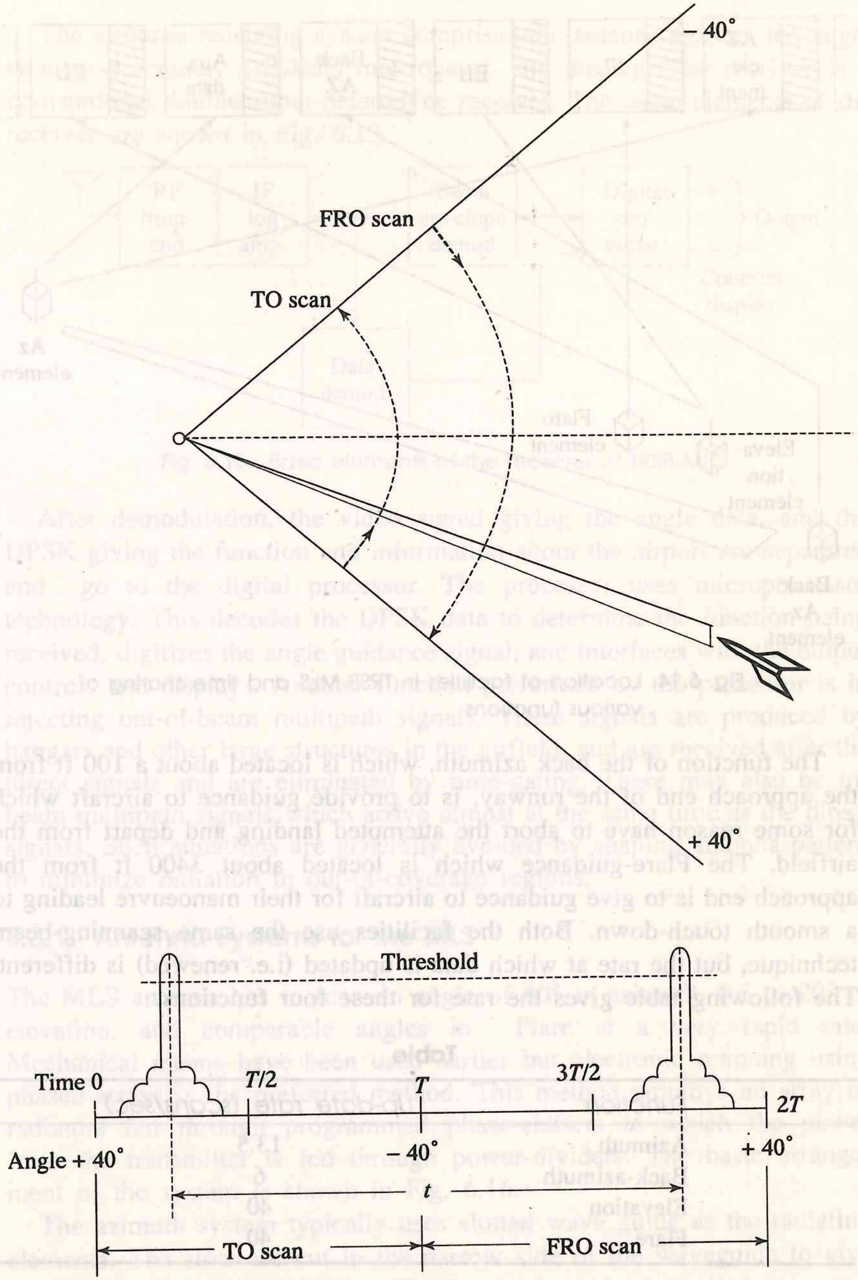


Fig. 6.13 Beam scanning technique in MLS

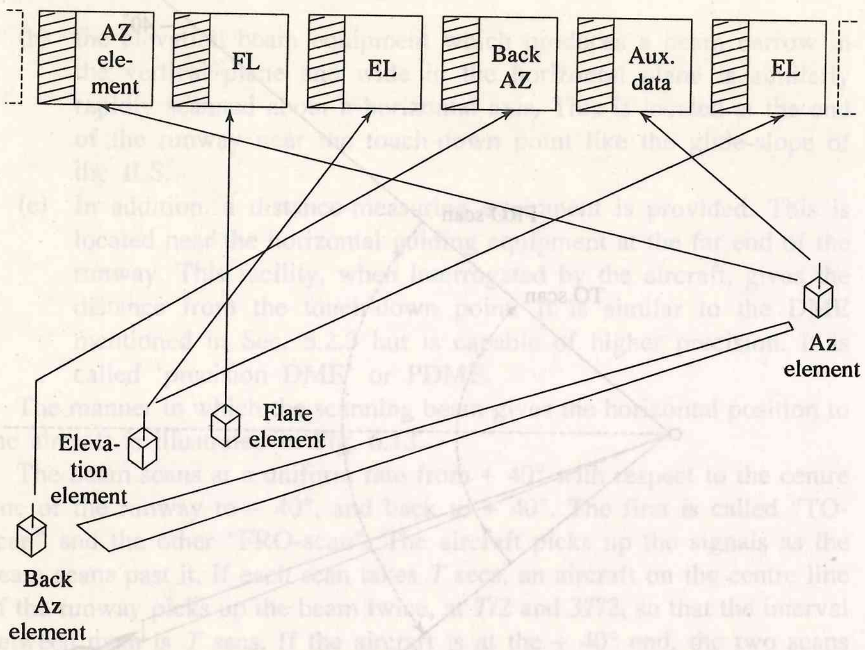


Fig. 6.14 Location of facilities in TRSB-MLS and time sharing of various functions

The function of the back azimuth, which is located about a 100 ft from the approach end of the runway, is to provide guidance to aircraft which for some reason have to abort the attempted landing and depart from the airfield. The Flare-guidance which is located about 3400 ft from the approach end is to give guidance to aircraft for their manoeuvre leading to a smooth touch-down. Both the facilities use the same scanning-beam technique, but the rate at which data is updated (i.e. renewed) is different. The following table gives the rate for these four functions.

Table

Function	*Up-date rate (scans/sec)*
Azimuth	13.5
Back-azimuth	6
Elevation	40
Flare	40

As shown in Fig. 6.14, the time allotted to each function starts with a preamble message (shown shaded) in coded form which identifies the function (azimuth, elevation, etc.). The receiver decodes this and each function is processed independently. The data obtained is intermittent, but because of the high rate of renewal, the guidance information has very little noise.

The airborne receiving system comprises the antenna system, the angle receiver-processor, a PDME Interrogator, and display. The receiver is a conventional double super-heterodyne receiver. The basic elements of the receiver are shown in Fig. 6.15.

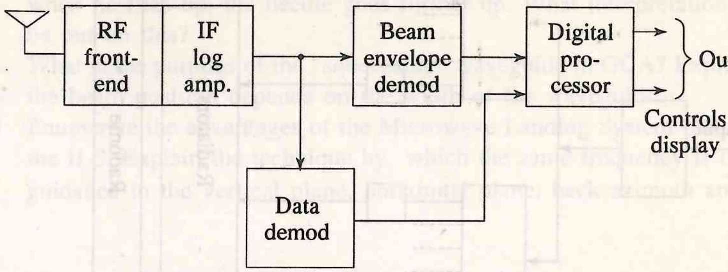


Fig. 6.15 Basic elements of the receiver of TRSB-MLS

After demodulation, the video signal giving the angle data, and the DPSK giving the function and information about the airport are separated and go to the digital processor. The processor uses microprocessor technology. This decodes the DPSK data to determine the function being received, digitizes the angle guidance signal, and interfaces with the output controls and displays. Another function performed by the processor is in rejecting out-of-beam multipath signals. These signals are produced by hangars and other large structures in the airfield, and are received after the direct signals and are eliminated by time-gating. There may also be in-beam multipath signals which arrive almost at the same time as the direct signals. Such situations are generally avoided by shaping antenna pattern to minimize radiation in out-of-coverage regions.

6.3.1 Antenna Systems for the MLS

The MLS antenna has to scan an angle of 80° in azimuth, 0.5 to 20° in elevation, and comparable angles in Flare at a very rapid rate. Mechanical means have been used earlier but electronic scanning using phased-arrays is the preferred method. This method employs an array of radiators fed through programmed phase-shifters to which the power from the transmitter is fed through power-dividers. The basic arrangement of the system is shown in Fig. 6.16.

The azimuth system typically uses slotted wave guide as the radiating elements. The slots are cut in the narrow side of the waveguide to give a vertically polarised radiation. The beam shape is controlled to ensure a sharp cut-off at negative angles of radiation and thereby avoid the effects of reflection from the ground. The beam width of the azimuth antenna depends on the runway length, as it has to provide adequate accuracy near the touch-down point. Runways of 15,000 ft require a 1° beam width antenna, 8000 ft runways require a 2° beam width while a

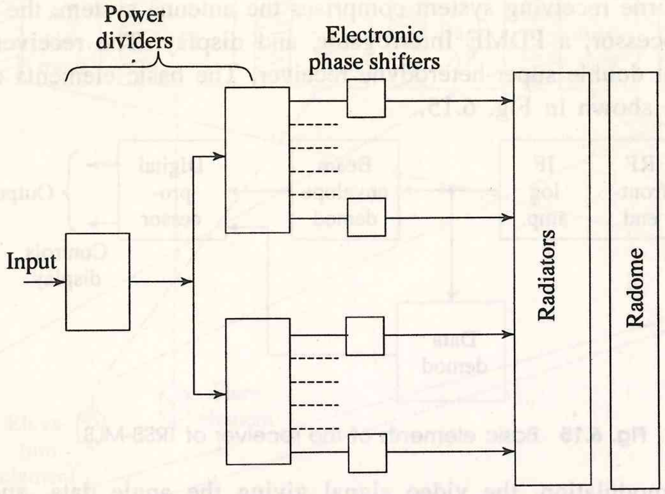


Fig. 6.16 The antenna system

runway of 5000 ft or less requires a 3° beam-width. The antenna for this last may have 40 waveguide radiators. Phase shifters used with the MLS are typically 4-bit diode phase shifters which give phase increments in steps of $22\frac{1}{2}°$. To keep the beam scan rate uniform, the phase shifters will have to be incremented at a non-uniform rate. This function, as well as the transfer of power in keeping with signal format are done by a unit called the Beam Steering Unit (BSU).

The vertical guidance antenna consists typically of a collinear array of dipoles in a shaped ground-plane. The radiating elements are generally printed dipoles, horns or micro-strip patches. Considerable development effort has been devoted to designing antenna arrays having less number of radiating elements (thinned arrays) without sacrificing the performance.

The MLS being an all-weather landing system which gives guidance at the most critical phase of flight, i.e. approach and landing, reliability is of paramount importance. Consequently, provision is made for ensuring integrity and reliable operation by continuous monitoring of its performance.

The MLS is planned to replace the ILS in stages in course of time.

QUESTIONS AND PROBLEMS

1. The Localiser gives two equi-signal paths [*XO* and *XO′* in Fig. 6.2] of which *XO* is the normal approach path. Can the pilot use *X′O* to come to the runway? If so, how does he have to interpret the indicator readings?
2. The following indications appear on the cross-pointer indicator [Fig. 6.7]

(i) The horizontal indicator is above the horizontal line.
(ii) The vertical needle is to the left of the central vertical line.

Indicate what action the pilot has to take.

3. A pilot descending along the glide-path finds the "fly-up" indication but when he flies up, the needle goes further up. What interpretation should be put on this?
4. What is the purpose of the "squeezable" waveguide in GCA? Explain how the beam position depends on the width of the waveguide.
5 Enumerate the advantages of the Microwave Landing System [MLS] over the ILS. Explain the technique by which the same frequency is used for guidance in the vertical plane, horizontal plane, back azimuth and flare.

7

Doppler Navigation

A self-contained dead-reckoning navigation system requires some device within the craft for finding its velocity vector with respect to some reference direction (such as the true North) on the surface of the earth. The Doppler navigation system is 'self-contained' in this sense. It employs the Doppler effect to determine the velocity of the craft in a frame of coordinates fixed with respect to the aircraft. The velocity with respect to the conventional earth coordinates such as the true or magnetic North is obtained by combining this information with the direction of the aircraft which may be obtained from a gyrocompass. The complete Doppler navigation equipment generally includes a computer which automatically computes navigational data required, such as the track, the distance covered, etc. by combining the velocity data obtained by Doppler effect with the directional reference provided by auxiliary means. Doppler navigation is used only in aircraft.

7.1 THE DOPPLER EFFECT

The equipment used for Doppler navigation is a Doppler radar carried in the aircraft which directs a beam of electromagnetic waves towards the earth. Some of the energy re-radiated by the earth towards the aircraft is received and comparison is made between the frequencies of the transmitted and received signals. When the aircraft has a component of velocity in the direction of the beam, the difference frequency (called the Doppler shift) is nearly proportional to the velocity component. This is Doppler effect, which holds good for all types of waves, though in this case we are concerned with electromagnetic waves.

Let the radar be moving with a velocity which has a component v_r in the direction of the beam (Fig. 7.1). Let f_t be the transmitted frequency. The frequency, as measured at the stationary reflecting object (E) is $f_t\left(1+\dfrac{v_r}{c}\right)$ where c is the velocity of electromagnetic waves. The reflected

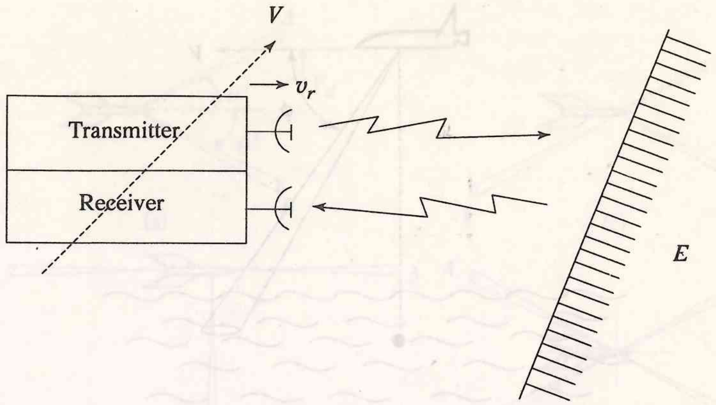


Fig. 7.1 Principle of Doppler radar

waves have this frequency but when they are received by the radar, because of the motion of the receiver towards the object, the received frequency is increased in the ratio $\left(1+\dfrac{v_r}{c}\right)$. The received frequency f_r is, therefore, given by:

$$f_r = f_t\left(1+\frac{v_r}{c}\right)^2 = f_t\left(1+\frac{2v_r}{c}+\frac{v_r^2}{c^2}\right) \tag{7.1}$$

The term $\dfrac{v_r^2}{c^2}$ is almost always negligible because of the high value of c.

So the 'Doppler shift' f_D may be taken to be:

$$f_D = f_r - f_t \approx f_t\,\frac{2v_r}{c} = \frac{2v_r}{\lambda} \tag{7.2}$$

where λ is the wavelength of the transmitted signal. By mixing the received signal with a part of the transmitted signal, f_D may be obtained and v_r determined from Eq. 7.2. Note that f_D may be either positive or negative, depending on the sign of v_r. If the radar is moving towards the object (Fig. 7.1), the received frequency is higher, and the Doppler shift is positive. If it is moving away from the object, the received frequency is lower and Doppler shift is negative.

7.2 BEAM CONFIGURATIONS

Consider an aircraft flying over the earth, transmitting electromagnetic waves in a narrow beam making an angle ϕ with the horizontal (Fig. 7.2). If the aircraft is in level flight, and the beam is directed in the vertical plane containing the forward velocity V of the aircraft, the component of

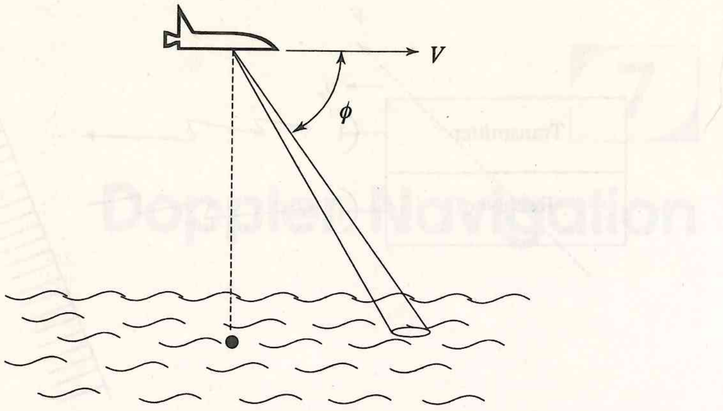


Fig. 7.2 Doppler radar in an aircraft in level flight

the velocity in the direction of the beam is $V \cos \phi$ and the Doppler shift is $\frac{2V\cos\phi}{\lambda}$. In this case, one beam is sufficient to obtain V, as ϕ is a known angle. But if the aircraft has a velocity component perpendicular to this, i.e. to the longitudinal axis of the aircraft, it would not contribute to the Doppler shift, as the component of this velocity along the beam is zero. Therefore, more than one beam is required to obtain the velocity in the general case. The aircraft velocity has, in general, three components with reference to a system of orthogonal coordinates fixed with reference to the aircraft frame. These are the forward (or heading) velocity along its longitudinal axis, the drift velocity perpendicular to it and the vertical velocity. In general, therefore, three beams are necessary to obtain the three components. In some cases, four beams are used. Some of the configurations used in Doppler radar are shown in Fig. 7.3 in plan.

In all cases, the position of the set of beams, called the 'beam cluster' has a fixed relation to the antenna system, and we may define a longitudinal axis and a transverse axis of the cluster, as shown by dotted lines in Fig. 7.3 (d). The antenna system (and consequently the beam cluster) may be fixed to the aircraft frame or it may be put on a stabilized mount corrected for pitch and roll of the air frame. If the antenna is not stabilized, the velocity components obtained from the Doppler shifts pertain to the aircraft reference coordinates as shown in Fig. 7.4(a). What is required for navigation is the set of velocity components with respect to earth. With a stabilized antenna [Fig. 7.4(b)] this is directly obtained. If a fixed antenna system is employed, the velocity components measured have to be converted to the earth system of coordinates by taking into account the pitch and roll angles.[20] The effect of pitch and roll is discussed qualitatively for some of the beam configurations in the next paragraph.

Fig. 7.3(a) shows a two-beam system, in which the beams, (marked A and B) when projected to a horizontal plane make an angle θ_0 with the longitudinal axis of the antenna. (This will coincide with the heading

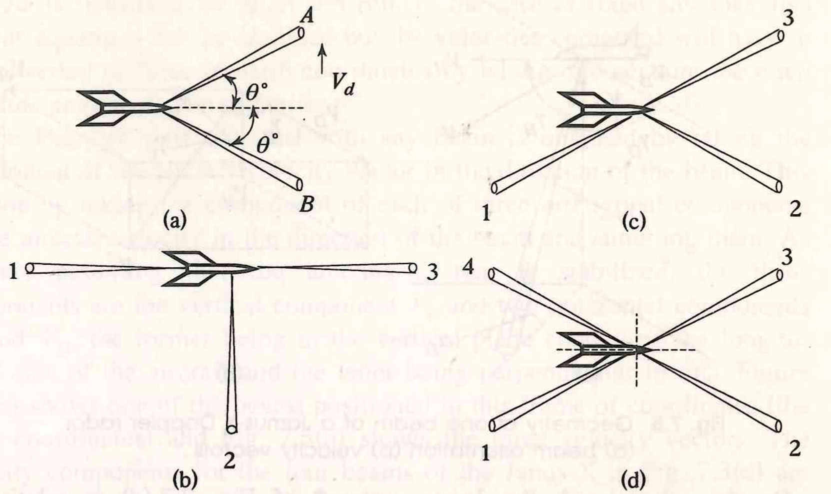


Fig. 7.3 Doppler radar beam configuration [(a) two-beam non-Janus system, (b) three-beam Janus-T, (c) three-beam Janus-λ, (d) four-beam Janus (Janus-X)]

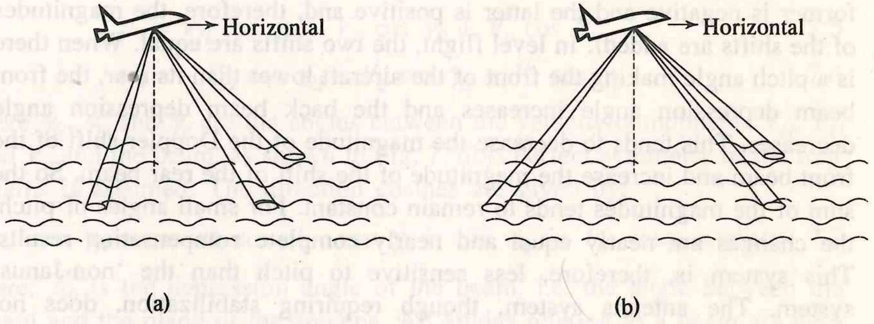


Fig. 7.4 Antenna system (a) fixed to the aircraft (b) and stabilized for pitch and roll

shifts obtained with the two beams are equal. When there is a drift in the direction indicated in the figure, the Doppler shift of A increases and that of B decreases, both in proportion to the drift velocity. By taking the sum of the two frequencies, the heading velocity is obtained, and by taking the difference, the drift velocity is obtained. A vertical component of the velocity, by itself, changes both the Doppler frequencies equally and this may be wrongly interpreted as a forward (or backward) velocity depending on the sign of the change. So this beam configuration is subject to errors unless the vertical velocity information is obtained by other means and used to correct the shift. Another disadvantage of this antenna system is that it is sensitive to the vertical attitude of the aircraft. Pitching, for example, results in the change of the angles γ of Fig. 7.5 and the

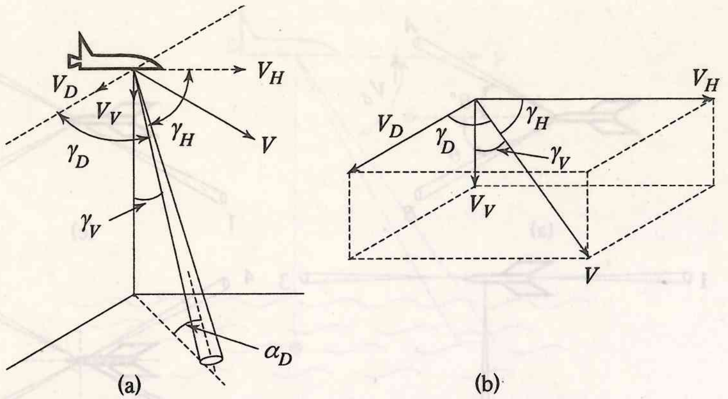


Fig. 7.5 Geometry of one beam of a Jamus-X Doppler radar (a) beam orientation (b) velocity vectors

On the other hand, the Janus system* of Fig. 7.3 (d) is relatively insensitive to pitch and roll of the aircraft as the following qualitative reasoning will show. With this Janus system, the Doppler shift of the rear beam is subtracted from that of the front beam. (It is to be noted that the former is negative and the latter is positive and, therefore, the magnitudes of the shifts are added). In level flight, the two shifts are equal. When there is a pitch angle making the front of the aircraft lower than its rear, the front beam depression angle increases and the back beam depression angle decreases. This tends to decrease the magnitude of the Doppler shift of the front beam and increase the magnitude of the shift of the rear beam. So the sum of the magnitudes tends to remain constant. For small angles of pitch, the changes are nearly equal and nearly complete compensation results. This system is, therefore, less sensitive to pitch than the 'non-Janus' system. The antenna system, though requiring stabilization, does not require such a high degree of stabilization as the 'non-Janus' system.

Another feature of the Janus system is that, with a stabilized antenna, the Doppler shift due to vertical velocity cancels out, because the vertical velocity components of the shift in the forward and backward beams are equal and cancel out when the difference is taken.

Two other beam configurations are shown in Fig. 7.3(b) and (c). These are called the three beam Janus-T and Janus-λ respectively. The angle θ_0 is the same for the two forward and one backward beam of Janus-λ. The configuration of Fig. 7.3(d) is called 'Janus-X'.

7.3 DOPPLER FREQUENCY EQUATIONS

The Doppler shifts obtained with each of the four beams of the Janus-X antenna will now be calculated. For simplicity we will assume that the antenna is stabilized for pitch and roll. In the case of fixed antennas also similar equations can be obtained but the velocities computed will have to be converted to those in earth coordinates by taking into account the pitch and roll angles, as stated earlier.

The Doppler shift obtained with any beam is obtained by taking the component of the aircraft velocity vector in the direction of the beam. This is done by taking the component of each of three orthogonal components of the aircraft velocity in the direction of the beam and summing them. As we are assuming that the antenna system is stabilized, the three components are the vertical component V_v and two horizontal components V_H and V_D, the former being in the vertical plane containing the longitudinal axis of the aircraft and the latter being perpendicular to this. Figure 7.5 (a) shows one of the beams positioned in this frame of coordinates (the earth coordinates) and Fig. 7.5(b) shows the three velocity vectors. The velocity components for the four beams of the Janus-X in Fig. 7.3(d) are as given in Eqs. 7.3 to 7.6.*

$$v_1 = -V_H \cos\gamma_H + V_D \cos\gamma_D + V_v \cos\gamma_v \qquad 7.3$$

$$v_2 = V_H \cos\gamma_H + V_D \cos\gamma_D + V_v \cos\gamma_v \qquad 7.4$$

$$v_3 = V_H \cos\gamma_H - V_D \cos\gamma_D + V_v \cos\gamma_v \qquad 7.5$$

$$v_4 = -V_H \cos\gamma_H - V_D \cos\gamma_D + V_v \cos\gamma_v \qquad 7.6$$

Here γ_H, γ_D and γ_v are the angles between the velocity components V_H, V_D and V_v and the beam, as shown in Fig. 7.5(b). Perfect symmetry of the four beams is assumed. The direction cosines are given by:

$$\cos\gamma_H = \cos\alpha_0 \cos\theta_0 \,;\; \cos\gamma_D = \cos\alpha_0 \sin\theta_0;\; \cos\gamma_v = \sin\alpha_0 \qquad 7.7$$

Here, α_0 is the depression angle of the beam, i.e. the angle between the beam and the plane of the antenna. All angles relating to a beam are with the 'beam centroid' as reference. The controid will be defined later.

The relative velocity components (v_1, v_2, v_3, v_4) are obtained by multiplying the corresponding Doppler shifts f_{D1}, f_{D2}, f_{D3} and f_{D4} assumed to be available individually, by the factor λ/2. The solution of simultaneous Eqs. 7.3 to 7.6 gives the velocity components V_H, V_D and V_v. For the Janus-X arrangement, the solutions are readily obtained by inspection:

$$V_H = \frac{v_2 + v_3 - (v_1 + v_4)}{4\cos\gamma_H} = \frac{\lambda}{8}\,\frac{(f_{D_2} + f_{D_3}) - (f_{D_1} + f_{D_4})}{\cos\alpha_0 \sin\theta_0} \qquad 7.8$$

$$V_D = \frac{(v_1 + v_2) - (v_3 + v_4)}{4\cos\gamma_D} = \frac{\lambda}{8}\,\frac{(f_{D_1} + f_{D_2}) - (f_{D_3} + f_{D_4})}{\cos\alpha_0 \sin\theta_0} \qquad 7.9$$

---

* Janus is the name of a Greek god who could look both ways. The name is used as

*Any three of these equations are sufficient to solve for V_H, V_D and V_V. But when all the four beams are present, as in Janus X, the solution take a form which is simple

$$V_v = \frac{v_1 + v_2 + v_3 + v_4}{4\cos\gamma_D} = \frac{\lambda}{8}\frac{(f_{D_1} + f_{D_2} + f_{D_3} + f_{D_4})}{\sin\alpha_0} \qquad 7.10$$

Similar solutions may be obtained for the 3-beam Janus-λ antenna system, by using only Eqs. 7.3, 7.4 and 7.5. Actually, in Janus-X system, only $(f_{D_3} - f_{D_1})$ and $(f_{D_2} - f_{D_4})$ may be available due to the nature of implementation of the Doppler shift measurement. In this case, only V_H and V_D can be found and V_v cannot be determined.

An important parameter of a Doppler radar is the angle between the longitudinal axis of the antenna and the beam centroid, i.e. the angle γ_H in the Janus systems. The smaller this angle, larger is the Doppler shift for a given velocity. But, for small values of γ_H, the signal returned from water surface becomes very small. This puts a lower limit to the value of this angle. A compromise is effected between the requirements of high sensitivity and appreciable return from the surface of water. The angle has generally a value between 65° and 80°. The value of the angle θ_0 is governed by sensitivity to drift velocity. Larger values of θ_0 give greater sensitivity to drift.

7.4 TRACK STABILIZATION

Equations 7.8 to 7.10 pertain to an antenna system in which the principal axis of the antenna system is in the same vertical plane as the aircraft longitudinal axis or heading. Such an antenna system is said to be *heading-stabilized*. In this case, the two horizontal components of the velocity, V_H and V_D are separately computed and the actual direction of motion of the aircraft (the 'track'), can be computed. The angle between the heading and track is called *drift-angle* (δ) and is given by the relation:

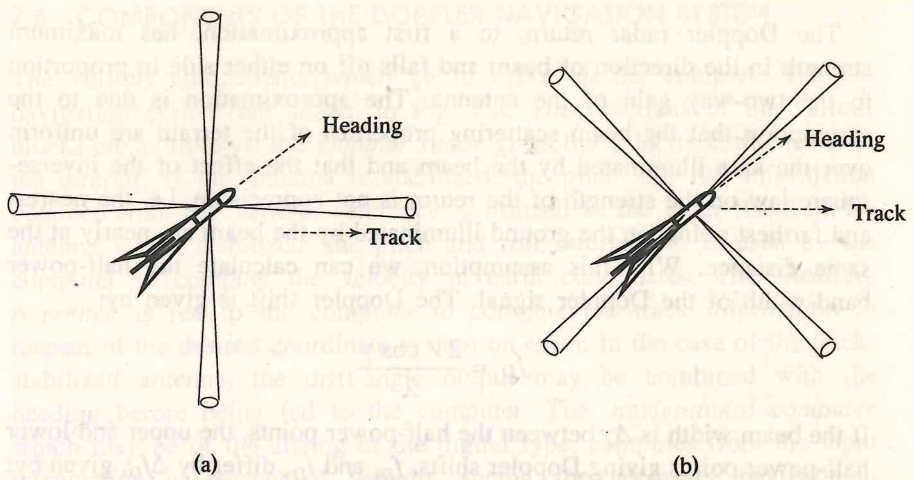
$$\tan\,\delta = \frac{V_D}{V_H} \quad \text{or} \quad \delta = \text{arc}\tan\frac{V_D}{V_H} \qquad 7.11$$

and the track speed $\quad V_g = (V_H^2 + V_D^2)^{1/2} = V_H \sec\,\delta$

There is an alternative type of mechanization, in which the antenna system is rotated about the vertical axis by a servo-system which is actuated by the drift component V_D. The antenna system is turned until the drift component reduces to zero. The track speed is given in this case by V_H itself. The axis of the antenna is then oriented in the direction of the track. This type of antenna is called *track-stabilized* or *drift-angle stabilized* antenna. The orientation of the beam cluster, when there is a drift, is illustrated for the two types of antennas in Fig. 7.6.

7.5 DOPPLER SPECTRUM

The Doppler shift in the above analysis has been taken as the frequency corresponding to the 'beam centroid'. If the beams were very narrow, antenna beams have finite width (generally about 4°) and cover a number of values of depression angle. The returns are consequently of different frequencies from different parts of the beam, and the resultant shift has the nature of a narrow-band frequency spectrum. The various scattering elements in the area illuminated by the beam give returns at different frequencies and in random phase. The spectrum, therefore, resembles that of narrow-band noise, superposed on white Gaussian noise which is always present. Figure 7.7 depicts the type of spectrum obtained with Doppler radar. The two small peaks on either side of the main peak are caused by side lobes in the antenna pattern and do not interfere with the frequency measurements. The shape of the main spectrum is approximately Gaussian. The amplitude of the peak is a function of radar parameters (transmitted power, antenna aperture, range, etc.) and the back-scattering property of the terrain. The Doppler frequency which has been referred to so far is the mean of this spectral distribution. The corresponding angle defines the 'beam centroid'.



(a) (b)

Fig. 7.6 Beam configuration [(a) antenna heading-stabilized, (b) antenna track-stabilized]

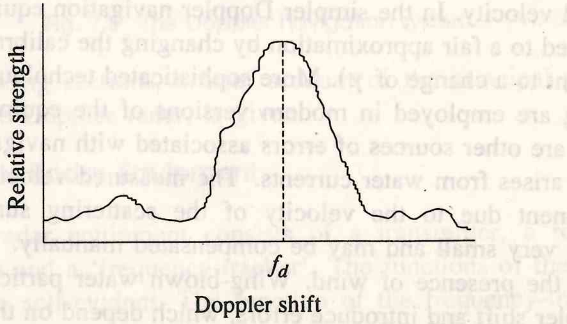


Fig. 7.7 The Doppler spectrum

The Doppler radar return, to a first approximation, has maximum strength in the direction of beam and falls off on either side in proportion to the two-way gain of the antenna. The approximation is due to the assumption that the beam scattering properties of the terrain are uniform over the area illuminated by the beam and that the effect of the inverse-square-law on the strength of the return is not appreciable, i.e. the nearest and farthest points on the ground illuminated by the beam are nearly at the same distance. With this assumption, we can calculate the half-power band-width of the Doppler signal. The Doppler shift is given by:

$$f_D = \frac{2V\cos\gamma}{\lambda}$$

If the beam width is $\Delta\gamma$ between the half-power points, the upper and lower half-power points giving Doppler shifts, f_{D_1} and f_{D_2} differ by Δf_D, given by:

$$f_{D_1} - f_{D_2} = \frac{2V}{\lambda}\left[\cos\left(\gamma - \frac{\Delta\gamma}{2}\right) - \cos\left(\gamma + \frac{\Delta\gamma}{2}\right)\right]$$

$$\Delta f_D \approx \frac{2V}{\lambda}\cdot\sin\gamma\,\Delta\gamma$$

The width of this spectrum is proportional to the central Doppler frequency. The relative width of the spectrum, $\Delta f_D/f_D$ is given by:

$$\frac{\Delta f_D}{f_D} = \frac{\sin\gamma\,\Delta\gamma}{\cos\gamma} = \tan\gamma\,\Delta\gamma \qquad 7.12$$

This quantity is generally between 15% and 25%.

The assumption of uniform back-scattering properties of the terrain holds good over land but is not valid over still water, where the back-scatter changes with the value of γ, being higher for the higher values of γ and decreasing as γ decreases. This tends to distort the spectrum of the return, which will no longer be symmetrical but tends to have a peak on the lower side of the frequency corresponding to the beam centroid. This could cause errors (called 'over-water calibration shift' errors) in the computed velocity. In the simpler Doppler navigation equipment, the error is corrected to a fair approximation by changing the calibration by a switch (equivalent to a change of γ). More sophisticated techniques, such as lobe-switching are employed in modern versions of the equipment.[20]

There are other sources of errors associated with navigation over water. The first arises from water currents. The measured velocities then include a component due to the velocity of the scattering surface. These are generally very small and may be compensated manually. The second type arises in the presence of wind. Wing-blown water particles contribute to the Doppler shift and introduce errors, which depend on the wind direction and speed. Automatic correction for these, taking into account the surface

7.6 COMPONENTS OF THE DOPPLER NAVIGATION SYSTEM

The various components which go to make up the complete Doppler navigation system are shown in Fig. 7.8. The functions of the various blocks are as follows: the *Doppler radar* gives the velocity components of the aircraft. If the antenna is stabilized, the pitch and roll information obtained from the *vertical reference* is utilized in the radar itself. If the antenna is not stabilized, the pitch and roll information is used by the computer to compute the velocity in earth coordinates. The *heading reference* is fed to the computer to compute the track information in respect of the desired coordinate system on earth. In the case of the track-stabilized antenna, the drift-angle output may be combined with the heading before being fed to the computer. The *navigational computer* which may be of the analog or the digital type, computes from the input information in respect of velocity vector, the necessary navigational information. The desired course and distance to be travelled may also be given as inputs to the computer and the distance to be covered, present position, track-angle error and other data for use in navigation may be obtained as outputs of the computer and suitably displayed.

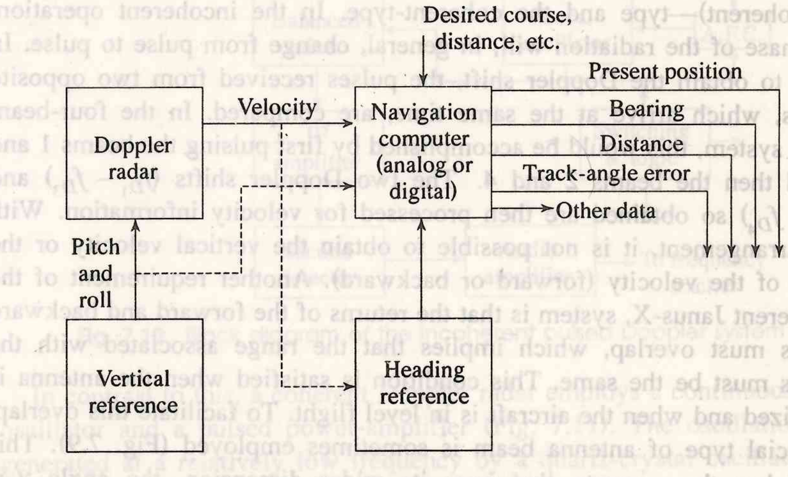


Fig. 7.8 The Doppler navigation system

In the following sections, a brief account of the principal part of the system, viz. the Doppler radar, is given.

7.6.1 Doppler Radar Equipment

The Doppler radar equipment consists of a transmitter, a receiver, an antenna system and a 'frequency-tracker'. The functions of the first three components are self-evident. The function of the frequency-tracker is to locate the centre frequency of the Doppler spectrum and give a sinusoidal output at that frequency. The transmitter operates in one of the two microwave bands—the *X*-band or the *K*-band, in the region of 9 GHz or 13 GHz, respectively. The Doppler radar may be either of the pulsed type or the continuous wave type.

In pulsed radars, the Doppler frequency is obtained by heterodyning the received pulses with either a continuous wave or a pulse from another beam. The output thus obtained consists of samples of the Doppler wave-form, each sample being of the duration of the pulse and the sampling rate being the pulse recurrence frequency. To reconstruct the wave-form, the number of samples per second (i.e. the pulse recurrence frequency (prf)) must be higher than twice the highest frequency in the wave-form. This is a consequence of the sampling theorem. For supersonic aircraft, with the radar frequency in the 10–13 GHz region, the highest Doppler frequency may be a few tens of kHz. This dictates the use of a prf of the order of 100 kHz, in contrast to search radars where the prf may be of the order of a few hundred Hz. As a result of the high prf, the Doppler shift can be measured unambiguously while the range cannot be determined unambiguously.

Pulsed Doppler radar may be one of the two types–the incoherent (or self-coherent)—type and the coherent-type. In the incoherent operation, the phase of the radiation will, in general, change from pulse to pulse. In order to obtain the Doppler shift, the pulses received from two opposite beams, which arrive at the same time, are compared. In the four-beam Janus system, this would be accomplished by first pulsing the beams 1 and 3 and then the beams 2 and 4. The two Doppler shifts $(f_{D_1} - f_{D_3})$ and $(f_{D_2} - f_{D_4})$ so obtained are then processed for velocity information. With this arrangement, it is not possible to obtain the vertical velocity or the sense of the velocity (forward or backward). Another requirement of the incoherent Janus-X, system is that the returns of the forward and backward beams must overlap, which implies that the range associated with the beams must be the same. This condition is satisfied when the antenna is stabilized and when the aircraft is in level flight. To facilitate this overlap, a special type of antenna beam is sometimes employed (Fig. 7.9). This beam has the property that over its wider dimension, the angle γ is constant, thereby ensuring that the Doppler shift associated with the various parts of the beam has the same value. But because of the spread in the area of ground illuminated, the pulse returns are wider in time and overlap is facilitated.

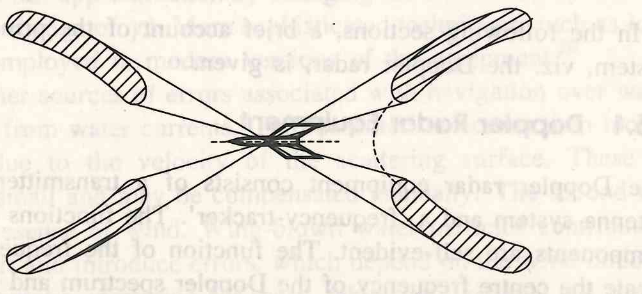


Fig. 7.9 Shaping of beam to facilitate pulse overlap

The block diagram of an incoherent system is shown in Fig. 7.10. A pulsed magnetron is used as the transmitter and this is switched to the beam pairs (1–3, 2–4) sequentially. A duplexer is used to permit common transmit-receive antenna operation. The received signal is applied to a super-heterodyne receiver, the output of which is the Doppler frequency signal. Automatic frequency control (AFC) of the local oscillator is necessary at these frequencies and, therefore, a sample of the transmitted signal is taken from a directional coupler and applied to the AFC circuit.

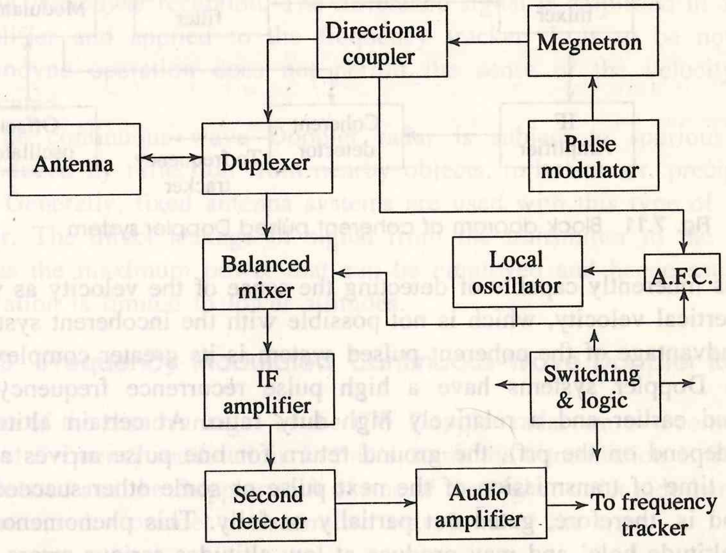


Fig. 7.10 Block diagram of the incoherent pulsed Doppler system

In contrast to this, a coherent Doppler radar employs a continuous wave oscillator and a pulsed power-amplifier (Fig. 7.11). The oscillations are generated at a relatively low frequency by a quartz-crystal oscillator and the frequency is stepped up by a chain of multipliers using step-recovery diodes or varactors. The local oscillator frequency is generated by heterodyning the oscillations at the transmission frequency with an oscillator at the intermediate frequency. The output of the mixer is, therefore, centred at the intermediate frequency (IF). The mixer and IF amplifier are followed by a coherent detector to which the other input is a reference frequency voltage. This reference frequency is obtained by mixing the IF with an 'offset' oscillator output and taking the difference frequency output. By appropriately setting the offset oscillator, both negative and positive Doppler shifts are obtainable. The coherent Doppler

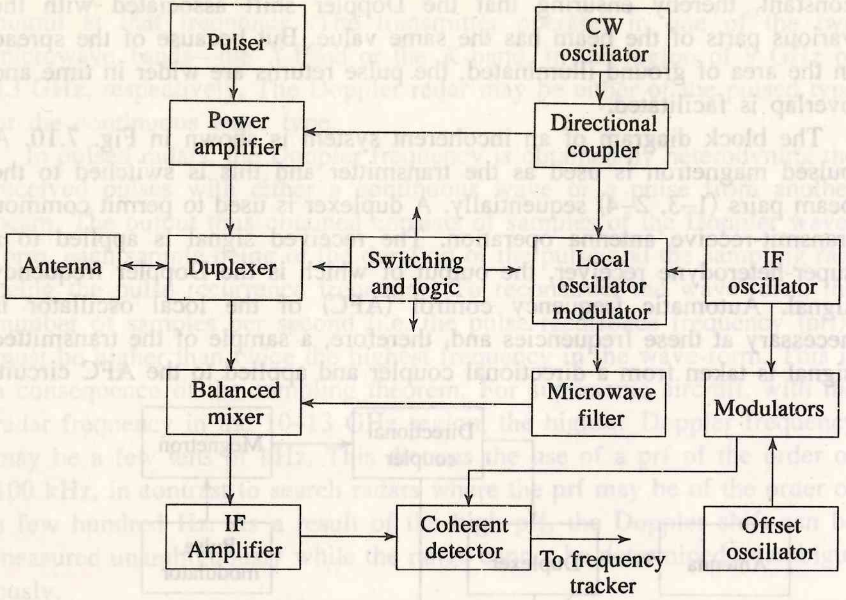


Fig. 7.11 Block dagram of coherent pulsed Doppler system

system is inherently capable of detecting the sense of the velocity as well as the vertical velocity, which is not possible with the incoherent system. The disadvantage of the coherent pulsed system is its greater complexity.

Pulse Doppler systems have a high pulse recurrence frequency as mentioned earlier and a relatively high duty ratio. At certain altitudes (which depend on the prf), the ground return for one pulse arrives at or near the time of transmission of the next pulse or some other succeeding pulse and is, therefore, gated out partially or fully. This phenomenon is called 'altitude hole' and may produce at low altitudes serious errors due to signal elimination and spectral weighting.[20] Altitude holes may be eliminated by changing or 'wobbling' the prf or by making the repetition frequency depend on the range in such a manner that the returns do not arrive at the time of transmission of a pulse.

7.6.2 Continuous Wave Doppler Radar

Continuous waves may also be used in Doppler radar. In this case, separate transmitting and receiving antennas are required for preventing the transmitter output from entering the receiver. The antennas also generally have to be separated physically. Except for this, the equipment is simpler, as the pulsing of the transmitter is avoided and a reference signal (part of the transmitter output) is always available (Fig. 7.12). The block diagram shows the essential parts of the system. The Doppler difference frequency is obtained by direct heterodyning of the transmitted and received signals.

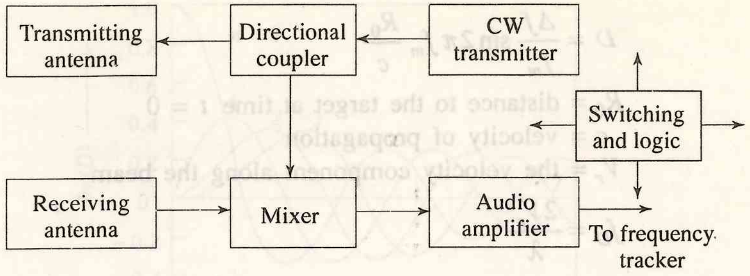


Fig. 7.12 Block diagram of continuous wave Doppler radar

This is equivalent to having an intermediate frequency of zero, and is called *homodyne* reception. The difference signal is amplified in an audio amplifier and applied to the frequency tracker. It is to be noted that homodyne operation does not permit the sense of the velocity to be indicated.

The continuous wave Doppler radar is subject to spurious effects introduced by reflection from nearby objects, turbulent air, precipitation, etc. Generally, fixed antenna systems are used with this type of Doppler radar. The direct leakage of signal from the transmitter to the receiver limits the maximum power that can be employed and hence satisfactory operation is limited to lower attitudes.

7.6.3 Frequency-Modulated Continuous-Wave Doppler Radar

Some of the disadvantages of the CW Doppler radar are overcome by the use of frequency modulation. In this system, the transmission is frequency modulated and the Doppler shift of one of the higher order side-band terms is measured. Consider a transmitted waveform of the type

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

where f_0 is the carrier frequency; f_m is the modulating frequency; and Δf, half the frequency excursion (maximum minus the minimum frequency).

This wave is returned, after a delay, with a change f_D in its frequecny. It can be shown[22] that the signals containing the difference frequency f_D are of the form:

$$\begin{aligned} v_0 = {} & \mathfrak{J}_0(D)\cos(2\pi f_D t - \phi_0) + 2\mathfrak{J}_1(D)\sin(2\pi f_D t - \phi_0)\cos(2\pi f_m t - \phi_m) \\ & - 2\mathfrak{J}_2(D)\cos(2\pi f_D t - \phi_0)\cos 2(2\pi f_m t - \phi_m) \\ & - 2\mathfrak{J}_3(D)\sin(2\pi f_D t - \phi_0)\cos 3(2\pi f_m t - \phi_m) \qquad 7.13 \end{aligned}$$

where $\mathfrak{J}_0$, $\mathfrak{J}_1$, $\mathfrak{J}_2$, etc. are Bessel functions of the first kind and order 0, 1, 2, etc.

$$D = \frac{\Delta f}{f_m} \sin 2\pi f_m \frac{R_0}{c}$$

R_0 = distance to the target at time $t = 0$

c = velocity of propagation

V_r = the velocity component along the beam

$$f_D = \frac{2V_r}{\lambda}$$

ϕ_o = phase shift corresponding to the delay $\dfrac{4\pi f_0 R_0}{c}$

ϕ_m = phase shift approximately equal to $\dfrac{2\pi f_m R_0}{c}$

This sinusoidal product terms give rise to a pair of frequencies of the type $nf_m \pm f_D$ by suppressed-carrier modulation of the nth order Bessel term nf_m by the Doppler shift f_D. The spectrum would appear as shown in Fig. 7.13.

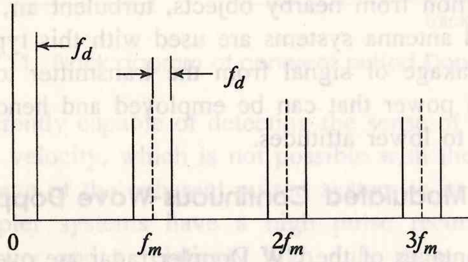


Fig. 7.13 Frequency spectrum of frequency modulated Doppler radar

The technique described above (i.e. taking the difference between the transmitted and received frequencies) is equivalent to homodyne reception and information regarding the sense of the velocity is lost. The Doppler frequency may be extracted from any of the side-bands by heterodyning with the corresponding harmonic of the modulating frequency. The amplitude of the Doppler signal depends on both the order of the Bessel side-band and the value of D. Bessel functions of the first kind and orders zero to 3 are shown in Fig. 7.14, as a function of D. The advantage of employing frequency modulation and choosing one of the higher order Bessel sidebands is evident from this. When the order of the Bessel sideband is 2 or higher, the corresponding amplitude of $\Im_n(D)$ is very low for low values of D. Since D is proportional to sin $2\pi f_m \frac{R_0}{c}$, this implies that direct feed from transmitter and returns from objects very near (low R_0) will give very low outputs and the disadvantage of CW Doppler is eliminated. However, this also implies that low altitude performance is unsatisfactory. In addition, FM-CW technique has the drawback

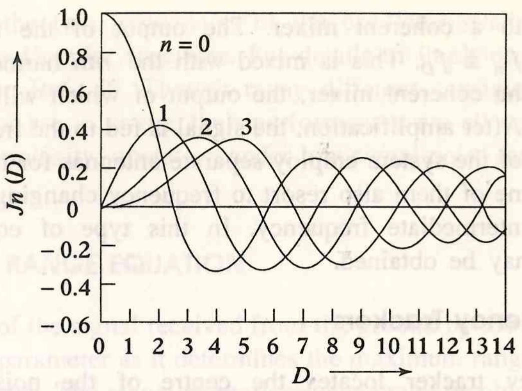


Fig. 7.14 Bessel function of orders 0 to 3

of 'altitude holes' in common with the pulsed system. The presence of altitude holes arises from the fact that when $2\pi f_m \frac{R_0}{c}$ is a multiple of π, say $m\pi$, D becomes zero. Since $R_0 = h$ cosec ψ where ψ is the depression angle, at certain altitudes given by $mc/2$ (f_m cosec ψ), the signals become zero. This disadvantage is overcome by either making f_m so low that even at the maximum usable altitude, the hole is not reached or by making f_m very high and wobbling the modulation frequency.

A block diagram of FM-CW Doppler radar is shown in Fig. 7.15. This shows a system using a common antenna for transmission and reception and homodyne reception. The received signal is mixed with sample of the transmitted signal in a balanced mixer and the desired side-band is filtered

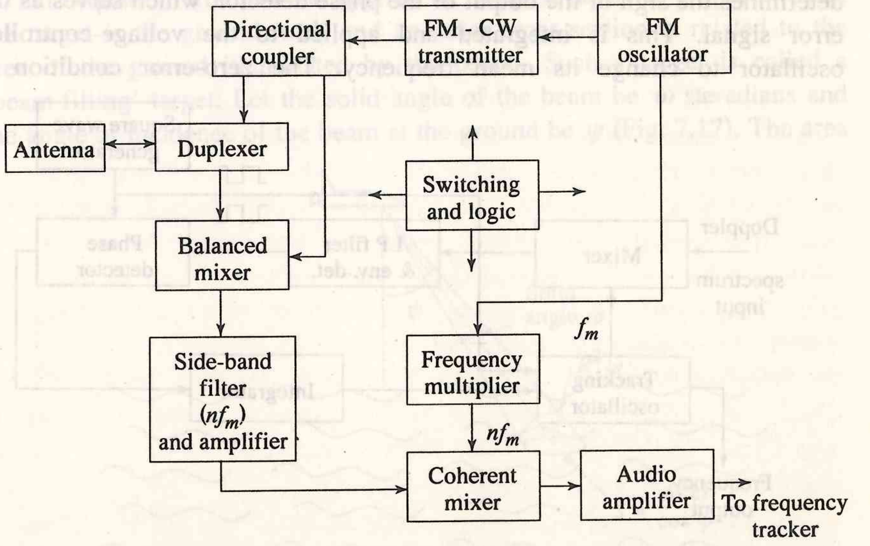


Fig. 7.15 Block diagram of FM-CW Doppler radar

and applied to a coherent mixer. The output of the filter will have frequencies $nf_m \pm f_D$. This is mixed with the nth harmonic of the FM oscillator in the coherent mixer, the output of which will then be at the frequency f_D. After amplification, the signal is fed to the frequency tracker.

Variations of the system employ separate antennas for transmission and reception. Some of them also resort to frequency changing and amplification at the intermediate frequency. In this type of equipment, sense information may be obtained.

7.6.4 Frequency Trackers

The frequency tracker locates the centre of the noise-like Doppler spectrum and gives as output the pure signal of this frequency. There are various configurations of the frequency tracker but most of them employ a tracking oscillator, the frequency of which is compared with the spectrum in a discriminator-type device and the error signal generated when the oscillator deviates from the centre of the spectrum is fed back to correct the frequency of the oscillator. One configuration (called the 'two-filter' tracker) is shown in Fig. 7.16. In this arrangement, actually a single filter is used but the oscillator frequency is switched by a square wave and takes on alternately two values which are separated by the spectrum width. The oscillator output is mixed with the Doppler signal and then passed through a low-pass filter and envelope detector. The output of the filter is a square wave which is applied to the phase detector, the other input to which is the output of the square wave oscillator. The phase of the first input in relation to the second (i.e. in-phase or 180° out of phase) determines the sign of the output of the phase detector, which serves as the error signal. This is integrated and applied to the voltage-controlled oscillator to change its mean frequency. The zero-error condition is obtained when the two frequencies of the oscillator straddle the centre frequency of the Doppler spectrum. For details of implementation of this arrangement, see Ref. 20. Though many different configurations of the frequency tracker are in vogue, their performances are all nearly the same in respect of complexity, operation under low signal-noise ratio conditions, etc.

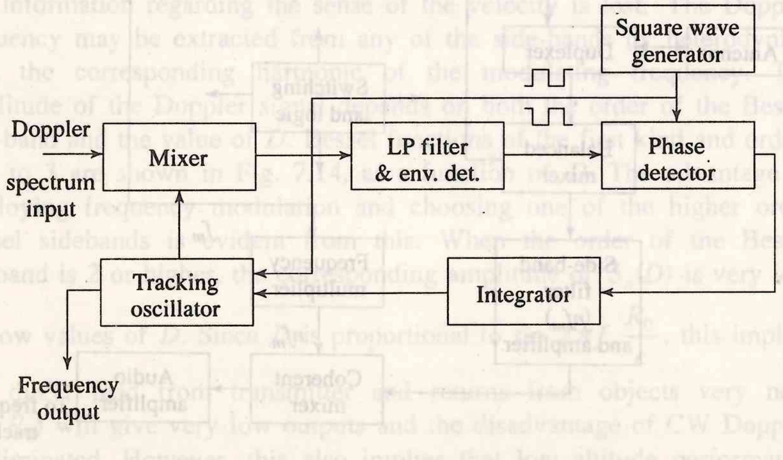


Fig. 7.16 Block diagram of a two-filter frequency tracker

7.7 DOPPLER RANGE EQUATION

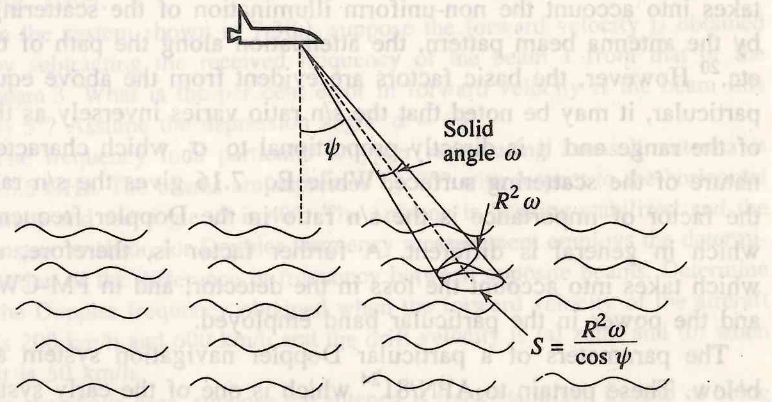
The magnitude of the signal received from the ground by the Doppler radar is an important parameter as it determines the maximum range at which the equipment operates. The signal strength received may be obtained from the radar range equation suitably modified to the situation pertaining in Doppler radar. In its simplest form, the radar range equation is given by [Ref. 23 Eq. 26-3]:

$$P_r = \frac{P_T \, G_T \, S \, A_e}{(4\pi R^2)^2} \tag{7.14}$$

where

P_r = received signal power
P_T = power of the transmitted signal
(the peak power in pulsed radar systems)
G_T = gain of the transmitting antenna
S = radar cross-section of the target
A_e = effective area of the receiving antenna and
R = distance of the target

This equation will now be applied to Doppler radar. In Doppler radar, the target is the ground itself and its radar cross-section is related to the area of the ground illuminated by the beam. Such a target is called a 'beam-filling' target. Let the solid angle of the beam be ω steradians and the angle of incidence of the beam at the ground be ψ (Fig. 7.17). The area



of the ground illuminated by the beam is then $\frac{R^2\omega}{\cos\psi}$. The scattering cross-section of the target is proportional to this area. If we denote by σ_0 the scattering cross-section per unit area illuminated, the factor S in the range Eq. 7.14 is replaced by $\frac{\sigma_0 R^2\omega}{\cos\psi}$. The effective area of the antenna A_e is related to the antenna gain by the relation

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

where λ is the wavelength of the radiation. Noting that the gain of an antenna is related to the solid angle of the beam by the relation [Ref. 22 Eq. 7.4] $G = \frac{4\pi}{\omega}$, the effective area reduces to $A_e = \frac{\lambda^2}{\omega}$. Substituting these in the range equation,

$$P_r = \frac{P_T G_T}{16\,\pi^2 R^4}\cdot\frac{\sigma_0\,\omega R^2}{\cos\psi}\,\frac{\lambda^2}{\omega}$$

$$= \frac{P_T G_T \lambda^2 \sigma_0}{16\pi^2 R^2 \cos\psi} \qquad 7.15$$

the signal/noise ratio at the IF [(s/n)/IF] is obtained by dividing the received signal power by the available noise power $k T_0\, B_{\text{IF}}\, F$, where B_{IF} is the IF band-width, F is the noise figure of the receiver, k the Boltzmann constant and T_0 the standard temperature of 290 °K. Substituting

$$(s/n)_{\text{IF}} = \frac{P_T G_T \lambda^2 \sigma_0}{16\pi^2 R^2\, k\, T_0\, B_{\text{IF}}\, F \cos\psi} \qquad 7.16$$

This equation could be further modified by including a factor which takes into account the non-uniform illumination of the scattering surface by the antenna beam pattern, the attenuation along the path of the beam, etc.[20] However, the basic factors are evident from the above equation. In particular, it may be noted that the s/n ratio varies inversely as the square of the range and it is directly proportional to σ_0 which characterizes the nature of the scattering surface. While Eq. 7.16 gives the s/n ratio at IF, the factor of importance is the s/n ratio in the Doppler frequency band, which in general is different. A further factor is, therefore, necessary which takes into account the loss in the detector, and in FM-CW systems and the power in the particular band employed.

The parameters of a particular Doppler navigation system are given below. These pertain to APN/81[24] which is one of the early system to be put to use. It is a pulsed system using four beam Janus antenna, with beam widths of 3.5° in the narrow direction and 35° in the wide direction. As in many other navigation systems, there are two modes of operation—a normal mode when the received signals are of adequate strength and coherence, and a memory mode when these conditions do not exist.

The frequency of operation is between 8.7 and 8.9 GHz and pulse repetition frequency is 50 kHz, with a pulse width having a mean value of 0.9 μsec. The high value of the prf enables unambiguous determination of Doppler shift with aircraft speeds up to 700 knots (approximately 1300 km/hr). The peak power of the transmitter is 1100 W and the average power (nominal) is 50 W. The equipment can operator up to an altitude of 20,000 m.

7.8 ACCURACY OF DOPPLER NAVIGATION SYSTEMS

The overall Doppler navigation system accuracy depends upon the accuracy of the measurement of ground speed and the heading accuracy. Computational error may be a factor of importance if analog computers are used, but the present tendency is to use digital computation, in which the computational errors can be reduced to negligible proportions. The earlier systems had an overall accuracy of position better than 1% of the distance travelled over fairly long distances. It is expected that with improved heading references and negligible computational error, this may be reduced to 0.25% of the distance travelled.

QUESTIONS AND PROBLEMS

1. Calculate the Doppler shift in each of the beams in a Doppler radar of the Janus-X type shown in Fig. 7.3(d) given $\alpha_o = 60°$, $\theta_o = 45°$, frequency of operation 9 GHz and assuming that the aircraft is on level flight with a forward velocity of 500 km/h and a drift velocity of 100 km/h.
2. Develop equations similar to 7.3 to 7.6 for the Janus-λ system of Fig. 7.3(c).
3. In the system shown in 7.3(b), suppose the forward velocity is obtained by subtracting the received frequency of the beam 1 from that of the beam 3. What is the per cent error in forward velocity if the beam tilts by 5°? Assume the depression angle $\alpha = 60°$.
4. The frequency in a particular Doppler radar using Janus-X antenna is 13.5 GHz. The beams are depressed by 70° with respect to the horizontal plane and the angle θ is 40°. The antenna is heading-stabilized and the instrumentation for Doppler frequency measurement employs the determination of the difference in frequency between opposite beams. Determine the Doppler frequency obtained when the forward velocity of the aircraft is 200 km/h and 600 km/h and the drift velocity is (a) zero, and (b) when it is 50 km/h.
5. In the above case, suppose the antenna is track-stabilized instead of being [illegible]. What would be the Doppler shifts ?

8

Inertial Navigation

Inertial navigation is a system of dead-reckoning navigation in which the instruments in the craft determine its acceleration and by successive integration, obtain its velocity and displacement. The system is entirely self-contained and can be used both on earth, under the sea and in space. Its earliest use appears to have been in nuclear-powered submarines. The earlier equipment was bulky and heavy but subsequent developments have reduced the size and weight to such an extent that it is now being used in some aircraft. The principal advantages of the inertial system are its self-contained nature, immunity from jamming, absence of any radiation and the fact that it is useable at all latitudes. It has a high degree of accuracy, though, like all dead reckoning systems, its accuracy decreases with time.

8.1 PRINCIPLES OF OPERATION

The essential elements of the inertial navigator are *accelerometers* which determine the acceleration of the craft and a set of *gyroscopes* which maintain the directions of these accelerations along the desired coordinates. In addition, a computer is required to integrate the acceleration with respect to time, to obtain the velocity of the craft and the distance it has travelled along the chosen coordinates.

For simplicity, let us consider first navigation in space free from gravitational field. Let x_1, x_2 and x_3 represent the positions of the vehicle in a set of orthogonal cartesian coordinates fixed with reference to the 'fixed stars'. Such a set of coordinates is called an inertial frame of reference. The three components of acceleration are then $\ddot{x}_1$, $\ddot{x}_2$ and $\ddot{x}_3$ (where each dot represents a differentiation with respect to time). If these accelerations are determined, the position of the vehicle at any time t can

$$\left.\begin{aligned} \text{velocity} &= \dot{x}_1(t) = \int_0^t \ddot{x}(t)\,dt + \dot{x}_1(0) \\ \text{displacement } x_1(t) &= \int_0^t \dot{x}_1(t)\,dt + x(0) \\ &= \int_0^t \left[\int_0^t \ddot{x}_1(t)\,dt\right] dt + \dot{x}_1(0)t + x_1(0) \end{aligned}\right\} \qquad 8.1$$

where $\dot{x}_1(0)$ and $x_1(0)$ are respectively the initial velocity and position along the x_1 axis. In a similar manner, $x_2(t)$ and $x_3(t)$ can be determined. Note that the above expressions imply that the initial velocity and position are known.

The presence of a gravitational field changes the situation. The vehicle then accelerates under the influence of both the gravitational force and its own thrust. Accelerometers (Sec. 8.3.1) do not respond to gravitational forces as these act both on the case of the accelerometers and on the 'proof mass' inside them. The thrust of the vehicle acts on the case of the accelerometer, which is attached to the body of the vehicle, but not on the proof mass and consequently the accelerometers respond to the thrust acceleration. However, if the gravitational acceleration is known, it may be taken into account and the true acceleration of the vehicle in the inertial frame of reference may be determined. Equations 8.1 then become applicable. In general, it can be assumed that the gravitational acceleration can be computed if the position of the vehicle in the inertial frame of reference is known. The block diagram in Fig. 8.1 illustrates the inertial

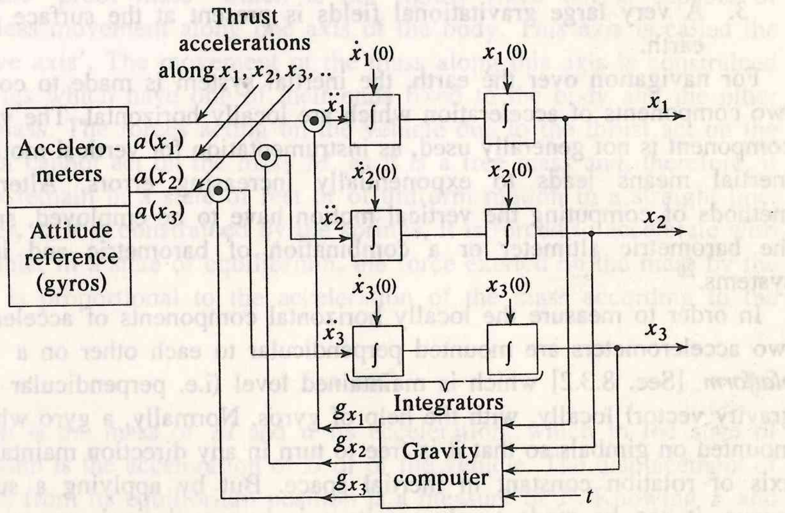


Fig. 8.1 Block diagram of the general inertial navigation system

navigation system in its most general form. The input to the integrators in this case is the sum of the thrust acceleration a_x and the gravitational acceleration g_x along the particular axis. The components of gravitational acceleration are, in turn, computed from the position of the vehicle and time. This last factor is to take account of the fact that the gravitational field at any point changes with time. The 'attitude reference' consists of a set of gyros which maintain the orientation of the accelerometers along the required axes. This general block diagram may be modified to take into account the requirements of particular situations.

8.2 NAVIGATION OVER THE EARTH

We are interested here principally in navigation of a craft over the earth either on the surface of the earth as in the case of sea-going vessels or close to the earth as in the case of aircraft. Navigation over the earth presents the following features:

1. The system of coordinates should be fixed with reference to earth. As earth is rotating, the system of axes fixed with respect to earth is a non-inertial system. (The orbital motion of the earth does not make it a non-inertial system, as the orbital motion is one of 'free fall' under the influence of gravitational forces. A system of axes centred at the earth and fixed with reference to the fixed stars, is an inertial system of axes.)
2. The coordinate system most convenient for use is latitude and longitude, which constitute a set of curvilinear coordinates on a spheroidal surface.
3. A very large gravitational fields is present at the surface of the earth.

For navigation over the earth, the inertial system is made to compute two components of acceleration which are locally horizontal. The vertical component is not generally used, as instrumentation for vertical motion by inertial means leads to exponentially increasing errors. Alternative methods of computing the vertical motion have to be employed, such as the barometric altimeter or a combination of barometric and inertial systems.[20]

In order to measure the locally horizontal components of acceleration, two accelerometers are mounted perpendicular to each other on a *stable platform* [Sec. 8.3.2] which is maintained level (i.e. perpendicular to the gravity vector) locally, with the help of gyros. Normally, a gyro which is mounted on gimbals so that it is free to turn in any direction maintains its axis of rotation constant in inertial space. But by applying a suitable torque, it can be made to change its axis in any desired manner [Sec. 8.3.2]. The gyros are mounted on the stable platform and the orientation of the latter is controlled by torquing signals applied to the gyros. In this manner, the platform may be maintained locally level. By integration of the components of acceleration, the velocity and displacement along the desired coordinates are obtained. Before going into the details of the method, the components making up the system will be described.

8.3 COMPONENTS OF AN INERTIAL NAVIGATION SYSTEM

8.3.1 Accelerometers

These instruments measure the thrust acceleration of the vehicle. Various types of accelerometers have been developed for inertial navigation. To illustrate the principle of operation of accelerometers in general, the simple arrangement shown in Fig. 8.2 will be considered. The accelerometer consists essentially of the body B, attached to the vehicle, and mass M called the 'proof mass' which is not rigidly fixed but is capable of frictionless movement along one axis of the body. This axis is called the 'sensitive axis'. The movement of the mass along this axis is constrained by springs which have one of their ends fixed to the body and the other to the mass. The forces acting on the vehicle due to the thrust act on the body but cannot act on the mass M, as it is a free mass and, therefore, it tends to remain in a state of rest or of uniform motion in a straight line. However, as it is constrained by the springs, it is forced to accelerate with the vehicle. In a state of equilibrium, the force exerted on the mass by the springs is proportional to the acceleration of the mass according to the relation

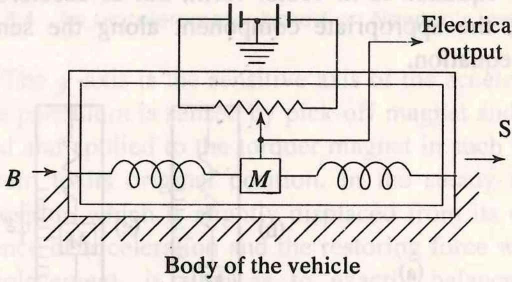


Fig. 8.2 A simple accelerometer

$$F = ma \qquad 8.2$$

where m is the mass of M and a its acceleration, which in the state of equilibrium is the acceleration of B or of the vehicle. The displacement of the mass from its equilibrium position is a measure of F. Knowing F and m, the acceleration can be determined. Figure 8.2 shows an arrangement for obtaining an electrical output proportional to the displacement. Note that an acceleration to the right results in a displacement of the mass to the left within the body B and *vice versa*. Ideally, accelerometers should be insensitive to accelerations orthogonal to the sensitive axis.

Consider such an accelerometer in a state of rest on a table. If the sensitive axis is horizontal, no acceleration is indicated (Fig. 8.3). The component of g along the sensitive axis is then zero. If allowed to stand vertically, the mass moves down under the action of gravity and an upward acceleration $(-g)$ is indicated. If allowed to fall freely, the indicated acceleration is zero, while if accelerated upwards with an acceleration of magnitude $2g$, the indication is $3g$. This illustrates the rule that the indicated acceleration is $a-g$ where a represents the true acceleration and g the gravitational acceleration. In vector form $f = a - g$, where f is the indicated acceleration. To obtain the true acceleration, the gravitational acceleration is added to the indicated acceleration, as has been shown in Fig. 8.1. The above equation is in vector form, but as accelerometers are single axis devices, the appropriate component along the sensitive axis satisfies the above equation.

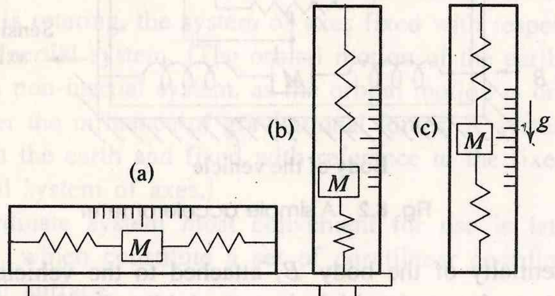


Fig. 8.3 Effect of gravity on the accelerometer output [(a) resting on table horizontal, (b) vertical, and (c) free fall]

The accelerometer discussed above is not a practical one and is given for purposes of illustration. But many accelerometers in use make use of Newton's law $F = ma$ in some form. A common variant is Newton's law in its polar form

$$T = I\ddot{\theta} \tag{8.3}$$

where T is the torque, I the moment of inertia and $\ddot{\theta}$ the angular acceleration of the body. In addition, in practice, the moving body is not allowed to move but is restored to its normal position by a feed-back system which causes a force (or a torque) to be exerted on the mass electromagnetically. The magnitude of this restoring force is then a measure of the acceleration. One such system[20] is illustrated in Fig. 8.4. Here the proof mass is provided by the pendulum P pivoted to move along

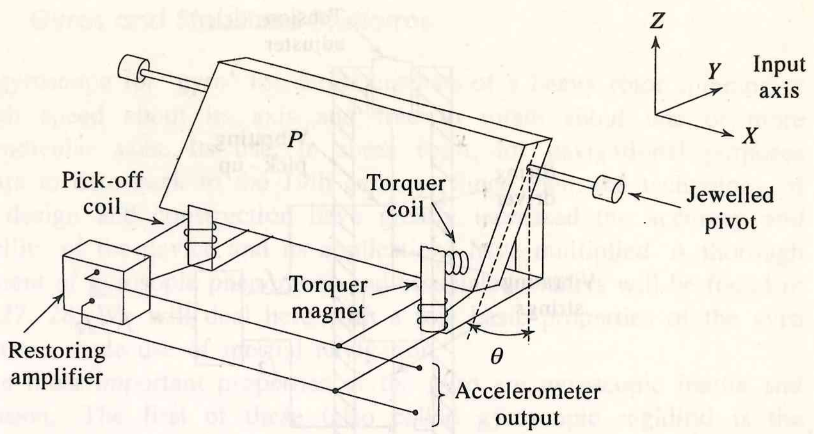


Fig. 8.4 An accelerometer based on Newton's law in polar form

the x-axis. The y-axis is the sensitive axis of the accelerometer. The movement of the pendulum is sensed by pick-off magnet and the resulting signal is amplified and applied to the torquer magnet in such a sense as to restore the pendulum to its original position. In the steady state, the pendulum attains a position which is slightly displaced from its equilibrium position in the absence of acceleration and the restoring force which is proportional to the displacement, is such as to exactly balance the force due to acceleration. This relation may be put in the form

$$k\theta = mbf_y \tag{8.4}$$

where

mb = pendulosity of the moving part (g-cm)
f_y = the component of the acceleration along the y-axis (or sensitive axis)
θ = the small displacement of the pendulum and
k = a constant which takes into account the servo-amplifier gain and any spring stiffness that may be present.

The restoring amplifier output gives $k\theta$ and f_y may, therefore, be obtained. The above equation ignores all complicating factors such as the residual torque due to connecting wires, the effect of cross-axis acceleration, etc.

Another type of accelerometer employs two vibrating strings with a proof mass at the centre. In the absence of acceleration, the tensions of the two strings are equal and their frequencies of oscillation are the same (Fig. 8.5). The frequency of oscillation of each string is proportional to the square root of the tension of the spring. In the presence of an acceleration along the sensitive axis (which is coincident with the string) the tension on one is reduced and on the other is increased by the same amount. If the initial tension is T_0, the two frequencies in the presence of an acceleration

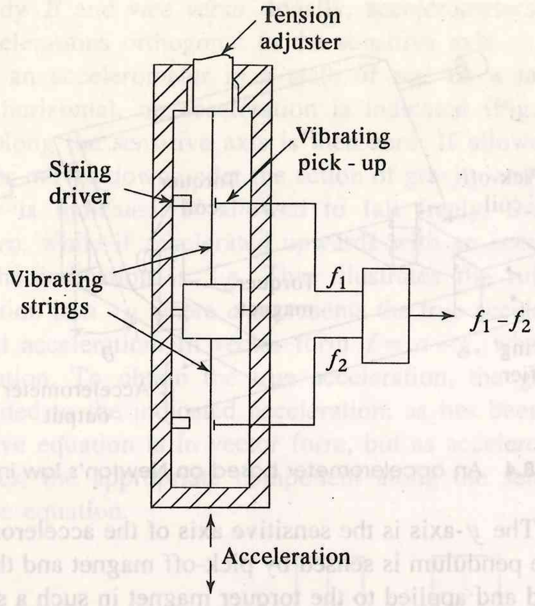


Fig. 8.5 Vibrating string accelerometer

frequency of the 1st string = $f_1 = k_1\sqrt{\mathcal{T}_0 + \Delta\mathcal{T}}$

frequency of the 2nd string = $f_2 = k_1\sqrt{\mathcal{T}_0 - \Delta\mathcal{T}}$

$$\text{and } f_1 - f_2 = k_1\sqrt{\mathcal{T}_0}\left[\frac{\Delta\mathcal{T}}{\mathcal{T}_0} + \frac{1}{8}\left(\frac{\Delta\mathcal{T}}{\mathcal{T}_0}\right)^3 + \dots\right] \qquad 8.5$$

where $\Delta\mathcal{T} \propto ma$ = proof mass × axial acceleration.
If $\mathcal{T}_0$ is large compared with the changes in tension $\Delta\mathcal{T}$ brought about by the acceleration, the difference frequency is proportional to the acceleration. If a higher accuracy is required, the other terms in the bracket can be used as correction terms in the computation of velocity and displacement. The design of these accelerometers must ensure that the support of the proof mass is such that only its acceleration along the sensitive axis affects the tension of the strings. Vibrating string accelerometers tend to exhibit sensitivity to cross-axis acceleration and to vibration. The main advantage of these devices is that the integration of acceleration is easily implemented by a counter which counts the number of cycles of the difference frequency in a given time.[26]

A third type of accelerometer employs a gyroscope with an unbalance. The precession of such a gyro is a measure of the acceleration. The actual rotation is a measure of the integral of acceleration, i.e. the velocity. This type of gyro is called the unbalanced integrating gyro. For details of this type of accelerometer Ref. 25 may be seen.

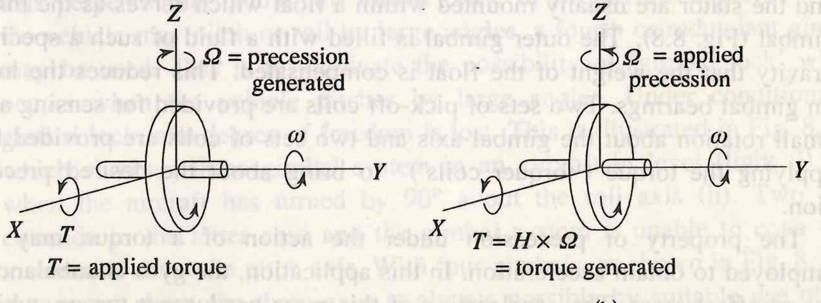
8.3.2 Gyros and Stabilized Platforms

The gyroscope (or 'gyro' for short) consists of a heavy rotor spinning at a high speed about its axis and free to rotate about one or more perpendicular axes. Its use, in some form, for navigational purposes appears to date back to the 19th century. Since then, the technology of gyro design and construction have greatly increased the accuracy and reliability of the device and its applications have multiplied. A thorough treatment of gyroscopic phenomena and associated matters will be found in Ref. 27, 28. We will deal here with a few basic properties of the gyro which are made use of inertial navigation.

The most important properties of the gyro are gyroscopic inertia and precession. The first of these (also called gyroscopic regidity) is the property by virtue of which the gyro tends to maintain its axis of rotation in a fixed direction in inertial space and resists any change. Precession is the property by virtue of which the plane of spin of the gyro changes at a constant rate when a torque is applied. The following vector relation holds between the rate of precession and applied torque:

$$\Omega = \frac{H \times \mathcal{T}}{|H|^2} \qquad 8.6$$

where Ω is the angular velocity of precession, H, the angular momentum of the body, (= moment of inertia I × spin rate ω_s), and T, the applied torque. This relation is illustrated in Fig. 8.6, where the torque and spin are orthogonal. When the spin is along the Y-axis, and a torque is along the X-axis, the gyro precesses about the Z-axis, as shown in Fig. 8.6 (a). (The precession is in a direction which tends to make the spin axis coincide with the torque). Conversely, when the gyro is made to precess in the Z-axis, this causes a torquing force to appear along the X-axis [Fig. 8.6(b)]. With a constant applied torque, the precession is constant or the axis of spin rotates at a constant rate. This is in contrast to what happens in the case of a non-spinning mass, where the application of a torque would make the body rotate at a linearly increasing rate:



(a) (b)

The gyro is mounted on gimbals, as shown in Fig. 8.7 so that its spin axis may be rotated about one or two other axes. The figure shows a gyro which is capable of rotation about two axes. The mass spins about the Z-axis on bearings fixed to the inner gimbal, the inner gimbal itself can rotate about the Y-axis on bearings fixed to the outer gimbal, which in turn can rotate about the X-axis on bearings fixed to the frame. Such a gyro is called a two-degree-of-freedom gyro. If the outer gimbal is fixed to the frame, only rotation about the Y-axis is possible and the gyro is called a single-degree-of-freedom gyro. This can sense the rotation of the frame along one axis. Three such gyros are required to sense the rotation along three axes. In navigational applications, the property of gyroscopic rigidity is not used directly, but small changes in the orientation of the gyro axis in relation to the gimbals is electrically sensed and a feed-back arrangement is used to precess the gyro in some desired manner. In this way, the gyro may be constrained to take up certain fixed orientation, say, in earth coordinates rather than in inertial coordinates.

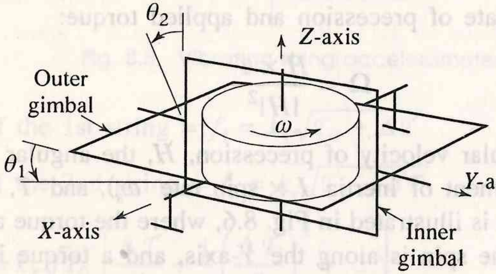


Fig. 8.7 Two-degree-of-freedom gyro

The modern gyro used for navigational purposes is a device of very high precision. Its spinning mass is electrically driven and is the rotor of an electric motor. To obtain a high degree of gyroscopic rigidity, the moment of inertia and angular velocity have to be high. The rotor is designed to have most of the mass concentrated near the periphery and speeds of rotation may be as high as 30,000 rev/min. The older gyros operated with ball bearings but later ones have self-lubricated gas bearings.[20] The rotor and the stator are usually mounted within a float which serves as the inner gimbal (Fig. 8.8). The outer gimbal is filled with a fluid of such a specific gravity that the weight of the float is compensated. This reduces the load on gimbal bearings. Two sets of pick-off coils are provided for sensing any small rotation about the gimbal axis and two sets of coils are provided for applying the torque ('torquer coils') to bring about the desired precession.

The property of precession under the action of a torque may be employed to obtain acceleration. In this application, the gyro is unbalanced by an offset mass. The acceleration of this mass produces a torque, which

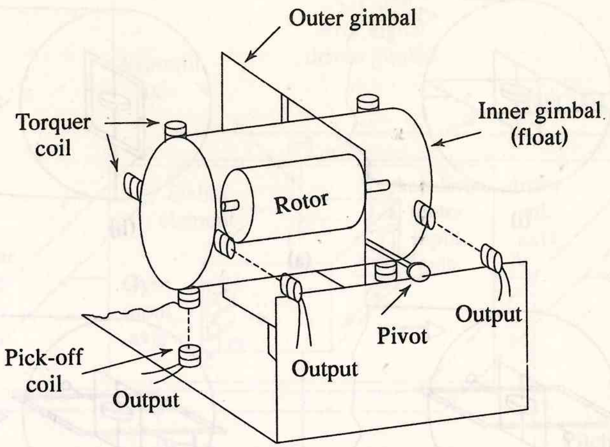


Fig. 8.8 Sketch of a gyro used in inertial navigation

The precessional rate (Ω) is then a measure of the acceleration and the angle precessed is proportional to the integral of acceleration, or velocity. Such gyros are called *integrating gyro accelerometers*. These are not generally used for navigation but find application in ballistic missiles.

Stable platforms serve the purpose of isolating the inertial sensing elements such as the accelerometers, from vehicle motion. The stabilized platform is mounted on gimbals to permit rotation of the vehicle with respect to the platform along the three axes. The platform is maintained at the desired orientation by gimbal servos which are actuated by error signals produced by gyros whenever there is a deviation of the orientation from the desired one. The three possible angular errors may be sensed either by two two-degree-of-freedom gyros or three single-degree-of-freedom gyros.

The gimbal system may be external of internal to the stable mount (or stable element). In the former, the stable element is surrounded by an inner gimbal to which the stable element is pivoted along one axis, the inner gimbal is pivoted to the second one along an axis at right angles to the first axis, and the second gimbal is pivoted to the third one along an axis perpendicular to the two former ones. In aircraft installations, where the vehicle may pitch or roll by large angles, a fourth or redundant gimbal may be used. This is to eliminate the possibility of 'gimbal lock' which occurs when the vehicle pitches by large angles. Under conditions of gimbal lock, one degree of freedom is lost. This is illustrated in Fig. 8.9(a) which shows a three gimbal system in an aircraft in level flight (i) and when the aircraft has turned by 90° about the roll axis (ii). Two axes coincide in the latter case and the gimbal system is unable to cope with movement about the pitch axis. With four gimbals, as shown in Fig. 8.9(b), movement about the three axes is always possible by suitable use of the

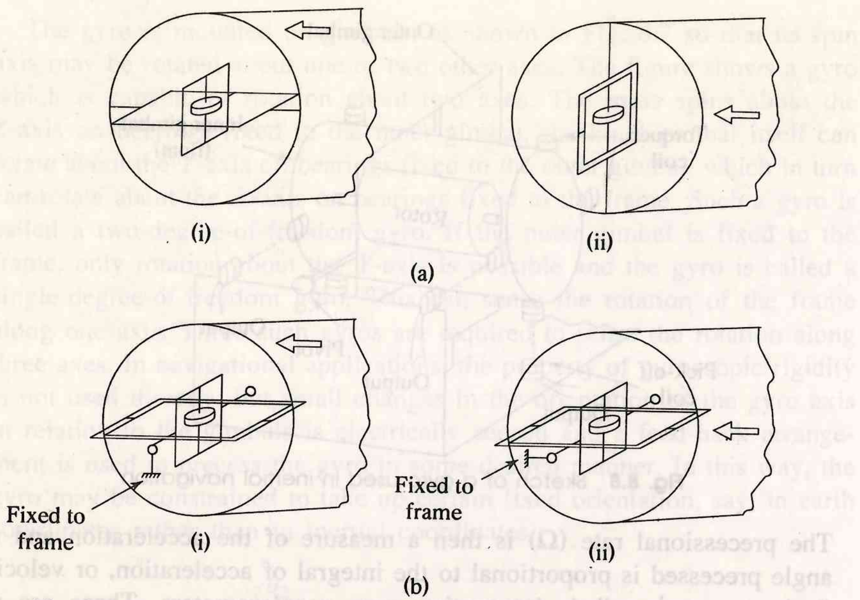


Fig. 8.9 Gimbal lock [(three-gimbal mount, (b) four-gimbal mount]

position (i) and with near 90° roll angle (ii). When the roll exceeds a certain amount, the fourth gimbal motor takes over and turns the second gimbal to a level position. Thus, the first two gimbals are maintained orthogonal. For further details of the fourth gimbal operation, see Ref. 20, 29. The four gimbal system is invariably used in aircraft when pitch and roll angles approaching 90° are likely during maneuvers. In ships and passenger aircraft, where this possibility does not exist, a three-gimbal system may be used. The third gimbal in the latter is replaced by the mounting structure for the platform in the carrying vehicle.

A four-axis stable platform is shown diagrammatically in Fig. 8.10. The inertial sensors and gyros are mounted on the stable element. The inner gimbal has an axis which is usually vertical. The second gimbal has an axis coinciding with the roll axis, the third gimbal axis is the pitch axis and the fourth axis again coincides with the roll axis. The platform and each of the gimbals are driven by a servomotor about one of the axes. The servomotors are actuated by the error signals obtained from the gyros, as shown for the vertical axis. The vehicle attitude data is also obtained for the three axes by synchros, resolvers or digital encoders.

8.4 EARTH COORDINATE MECHANIZATION

As stated earlier, navigation over the earth's surface require basically the measurement of two locally horizontal components of acceleration. For this purpose, the accelerometers are mounted orthogonally on a stable platform which is maintained locally level. In one type of mechanization,

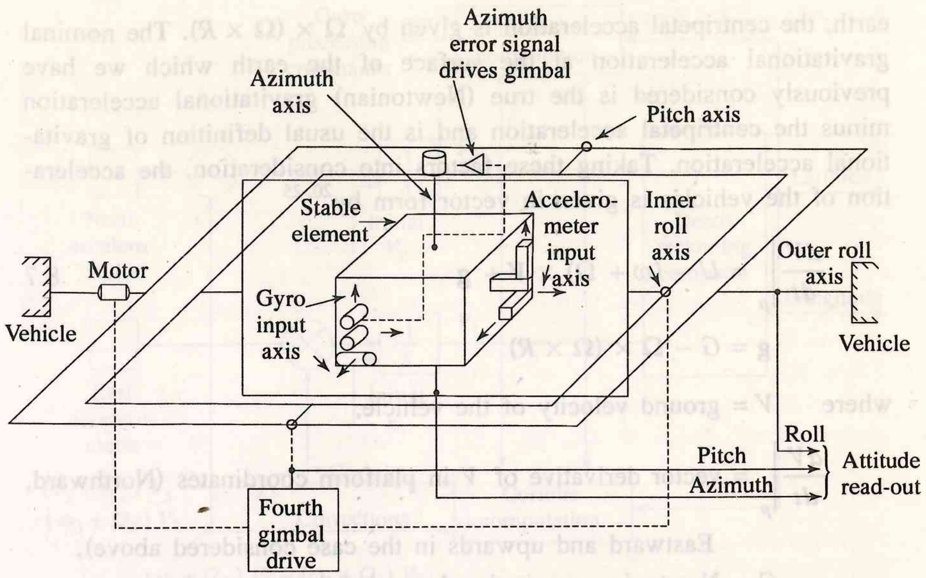


Fig. 8.10 A four-axis stable platform

North axis and the East axis, and the accelerations are integrated to determine the velocities and displacements along these coordinates. The latitude and longitude are then obtained directly. The gyros are precessed continuously to maintain the platform level and to keep the accelerometers pointing North and East. The signals necessary for precessing the gyros are computed, in turn, from position and velocity information.

The movement of a vehicle on a rotating earth gives rise to some fictitious (or apparent) accelerations called Coriolis accelerations which have to be allowed for to obtain the true acceleration of the vehicle in earth coordinates. In addition to this, centripetal acceleration is also present. If R is the vector representing the position of the object with respect to the centre of the earth (Fig. 8.11) as origin, and Ω the angular velocity of the

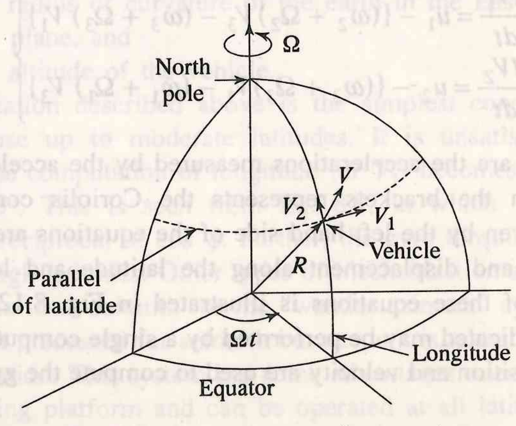


Fig. 8.11 The earth coordinate system

earth, the centripetal acceleration is given by $\Omega \times (\Omega \times R)$. The nominal gravitational acceleration at the surface of the earth which we have previously considered is the true (Newtonian) gravitational acceleration minus the centripetal acceleration and is the usual definition of gravitational acceleration. Taking these factors into consideration, the acceleration of the vehicles is given in vector form by[20, 25]

$$\left.\frac{dV}{dt}\right|_p = U - (\omega + \Omega) \times V + \mathbf{g} \qquad 8.7$$

$$\mathbf{g} = G - \Omega \times (\Omega \times R)$$

where V = ground velocity of the vehicle,

$\left.\frac{dV}{dt}\right|_p$ = vector derivative of V in platform coordinates (Northward, Eastward and upwards in the case considered above),

G = Newtonian gravitational acceleration,

U = acceleration vector measured by the accelerometers,

ω = inertial angular velocity of the platform

Ω = inertial angular velocity of the earth, and

g = gravitational acceleration (with a locally horizontal platform, this is normal to the platform and has no components along the latitude and longitude).

The vector equation may be reduced to three scalar equations, of which the two relating to the motion of vehicle in the horizontal plane are usually solved for navigational purposes. Indicating by subscripts 1, 2 and 3, the components along the East, North and upward axes, these two relevant equations take the following form:

$$\left.\begin{aligned}\frac{dV_1}{dt} &= u_1 - \{(\omega_2 + \Omega_2)V_3 - (\omega_3 + \Omega_3)V_1\}\\ \frac{dV_2}{dt} &= u_2 - \{(\omega_3 + \Omega_3)V_1 - (\omega_1 + \Omega_1)V_3\}\end{aligned}\right\} \qquad 8.8$$

Here, u_1 and u_2 are the accelerations measured by the accelerometers and the term within the brackets represents the Coriolis corrections. The accelerations given by the left hand side of the equations are integrated to obtain velocity and displacement along the latitude and longitude. The mechanization of these equations is illustrated in Fig. 8.12. The various computations indicated may be performed by a single computer in practice. Note that the position and velocity are used to compute the gyro precession commands.

The various quantities involved in the computations are as follo

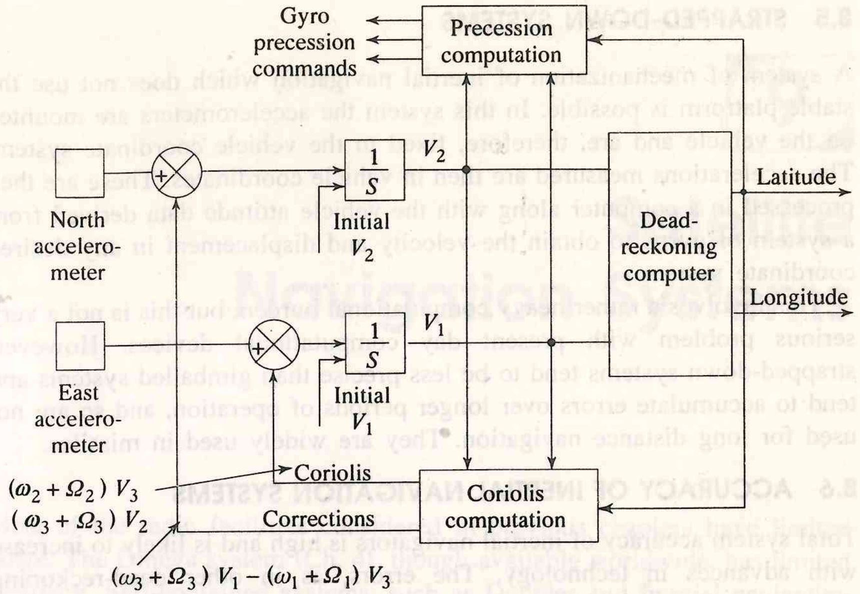


Fig. 8.12 Mechanization for navigation in the latitude longitude system of coordinates

$$\left.\begin{aligned}\Omega_1 &= 0 & \omega_1 &= \Omega_1 - \frac{V_2}{R_M + h}\\ \Omega_2 &= \Omega\cos\phi & \omega_2 &= \Omega_2 + \frac{V_1}{(R_P + h)\cos\phi}\\ \Omega_3 &= \Omega\sin\phi & \omega_3 &= \omega_2\sin\phi\end{aligned}\right\} \qquad 8.9$$

where

ϕ = latitude of the vehicle,

R_M = radius of curvature of the earth in the meridian plane,

R_p = radius of curvature of the earth in the East-West vertical plane, and

h = altitude of the vehicle.

The mechanization described above is the simplest conceptually and is suitable for use up to moderate latitudes. It is unsatisfactory at high latitudes, as the computation of longitude (or V_1) becomes inaccurate as ϕ approaches 90°. This is seen from Eq. 8.9, in which ω_2 has a term involving the reciprocal of cos ϕ. Further, the gyro torquing rates become excessive at high latitudes. Other types of mechanization are possible, one of which is the 'free-azimuth' or the "wander azimuth" system where the platform is not precessed for azimuth but account is taken of wander angle in the computation. This system has some advantages over the system with a North-pointing platform and can be operated at all latitudes.[20]

8.5 STRAPPED-DOWN SYSTEMS

A system of mechanization of inertial navigation which does not use the stable platform is possible. In this system the accelerometers are mounted on the vehicle and are, therefore, fixed to the vehicle coordinate system. The accelerations measured are then in vehicle coordinates. These are then processed in a computer along with the vehicle attitude data derived from a system of gyros to obtain the velocity and displacement in any desired coordinate system.

This involves a rather heavy computational burden, but this is not a very serious problem with present day computational devices. However, strapped-down systems tend to be less precise than gimballed systems and tend to accumulate errors over longer periods of operation, and so are not used for long distance navigation. They are widely used in missiles.

8.6 ACCURACY OF INERTIAL NAVIGATION SYSTEMS

Total system accuracy of inertial navigators is high and is likely to increase with advances in technology. The errors, as in other dead-reckoning systems, increase with time. They arise from various causes such as gyro-drift, accelerometer errors, linearity and scale factor errors of the gyro-torquer, computer errors and initial condition errors. Some errors are oscillatory and have periods of the earth rate Ω, the Schüler frequency $\omega_s = \sqrt{g/R}$ (84.4 minute period) and a frequency $2\ \Omega \sin \phi$. These frequencies arise out of the nature of the system mechanization.

Some reported values of error are in the range or 1 of 1.5 km for a flight period of 4 to 6 hr. Greater accuracies may still be expected, but there are some fundamental limits to the accuracy of an inertial navigation system. These arise mainly from the uncertainty of the gravitational field in the region of operation and angular errors arising from the precession of equinoxes and the migration of earth's poles.

QUESTIONS AND PROBLEMS

1. A gyro rotor has a mass of 0.3 kg and a radius of gyration of 1.5 cm. It rotates at a speed of 12000 rpm. Find at what rate it precesses when a torque of 1 kg/m is applied perpendicular to the spin axis?
2. Derive Eq. 8.8 from 8.7 and justify the expressions given in Eq. 8.9.
3. In the vibrating string accelerometer of Fig. 8.5, the axial acceleration is nearly proportional to the difference in frequenc, $f_1 - f_2$ (Eq. 8.5). If the acceleration is to be measured with an accuracy of 0.1%, what is the maximum value of the proportional change in frequency?
4. Explain the phenomenon of "gimbal lock" with an illustration. Why is the four gimbal mount necessary in aircraft?
5. For navigation over the earth, the stable platform has to be maintained locally level with the accelerometers measuring the accelerations along two coordinates, generally latitude and longitude. Discuss the mechanisa-

---

9

Satellite Navigation Systems

Most of the radio facilities considered in previous chapters have limited range. The Omega system (Ch. 4), though available worldwide, has limited accuracy. Self-contained systems, such as Doppler and Inertial navigation systems, though useable world-wide, have accuracies which degrade with time and require periodic updating. Celestial navigation using Star-trackers are suitable only for high flying aircraft. Two developments which have come into being later have overcome these drawbacks by using artificial satellites for navigation. These are dealt with in this chapter.

9.1 THE TRANSIT SYSTEM

The idea of using artificial satellites for navigation was conceived soon after the launch of Soviet satellite, Sputnik. In 1957, American scientists at the Applied Physics Laboratory demonstrated that they could establish the ephemeris* of the satellite by carefully measuring the Doppler shift of its transmission. Later it was realised that the reverse was also true, i.e., if the ephemeris of the satellite was known, a receiver on earth could determine its position by measuring the Doppler shift in the received frequency of a stable transmitter in the satellite. This was the beginning of the use of satellites for navigation. The first application of this method was by the US Navy for the navigation of their missile carrying ships. The Navy Navigation Satellite System (NNSS), also known as the "Transit System", came into operation in 1967. It was originally used exclusively by the US Navy but the system was released for general use in 1977. The system is suitable for slow moving vehicles.

The system consists of six satellites and some spares in polar orbits around the earth. The orbits are elliptic but the eccentricity is so small that

---

* The ephemeris consists of a set of 16 constants from which the exact position

they may be considered essentially circular. The period of the orbit is $1\frac{3}{4}$ hour, so that each satellites completes about 13.5 orbits in a day. In each orbit a satellite can be received for only 10 to 15 mins. by a ship, depending on its longitudinal separation from the receiver. All the satellites are provided with two very stable transmitters, operating at 400 MHz and 150 MHz. A receiver on the surface of the earth will find the frequencies higher when the satellite is approaching and lower when it is receding, due to Doppler effect. At the nearest approach, the Doppler shift is zero. If the Doppler shift is plotted against time, a curve, such as that shown in Fig. 9.1, is obtained.

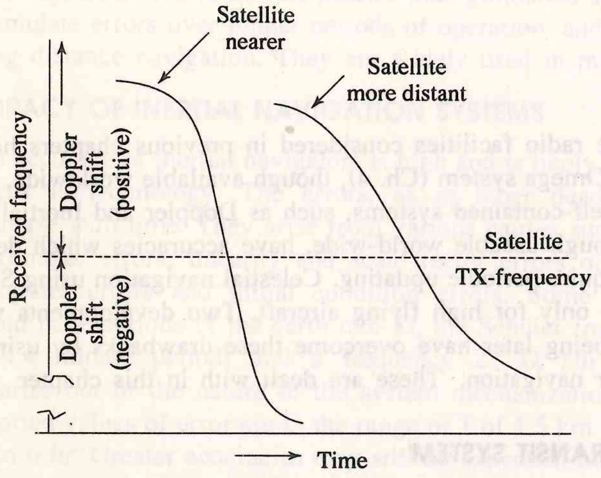


Fig. 9.1 Doppler profiles of satellites

This is called the Doppler profile. When the satellite orbit is near, the profile is steeper, and as it moves further away, the slope of the profile decreases. The actual slope is related to the distance of the satellite at its nearest approach or the longitudinal separation of the satellite from the receiver. The time of zero Doppler (i.e., the nearest approach) can be used to calculate the latitude of the receiver from the known ephemeris of the satellite. This, in principle, is the basis of position determination in the transit system.

But as this method involves the determination of the Doppler shift at a particular instant, it is not practicable. A slightly more indirect method is used. In this method, the number cycles of the beat frequency in successive two minute periods are counted. In order to avoid zero beat, the transmission frequency (previously taken as 400 MHz) is kept 399.968 MHz and this beats with a frequency of 400 MHz in the receiver. Then the previous zero beat is shifted to 32 kHz. This frequency is higher than the highest Doppler shift ever attained, so that the beat frequency between the received signal and the receiver frequency is always positive. Further, instead of measuring the instantaneous frequency. The number of cycles in successive periods of two minutes are counted. These are called Doppler counts. The satellite transmits time markers at two minute intervals to facilitate this measurement. Figure 9.2 shows a plot of the Doppler frequency against time, with the two minute intervals marked. The areas of each of the shaded intervals is proportional to the Doppler count of that interval. It can be shown (Ref. 32) that the Doppler count is a linear function of the difference of the distance of the satellite from the receiver at the beginning and the end of the two minute period, i.e., $(S_2 - S_1)$, $(S_3 - S_2)$, $(S_4 - S_3)$, etc. (Fig. 9.3). The count N_{12} in the interval between t_1 and t_2 is given by

$$N_{12} = (f_g - f_T)\,(t_2 - t_1) + \frac{1}{\lambda_g}\,(S_2 - S_1)$$

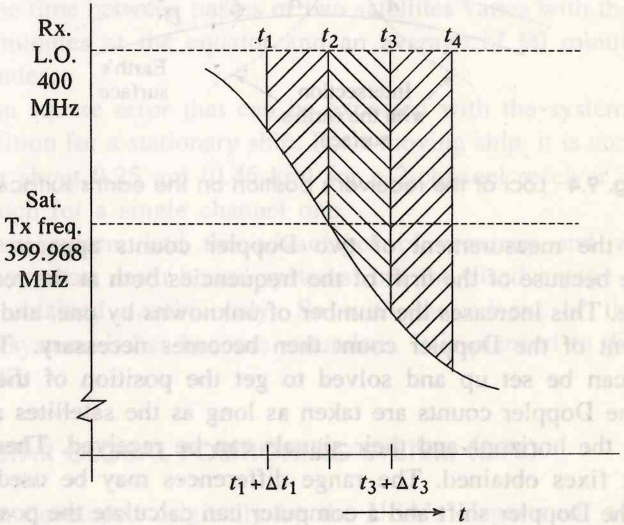


Fig. 9.2 Doppler counts

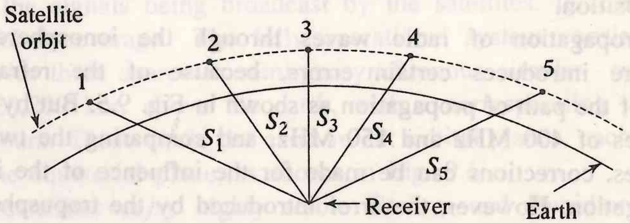


Fig. 9.3 Satellite positions at 2 min. interval

where f_g is the receiver oscillator frequency (400 MHz), f_t the satellite frequency (399.968 MHz), λ is the wavelength of transmission, S_1 and S_2

the distances of the satellite from the receiver and t_1 and t_2 (Fig. 9.2). This derivation assumes that the receiver is stationary. The locus of all points which have this constant difference in distance is a hyperboloid of revolution with A and B as foci (Fig. 9.4). Therefore, the receiver must be somewhere on the intersection of these hyperboloidal surfaces with the earth's surface as shown in Fig. 9.4. The next measurement between t_2 and t_3 gives another pair of hyperboloidal surfaces and their intersection with the earth's surface gives another pair of loci. The intersection of these loci give the possible positions of the receiver. Generally these points of intersection are so far apart that the selection of the correct one will not be a problem if the approximate position of the receiver is known.

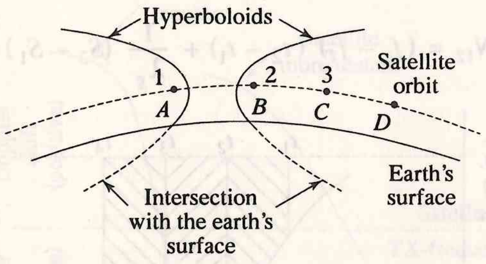


Fig. 9.4 Loci of the receiver's position on the earths surface

Though the measurement of two Doppler counts appears sufficient, errors arise because of the drift of the frequencies both at the receiver and the satellite. This increases the number of unknowns by one, and one more measurement of the Doppler count then becomes necessary. Then three equations can be set up and solved to get the position of the receiver. Actually the Doppler counts are taken as long as the satellites are visible (i.e. above the horizon) and their signals can be received. There may be five or six fixes obtained. The range differences may be used independently of the Doppler shift and a computer can calculate the position from these data and give an estimate of the ship's position. Then certain optimisation procedures can be employed to get the best estimate of the ships's position.

The propagation of radio waves through the ionosphere and the troposphere introduces certain errors, because of the refraction and bending of the path of propagation as shown in Fig. 9.5. But by using two frequencies of 400 MHz and 150 MHz, and comparing the two Doppler frequencies, corrections can be made for the influence of the ionosphere on propagation. However, the error introduced by the troposphere cannot be compensated by this method.

The transit system gives best results when the maximum elevation angle of the satellite is more than 10° to 15° and less than 70°. At least 4 to 5 counts of two minute duration are necessary. Lower angles give larger errors of latitude determination and larger angles give larger errors of longitude determination. As the number of satellites is low and the orbits are polar, the time between passes of two satellites varies with the latitude, being 110 minutes at the equator and an average of 90 minutes at the higher latitudes.

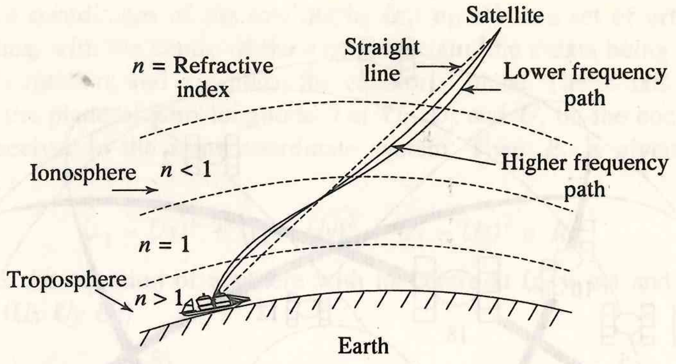


Fig. 9.5 Bending of rays in ionosphere and troposphere

The mean square error that can be expected with the system is about 80 m in position for a stationary ship. For a moving ship, it is considerably more, being about 0.25 nm (0.46 km) for a 2-channel receiver and about twice as much for a single channel one.

The transit system had the advantages of accuracy and worldwide coverage over most earth bound systems, but the disadvantage that fixes cannot be obtained continuously. So with the advent of the Global Positioning system, it has become redundant. It is planned to discontinue it from 1997.

9.2 NAVSTAR GLOBAL POSITIONING SYSTEM (GPS)

This system developed originally by the USA's Department of Defence (DoD), has come into widespread use both by the military, and with slightly reduced accuracy, by all users equipped with receivers capable of processing the signals being broadcast by the satellites. It continuously provides global coverage. The fully operational system consists of 21 satellites, with three spares in semi-geosynchronous circular orbits at a height of 20,200 km, above the earth, with a corresponding period of nearly 12 hours. These circle the earth in six orbital planes, each inclined at 55° to the equatorial plane, as shown in Fig. 9.6.

All the satellites carry highly stable Cesium and Rubidium atomic clocks which are ensured by the DoD to be synchronous with a common time standard, called the GPS time. They continually broadcast their identity code, ephemeris constants from which their current position can be

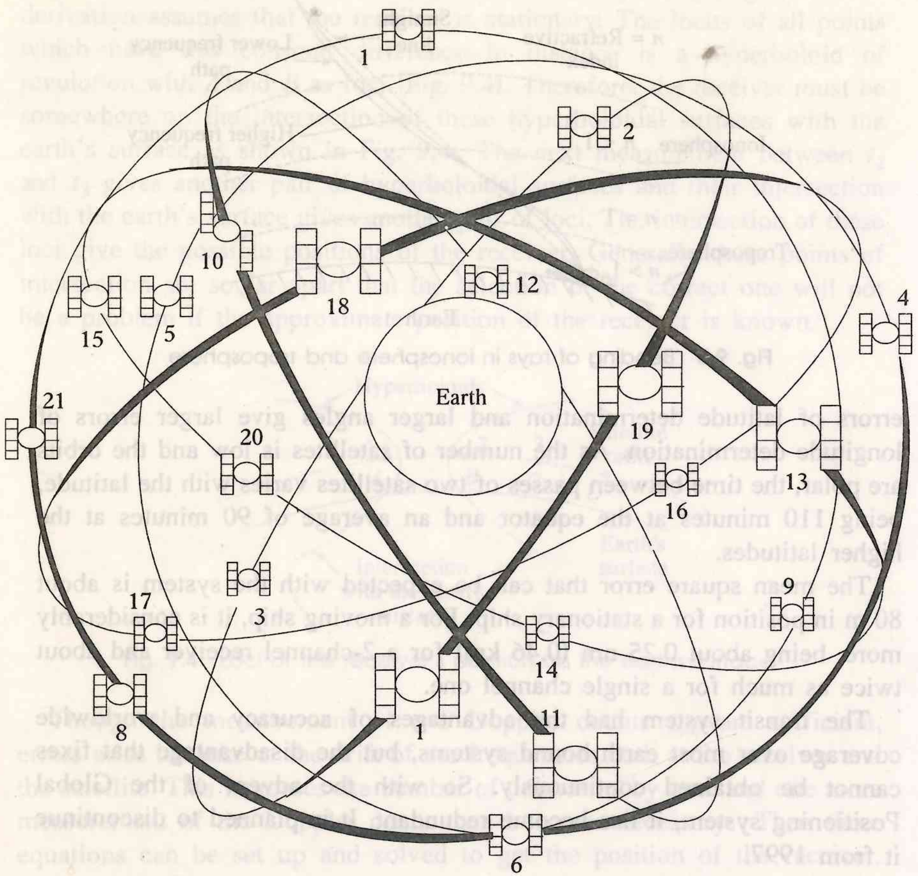


Fig. 9.6 The navstar satellite constellation (21 satellite are shown)

determined, their health status, and almanac constants from which the approximate position of all the satellites can be determined. These are contained in a data stream, the details of which are given in section 9.2.3. These broadcasts are at two frequencies in the *L*-band, $L_1 = 1575.42$ MHz and $L_2 = 1227.6$ MHz(*C/A* and *P* channels).

9.2.1 Basic Principles of Operation

the GPS receiver on (or near) the earth is provided with a high precision clock, which ideally, is synchronized with the satellite clocks. The signal sent by 1 (say S_1) arrives after a delay of Δt_1 which can be precisely determined as the clocks are perfectly synchronized. The distance of this satellite from the receiver is then R_1 given by $R_1 = c\ \Delta t_1$, where c is the velocity of electromagnetic waves.

Let the coordinates of the satellite be (x_1, y_1, z_1) in a set of orthogonal coordinates, with the centre of the earth as origin, the *z*-axis being the axis of earth's rotation, and *x-y* plane the equatorial plane. The *x*-axis may be taken in the plane of zero longitude. Let U_x, U_y and U_z be the coordinates of the receiver in the same coordinate system. Then R_1 is given by the equation

$$(x_1 - Ux)^2 + (y_1 - Uy)^2 + (z_1 - Uz)^2 = R_1^2 \qquad 9.1$$

This is the equation of a sphere with its centre at $(x_1 y_1 z_1)$ and passing through $(Ux\ Uy\ Uz)$.

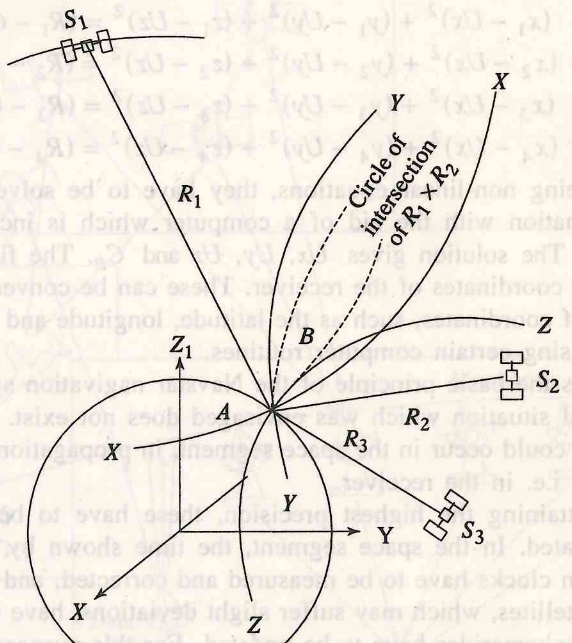


Fig. 9.7 Position finding by three satellites S_1, S_2, and S_3, at distances R_1, R_2 and R_3. Intersection of spheres of radius R_1 and R_2 shown dotted, *A* and *B* are the two possible positions of the receiver

If a similar measurement is made on a second satellite, say S_2, a second equation can be set up. The satellite is then somewhere on the intersection of these two spherical surfaces, which is a circle. A third measurement with a satellite S_3 defines another spherical surface. The intersection of this with the circle gives two points, one of which is the required position of the receiver. It is generally possible to select the correct one from these, the other being an extraneous solution.

This is an ideal situation which does not exist in practice. The first source of error is the clock in the receiver, which is invariably a crystal clock which is not as stable as the atomic clocks in the satellites. This introduces an error in the measured time, say δt. The corresponding error in the range is $C_B = c \times \delta t$, where c is the velocity of e.m. waves. This error is called the 'clock bias', and the range calculated by including this is called the 'pseudo-range'. However, as all measurements are made at the same time or within a short time of each other, the clock bias is the same for all measurements and C_B may be regarded as the fourth variable, in addition to Ux, Uy and Uz. Hence, if time delay measurements are made on four satellites, four equations can be set up, which can be solved for Ux, Uy, Uz and C_B. These equations are

$$\left.\begin{aligned}(x_1 - Ux)^2 + (y_1 - Uy)^2 + (z_1 - Uz)^2 &= (R_1 - C_B)^2\\ (x_2 - Ux)^2 + (y_2 - Uy)^2 + (z_2 - Uz)^2 &= (R_2 - C_B)^2\\ (x_3 - Ux)^2 + (y_3 - Uy)^2 + (z_3 - Uz)^2 &= (R_3 - C_B)^2\\ (x_4 - Ux)^2 + (y_4 - Uy)^2 + (z_4 - Uz)^2 &= (R_4 - C_B)^2\end{aligned}\right\} \qquad 9.2$$

These being non-linear equations, they have to be solved by successive approximation with the aid of a computer which is incorporated in the receiver. The solution gives Ux, Uy, Uz and C_B. The first three are the cartesian coordinates of the receiver. These can be converted to any other system of coordinates, such as the latitude, longitude and altitude over the geoid* using certain computer routines.

This is the basic principle of the Navstar nagivation system. However, the ideal situation which was envisaged does not exist. Deviations from the ideal could occur in the space segment, in propagation, and in the user segment, i.e. in the receiver.

For attaining the highest precision, these have to be allowed for or compensated. In the space segment, the time shown by the Cesium and Rubidium clocks have to be measured and corrected, and the actual orbits of the satellites, which may suffer slight deviations, have to be determined and the ephemerides have to be updated. For this purpose, certain ground facilities constituting the control segment of the system have been established. These consist of five monitoring stations not far from the equator at widely separated points on the earth, a Master Control station in the USA and a data uplink (also in the USA) for transferring the data to the satellites. The Master Control station is equipped with a hydrogen maser atomic clock, which is 50 to 100 times more stable than the Cesium clocks. The monitoring stations are located at precisely known positions distributed over the earth's surface, and have Cesium clocks which are checked with the Master Control station. The monitoring stations measure the position of the satellites by using the range equations (9.2) in the reverse, i.e. Ux, Uy and Uz are known and x_1, y_1 and z_1, the coordinates of the satellites are the unknowns. The data collected by them are sent to the Master Control station where they are analysed, and the corrections to the satellite clocks and ephemerides are sent to the satellites in a data uplink, operating in the S-band. The satellites, in turn, communicate these corrections back to the users in the data stream which they broadcast. So the GPS system may be regarded as consisting of three segments: (a) the satellites themselves (b) the control segment on the earth consisting of monitoring stations, the Master Control station and data uplink, and (c) the innumerable users of the system on or near the earth. This is illustrated in the Fig. 9.8.

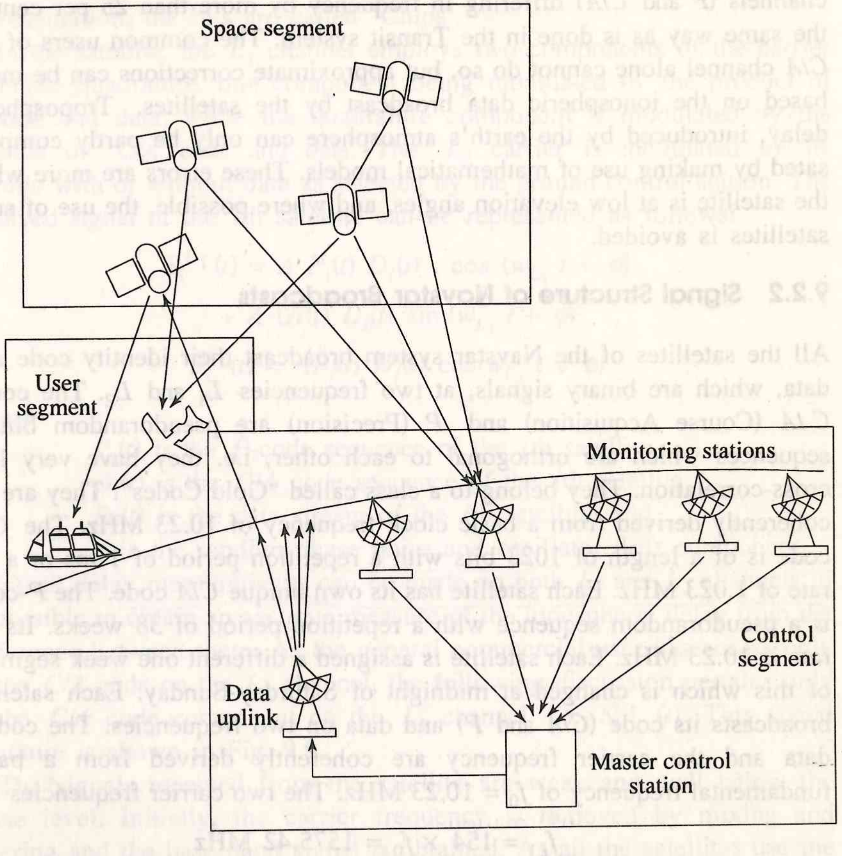


Fig. 9.8 The three segments of the GPS system

There are two other sources of error in the GPS, due to (1) Relativistic effects, and (2) Ionospheric and tropospheric distortion of the path of propagation from the satellite to the user.

The relativistic effects, which are predicted by Einstein's general and special theories of relativity, arise from the fact that the satellites are in a gravitational field which is only 6% of that on the earth and they are moving at 12,000 ft/s compared to a receiver on the earth, which may have a velocity of less than 1000 ft/s.

---

* Geoid is the surface constituted by the mean sea-level extended over the whole

These affect the clock rate. But as these can be predicted, the satellite clock frequency, which is nominally 10.23 MHz, is reduced to 10.22999999545 MHz. For very high dynamic user sets (with velocities exceeding 1000 ft/s), special relativistic effects due to velocity differences of the user with the satellites may have to be considered and compensated.

The ionosphere and troposphere also introduce errors, as in the Transit system, due to the bending of the path of propagation. The ionospheric delay can be compensated by users of the *P*-channel, who can use the two channels (*P* and *C*/*A*) differing in frequency by more than 25 per cent in the same way as is done in the Transit system. The common users of the *C*/*A* channel alone cannot do so, but approximate corrections can be made based on the ionospheric data broadcast by the satellites. Tropospheric delay, introduced by the earth's atmosphere can only be partly compensated by making use of mathematical models. These errors are more when the satellite is at low elevation angles, and where possible, the use of such satellites is avoided.

9.2.2 Signal Structure of Navstar Broadcasts

All the satellites of the Navstar system broadcast their identity code and data, which are binary signals, at two frequencies L_1 and L_2. The codes *C* /*A* (Course Acquisition) and *P* (Precision) are pseudorandom binary sequences which are orthogonal to each other, i.e. they have very low cross-correlation. They belong to a class called "Gold Codes". They are all coherently derived from a basic clock frequency of 10.23 MHz. The *C*/*A* code is of a length of 1023 bits with a repetition period of 1 ms or a bit rate of 1.023 MHz. Each satellite has its own unique *C*/*A* code. The *P*-code is a pseudorandom sequence with a repetition period of 38 weeks. Its bit rate is 10.23 MHz. Each satellite is assigned a different one week segment of this which is changed at midnight of Saturday-Sunday. Each satellite broadcasts its code (*C*/*A* and *P*) and data on two frequencies. The codes, data and the carrier frequency are coherently derived from a basic fundamental frequency of f_0 = 10.23 MHz. The two carrier frequencies are

$$f_{L1} = 154 \times f_0 = 1575.42 \text{ MHz}$$

$$f_{L2} = 120 \times f_0 = 1227.60 \text{ MHz}$$

The codes and data, which are all binary signals, are combined by Modulo-2 (exclusive *OR*) addition and modulate the carrier by quadrature phase shift, i.e. $\pm \pi/2$. So a change from 0 to 1 and 1 to 0 results in phase reversal. This is equivalent to multiplication by + 1 and − 1. Thus the modulation is a multiplication by a binary chain where 0 is represented by + 1 and 1 is represented by − 1. This is shown by the truth table below, where such a representation is shown in brackets.

Table

Code	*Data*	*Code + Data*	*Code × Data*
0 (+1)	0 (+1)	0 (+1)	+ 1
0 (+1)	1 (− 1)	1 (− 1)	− 1
1 (−1)	0 (+1)	1 (− 1)	− 1
1 (−1)	1 (− 1)	0 (+1)	+ 1

It can be seen that the above is binary phase shift keying (BPSK). In this representation, the bits are called 'Chips'.

In the satellite, the L_1 channel employs two components of the carrier in phase quadrature, one component being modulated by the product of *P*-code and data, while the quadrature component is modulated by the product of *C*/*A* code and data. The L_2 carrier is modulated by its *P*-code with or without data as selected by the ground control station. The received signal of the *i*th satellite can be represented as follows:

$$S_i^{L1}(t) = A\ P_i(t)\ D_i(t)\ .\ \cos\,(w_{L_1}\ t + \phi)$$

$$+\ A\ Gi(t)\ D_i(t)\ \sin\,(w_{L_1}\ t + \phi)$$

$$S_i^{L2}(t) = AP_i(t)\ D_i(t)\ \cos(w_{L_2}\ t + \phi)$$

where

$P_i(t)$ is the *P*-code sequence of the *i*th satellite,
$G_i(t)$ is the *C*/*A* code sequence of the *i*th satellite,
$D_i(t)$ is the data stream of the *i*th satellite, and
ϕ is the random phase noise and oscillator drift component.

If transit delay measurements can be made on both L_1 and L_2 channels, it is possible to obtain an accurate measure of the ionospheric delay from the difference between them. As the general commercial user has access only to the *C*/*A* code on the L_1 channel, the following discussion pertains only to the *C*/*A* code component on the L_1 channel, i.e. $S_i^{L1}(t)$. This signal structure is shown in Fig. 9.9.

The signals received from the satellite are weak and well below the noise level. Initially, the carrier frequency is removed by mixing and filtering and the base band signal is obtained. As all the satellites use the same frequency, the codes of all the visible satellites are present in this, in addition to noise. The receiver has to extract the desired signal and measure the delay. The technique employed to achieve this is to compute the correlation function of the incoming signal with a replica of the desired code which is available in the receiver by evaluating the integral

$$\frac{1}{N}\int_{1}^{N\tau} S_i\,(t)\,\{G_i\,(t + j\tau)\}\,dt$$

where $S_i\,(t)$ is the demodulated signal and $G_i(t)$ the replica or the desired code delayed by j chip lengths and (τ) is the chip length. The value of j

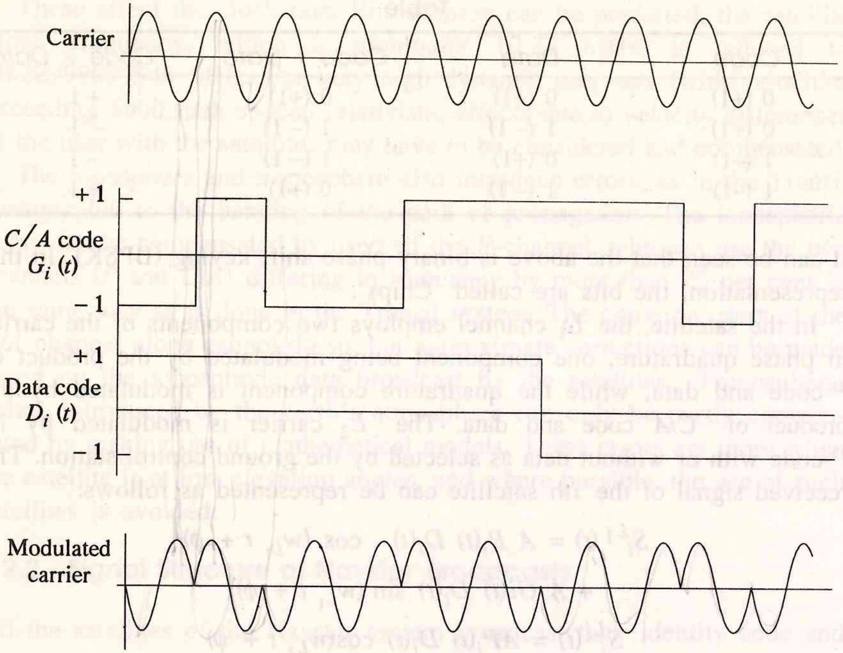


Fig. 9.9 Signal structure

is changed from 1 to *N*, i.e. one code length or 1023 chip lengths, and the integral is evaluated for each delay. When the two codes match, the integral jumps to a maximum value, as the integral is then the auto correlation function for the code, which is a Gold code. The locally generated code is locked at this value of delay, and the two signals move in step subsequently. The output of the correlator would be a dc of this maximum value, but as $S_i(t)$ has a factor $D_i(t)$ multiplying the code, the output changes sign whenever $D_i\,(t)$ changes sign. So the output of the correlator is the data signal $D_i(t)$. The circuit of the correlator is shown in Fig. 9.10 (a) and the auto correlation function in 9.10 (b).

There is a small amount of noise (not shown) in the autocorrelation function. Generally, it is possible to locate the peak of the autocorrelation function to a tenth of the pulse width, i.e. 0.1 μs by proper threshold setting which improves the position accuracy up to 30 m.

This measurement pertains to one code length, i.e. 1 ms. But the total delay has in addition an integral number of milliseconds. As one millisecond delay corresponds to a distance of 300 km, no ambiguity is likely to arise if the position of the receiver is even approximately known. If the resolution of this ambiguity is also necessary, it can be done in the initial satellite acquisition process, as part of obtaining the first fix.

The delay determination for the *P*-code follows a similar process, but the problem of ambiguity does not arise as the code is of one week

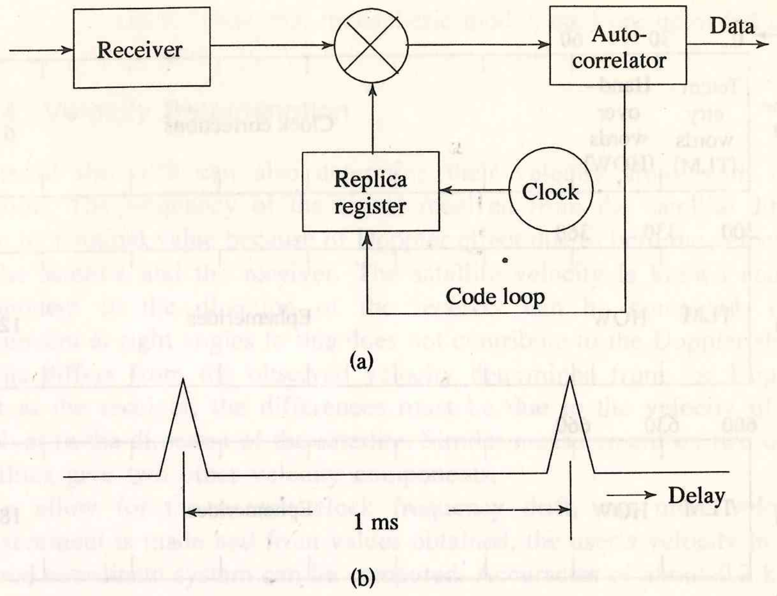


Fig. 9.10 (a) Circuit of the correlator
(b) The auto-correlation function

9.2.3 Data Message

The GPS navigation message is superimposed on the code chip through modulo-2 addition. It has a bit rate of 50 bits/s and has a length of 30 seconds. This is called a frame. Each frame is divided into five sub-frames of 6 seconds each containing 300 bits of information. The format of this transmission is shown in Fig. 9.11.

In each subframe, the Telemetry words (TLM) and Hand-Over Words (HOW) are generated within the satellite, and the rest are uploaded from the Master Control Station. Each subframe starts with the Telemetry Word (TLM) which is used for synchronization. This is followed by the Hand Over Word (HOW) which gives the number of 1.5 s periods which have elapsed since the start of the GPS week, i.e. the start of the *P*-code. It helps the users of the *P*-code to acquire the code in a short time as it indicates the time elapsed since the start of the code. The satellite time can also be obtained from this to resolve any time measurement ambiguity by the *C/A* code users. The information of the GPS message is listed below:

Subframe 1 contains satellite clock correction coefficients, satellite health, and age of data.
Subframe 2 and 3 contain the satellite ephemeris parameters from 45 minutes before to 15 minutes after the time of the message. These data are repeated every 30 seconds and changed every hour.

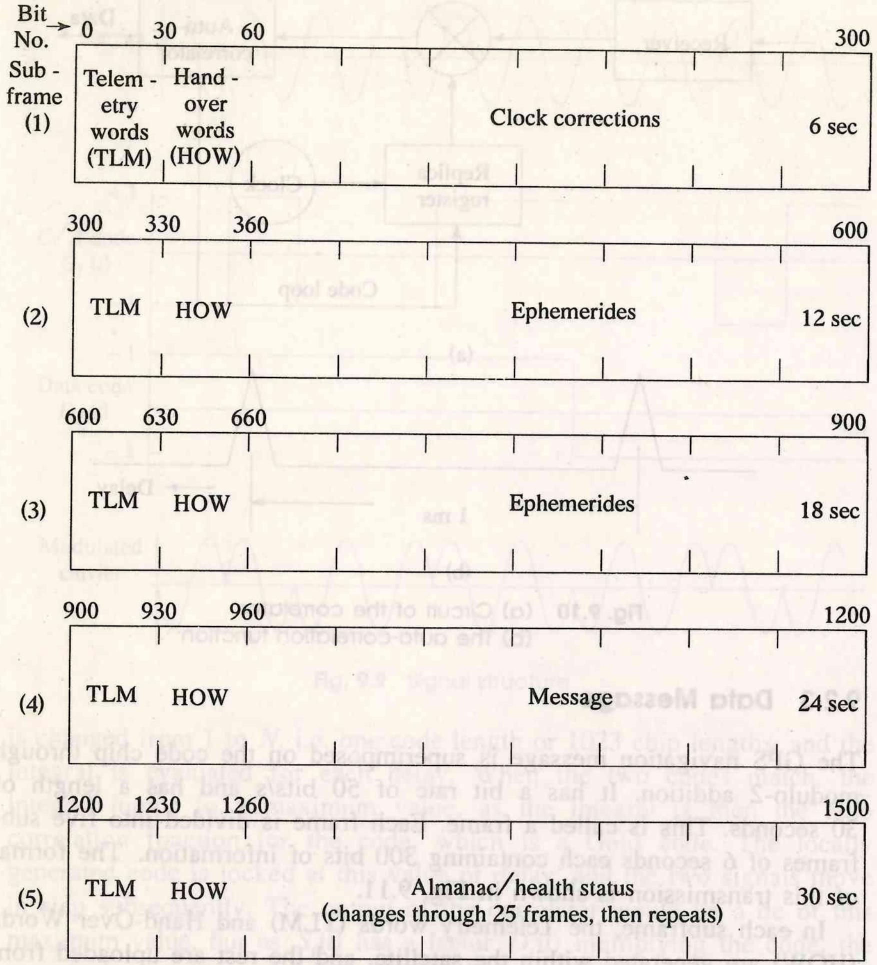


Fig. 9.11 Data Message Format

Subframe 4 contains navigation messages and satellite health status information.

Subframe 5 contains almanac information on all satellites. Almanac contains fewer constants pertaining to the satellite orbits and is therefore less accurate than the ephemeris. But it defines the orbits of the satellites sufficiently accurately for the user to chose the satellites most favourably situated to get an accurate position fix. Each 30 second frame provides the information about one satellite and the information is repeated after 25 frames, i.e. after 12.5 minutes. These data are up-loaded from the Master Control station about once a day and the satellite transmits one hour segments of these data. Other data in the data block frame (e.g. ephemeris, clock correction, ionospheric model, etc.) are uploaded once an hour.

9.2.4 Velocity Determination

Users of the GPS can also determine their velocity along with their position. The frequency of the signal received from the satellite differs from its nominal value because of Doppler effect due to both the velocities of the satellite and the receiver. The satellite velocity is known and its component in the direction of the receiver can be computed. (The component at right angles to this does not contribute to the Doppler shift). If this differs from the observed velocity determined from the Doppler shift at the receiver, the differences must be due to the velocity of the receiver in the direction of the satellite. Similar measurement on two other satellites give two other velocity components.

To allow for the users's clock frequency drift, one more velocity measurement is made and from values obtained, the user's velocity in any desired coordinate system can be computed. Accuracies of about 0.2 km/h can be obtained.

9.2.5 Accuracy of Position Determination

Position determination using Navstar satellites is subject to errors arising from two sources:

1. The average User Equivalent Range Error (UERE) along the line of sight vector connecting the satellite and the user, and
2. The instantaneous Geometrical Dilution of Precision (GDOP) which depends on the geometry of the four satellites used as seen from the receiver's position on the earth.

The user equivalent error arises from several sources. The first is from the space segment, due to perturbation in space, satellite clock instabilities and deviations from the ephemeris, etc. Theses are relatively small and the same for *P*-code and *C*/*A* code users. In the user segment, ionospheric and tropospheric delays and receiver noise are the main sources or error. In this case, compensation for ionospheric delays and more precise range measurements of *P*-code users give them a distinct advantage, and the errors are considerably less for them than for *C*/*A* code users.

Geometrical dilution of precision depends on the choice of satellites for making measurements. Very low elevation angles (5° or less) are generally avoided. In the remaining space, the error depends on the relative positions of the satellites. The small probable errors in range would define a small volume around a receiver, the shape and size of which is an indication of the degree of uncertainty in the measured position. As a general rule, the error depends on the volume of the tetrahedron formed by the four chosen satellites and the receiver. This must be as large as possible. Thus

The average error (mean-square error) is taken to be equal to the product of these two quantities (GDOP and UERE). This is typically 50 ft. or less for *P*–code users and about 150 ft for *C/A* code users. This accuracy can be deliberately degraded by the authorities (DoD), when they so desire, by introducing a jitter in the pulse train.

The accuracy of navigation solutions can be enhanced by sophisticated mathematical techniques such as Kalman fittering, which provide the best estimates of position from a selection of observed estimates and a prior knowledge of user dynamic models, etc.

9.2.6 Differential Navigation

If there are two locations, the position of one of which is known exactly, say in latitude and longitude, the position of the second location with respect to the first can be determined with great precision by the G.P.S. system if they are not too far apart, and if both use the same satellites for determining the position. This is called differential navigation. The reason for the higher precision is that the radiations from the satellites take nearly the same path, as the distance of, say 100 km is small compared with the distance from the satellites which are more than 20,000 km away. So, many of the errors which arise because of propagation, get cancelled. Position accuracies of the order of centimetres can be obtained. Thus, relative accuracy is much more than absolute accuracy. Occasions to use such methods arise in conducting surveys. Similar situations arise in landing of aircraft on carriers, in refuelling aircraft in flight, etc.

9.2.7 Navstar Receivers

Navstar receivers have to receive exceedingly weak signals coming from the satellites, amplify and process them and display the coordinates of the receiver. The early receivers were bulky, heavy, and expensive. Later versions are much more compact, some being so small that they are described as hand-held models. They all use modern advanced electronic techniques with Gallium arsenide computer chips, Application-specific Intergrated Circuits (ASIC's), highly efficient antennas, and digital processing techniques with efficient routines. All these advances have made the receiver not only smaller but less expensive.

A modern receiver can be broken down into five sub-assemblies.

1. Antenna and associated electronics,
2. The tracking loops,
3. The navigation processor,
4. Control and display unit, and
5. The power supply.

Figure 9.12 illustrates these major functions of the receiver sub-assemblies and indicates the functions performed by them. A brief account of their

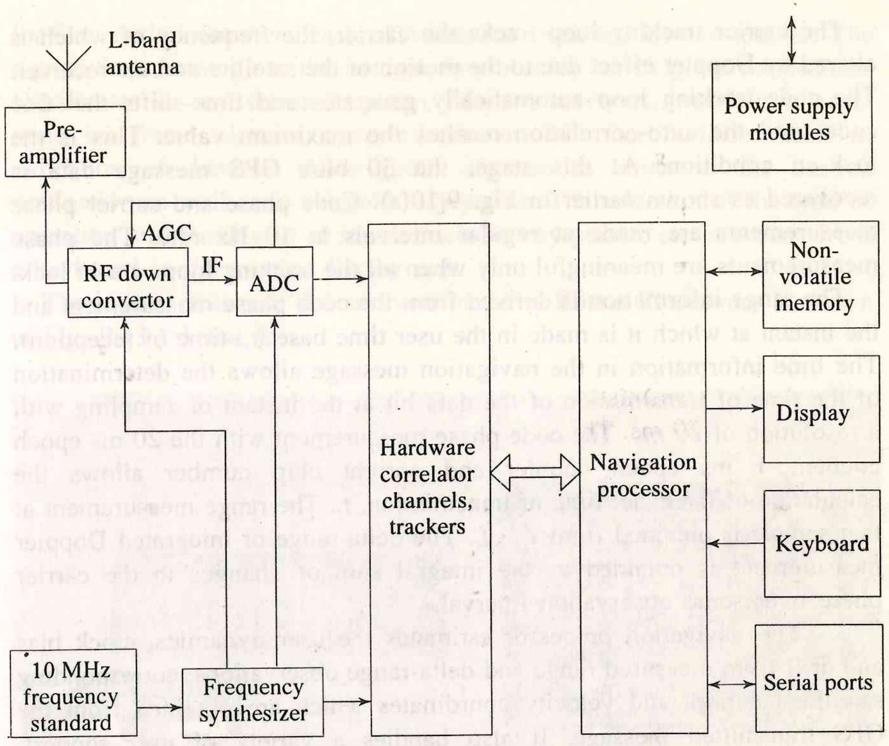


Fig. 9.12 Modern GPS servicing functional black diagrams

The satellites radiate a right-hand circular polarised beam and the receiver antenna is designed to pick up these. As the antenna may be situated away from the rest of the receiver, the signal picked up is preamplified by an amplifier located near the antenna and taken to the receiver in co-axial cable. This improves the signal/noise ratio which would otherwise be reduced by the cable. Sometimes, a ground plane is employed to isolate the antenna from multipath signals and other interference. Military aircraft may use two antenna, one on top and the other at the bottom of the aircraft as they sometimes fly up-side down.

The R.F. down-convertor brings down the RF signal to manageable intermediate frequency through one or more convertors. The IF signal is then converted to digital form and fed to the channel hardware. The frequency synthesizer coherently derives all system frequencies for the RF down-converter and channel hardware from stable crystal oscillator references.

2. The channel hardware, in conjuction with carrier and code tracking loops implemented in software, rotates the carrier phase to minimize the phase noise due to oscillator frequency drifts and user dynamics, and removes the code to get the GPS message.

The carrier tracking loop tracks the carrier, the frequency of which is altered by Doppler effect due to the motion of the satellite and the receiver. The code tracking loop automatically generates and time-shifts the *C/A* code until the auto-correlation reaches the maximum value. This is the lock-on condition. At this stage, the 50 bit/s GPS message data is recovered as shown earlier in Fig. 9.10(a). Code phase and carrier phase measurements are made at regular intervals at 10 Hz rate. The phase measurements are meaningful only when all the tracking loops are in lock.

The range information is derived from the code phase measurement and the instant at which it is made in the user time base (t_r–time of reception). The time information in the navigation message allows the determination of the time of transmission of the data bit at the instant of sampling with a resolution of 20 ms. The code phase measurement with the 20 ms epoch counter, 1 ms epoch counter and current chip number allows the calculation of the exact time of transmission, t_t. The range measurement at that sample is obtained from $t_r - t_t$. The delta-range or integrated Doppler measurement is obtained as the integral sum of changes in the carrier phase over some observation interval.

3. The navigation processor estimates the user dynamics, clock bias and drift from measured range and delta-range observations, corresponding satellite position, and velocity coordinates which are obtained from the GPS transmitted message. It also handles a variety of user support, functions, such as selection of satellites for better accuracies, coordinate transformation, routing calculation for the best way-point navigation, etc. (The way-point is an intermediate latitude-longitude combination through which the aircraft must fly in order to reach the desired destination).

The processor also stores the last position coordinates and almanac constants in a non-volatile memory, when the receiver is turned off. These can be used as a first estimate of position when the receiver is turned on again, as also the satellites which are most favourably situated for navigation, thus shortening the time required to start operation again. The serial ports are for entry and extraction of additional data.

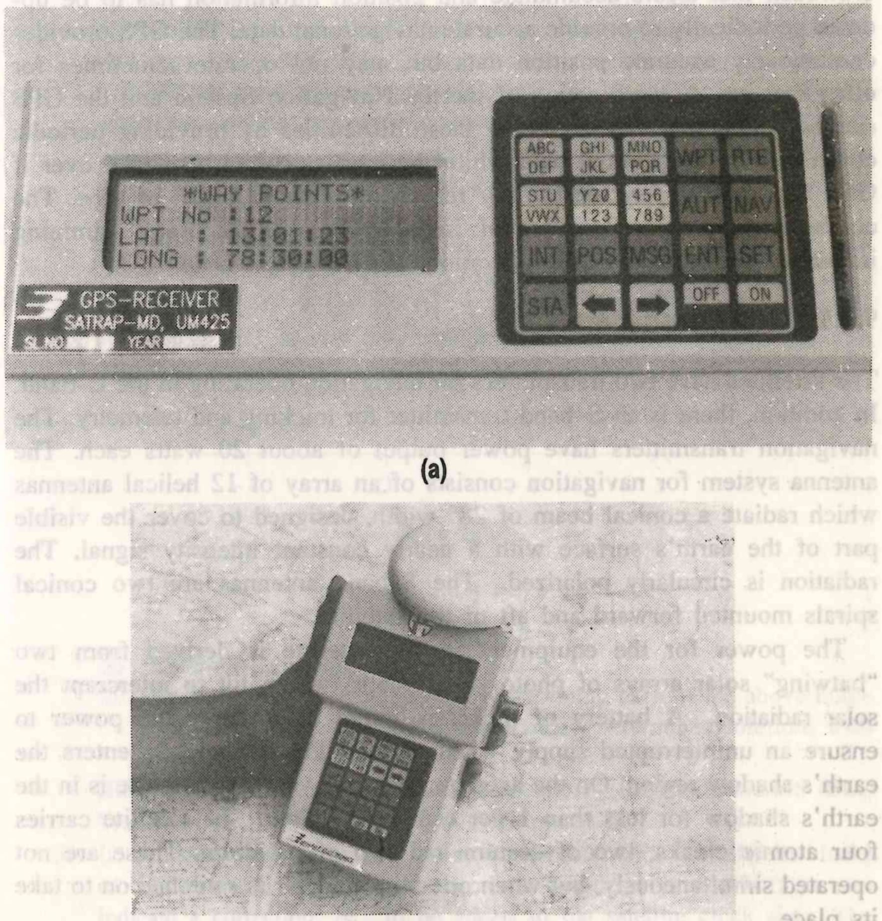
4. The control display unit is for feeding-in instructions regarding operating modes, coordinate systems, way points, etc. The displays give the current position and velocity, the distance made good, velocity made good, ground speed, etc. The displays are commonly LED's (light emitting diodes), LCD's (liquid crystal displays) and cathode-ray tubes.

5. The power supply is generally provided by rechargeable batteries or, sometimes, by disposable batteries. The electrical systems of vehicles in which they are used may also be used.

The receiver may take different forms depending on the application. High performance receivers are generally capable of processing the data from four or more satellites simultaneously. In this case, the signal from the down-converter goes through an analog to digital converter to four

(or five) digital channels, each of which can track a different satellite. The signals from these go to the navigation processor which gives the position coordinates in any selected system. Modern receivers, which use large-scale integrated circuits, have a selection of other facilities such as indication of distance made good, distance to cover to the desired destination as well as to selected way-points. The selection of the type of operation and feeding in of inputs such as way-points, etc. are done from a keyboard in the control display unit.

The picture of two of the receivers made by Bharat Electronics is shown in Fig. 9.13 (a) and (b).



(a)

(b)

Fig. 9.13 Modern receivers made by Bharat Electronics (a) Satrap-MD (medium dynamics) (b) Satrap-LD (low dynamics)

The above description refers to multichannel receivers which are suitable for exacting application but are expensive. When the dynamics of can also be used. In this case, the reception of signals from the satellites is done sequentially either at a low switching rate or a high switching rate. The movement of the satellites implies that the satellite positions are not simultaneously determined, and the position of the receiver also changes in the interval. This leads some reduction in the accuracy of position determination, but with vehicles with low velocities, this will not be serious. Such receivers will naturally be smaller and cheaper.

9.2.8 Integration of GPS with Inertial Navigation Systems

As indicates in chapter 8, the Inertial Navigation System suffers from the drawback that errors accumulate and position information has to be updated periodically to provide accurate navigational data. The GPS provides continuously accurate position data but may not operate sometimes for other reasons. A combination of inertial Navigation System and the GPS suitably integrated can overcome these difficulties by providing periodic corrections to the inertial data while the inertial system can take over if GPS becomes inoperative for any reason for a short length of time. The combination provides substantial improvements in accuracy, jamming immunity and high dynamic operation.

9.2.9 GPS Transmitters

The satellites have two transmitters for navigation operating in the L–band. In addition, there is an S–band transmitter for tracking and telemetry. The navigation transmitters have power output of about 20 watts each. The antenna system for navigation consists of an array of 12 helical antennas which radiate a conical beam of 28° width, designed to cover the visible part of the earth's surface with a nearly constant intensity signal. The radiation is circularly polarized. The S–band antennas are two conical spirals mounted forward and aft of the satellite.

The power for the equipment in the satellite is derived from two "batwing" solar arrays of photo-voltaic cells, which tilt to intercept the solar radiation. A battery of nickel-cadmium cells stores the power to ensure an uninterrupted supply when the satellite periodically enters the earth's shadow region. On the average, over one year, the satellite is in the earth's shadow for less than 1 per cent of the time. The satellite carries four atomic clocks, two of cesium and two of rubidium. These are not operated simultaneously, but when one fails, another is switched on to take its place.

9.2.10 The Russian Glonass System

The Russians have launched satellites for navigational purposes, similar to the Navstar system. The Russian system has 18 satellites moving in each. The planes are inclined at about 68° with respect to the equatorial plane. Each satellite radiates at two frequencies, like the Navstar satellites, but the frequencies of all the satellites are different. The codes, however, are the same. So they employ frequencies division multiplex, as against code division multiplex of the GPS system.

The accuracies of the two systems are comparable. At present, the same receiving equipment on earth cannot make use of both the systems.

QUESTIONS AND PROBLEMS

1. In the transit navigation system, in a particular case, the successive satellite positions are as shown in Fig. 9.3. Assume that in position 3 of the satellite is at the same latitude as the receiver, so that the received frequency is 399.968 MHz. Its longitude is such that the orbit projection is 2000 km from the receiver. Calculate the Doppler count registered when the satellite moves from position 2 to position 3 in the figure. [Ignore the curvature of the earth. Note that the satellite orbital velocity is 7.3 km/s and the altitude of the satellite orbit over the earth is 1100 km.]
2. Consider the following space position fixing problem pertaining to one dimension. In the figure shown below, two transmitters are located at S_1 (+4) and S_2 (–4) on the X-axis. A receiver R located at an unknown position in the X-axis measures its absolute distance from the two transmitters subject to an unknown bias b, as given the equations below:

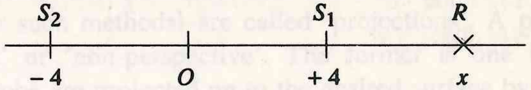


Fig. 9.13

$$[\, x - 4\,] + b = 2$$

$$[\, x + 4\,] + b = 4$$

(a) Determine the unknown x and bias b.

(b) Draw the pair of straight lines representing each of the above range equations and compare the algebraically determined solutions with the points of intersection.

[Ref: "An algebraic solution of GPS equations" Stephen Bancroft, IEEE Trans on AES, Vol AES–21, Jan 1985, pp. 56–59]

3. (a) A satellite transmits a signal at the nominal GPS time [by its clock] of t_{sv}. However, the clock corrections broadcast in the data stream indicate a correction Δt_{sv} to be added to the satellite clock time. The signal is received by the user at time Δt_u by the user's clock, which has got an error indicated by t_{bias}. Write the range equation for the satellite which takes these factors into consideration. Show by a sketch how these factors affect the measured delay.

(b) what are the advantages of using auto-correlation function to

4. (a) How is the user's velocity determined in the Navstar system?
   (b) A receiver is so situated that a satellite passes overhead at time $t = t_0$. At 15 minutes before this time, the Doppler shift in the received frequency measured by the receiver is 3500 Hz. Does the receiver have a velocity component in the direction of the satellite? [Take the satellite orbit radius to be 26,600 km, period 12 hours and earth's radius 6400 km].
5. Discuss the various factors affecting the accuracy of position determination by the Navstar GPS.

Appendix I

Maps and Charts

The first need of a navigator is a map of the region he is going to traverse, wherefrom he can get a knowledge of the position of departure and destination as well as of the intervening terrain. The maps used for navigation are generally called 'charts' and are printed to various scales to suit a veriety of navigational requirements.

The production of a map or a chart presents the problem of representing the spherical (or near-spherical) surface of the earth on a plane. A spherical surface is 'non-developable', i.e. it cannot be spread out on a plane surface on without distortion. Any method of representation of the spherical surface on a plane may be reduced to that of representing the parallels of latitude and the meridians of longitude by lines or curves, having a one-to-one correspondance with them, on a plane. These methods (there are many such methods) are called 'projections'. A projection can be 'perspective' or 'non-perspective'. The former is one in which the points on the globe are projected on to the desired surface by joining them to a point (called the origin) and projecting the line until it intersects the desired surface. This is also the sense in which the word perspective is commonly understood. The point of origin may be anywhere, even at infinity, but in the more common projection is at the centre of the globe. Non-perspective projections are modifications of perspective projections designed to represent some particular aspects of the region being projected, such as area, shape, etc.

All projections are compromises, as they cannot represent the spherical surface accurately in all respects. Some of the important factors to be considered in choosing a projection are: (a) the position of the region in relation to the earth's surface, (b) the direction which any point bears in relation to another, (c) the distance between any two given points in the region, (d) the shape of the region and (e) the accuracy with which the area is represented. All the above requirements are not satisfied by any projection and a choice has therefore, to be made keeping in mind the special requirements of any particular application. For navigation, factors (b) and (c) are important. The commonly used projections employed for tions. Before proceeding to a description of these, it is to be noted that there are two types of routes which can be taken by a craft in going from one point to another on the globe. One is the great circle path, which is the shortest distance between the two points and the other is the rhumb line, which is the line which cuts the meridians at a constant angle. Over short distances (up to 300 or 500 km), there is little difference between the two paths. A rhumb line would be followed if a vehicle navigates by the compass and keeps the compass reading constant.

Mercator's Projections

This projection, which is the most widely used one in marine navigation, is obtained by projecting the points on the globe on to the surface of a cylinder having its axis coincident with axis of the globe. When the surface of the cylinder is 'developed', the parallels of latitude and the meridians appear as set of parallel lines orthogonal to each other (Fig. I.1). In this projection, the distance between the parallels of latitude increases rapidly as one moves towards the poles, and the linear distance from the equator to a parallel is made to increase at such a rate that the scale along a meridian is the same as the scale along the parallel *at that point*. The poles can never appear on this projection. The advantage of Mercator's projection is that a straight line connecting any two points is a rhumb line. The shapes of small areas are well preserved at all latitudes but the scale increases with latitude and, therefore, large areas particularly at the higher latitudes, appear distorted.

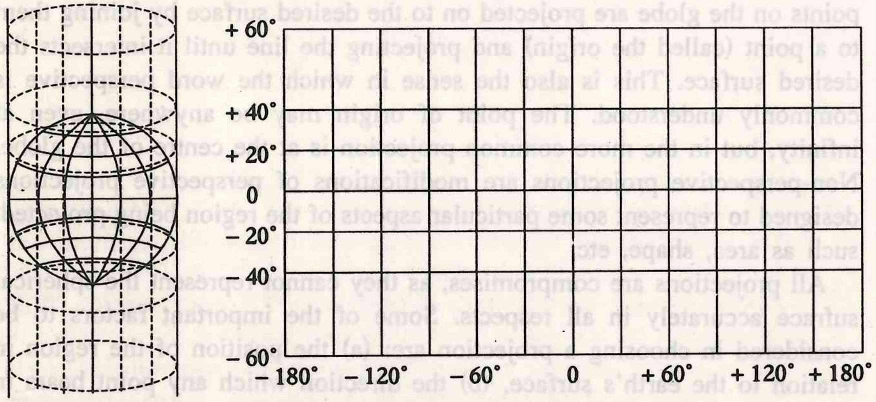


Fig. I.1 Mercator's projection

Mercator's projection is useful when navigation along a rhumb line is adequate, which is generally true for short distances. For navigation over long distances when the great circle path is to be followed, the desired route (which need not be straight on Mercator's projection) is determined

lines are substituted for these. Navigation can then be done by maintaining the required constant bearings over the segments.

Lambert's Conformal Conic or Lambert Projection

In this projection, a cone coaxial with the earth's axis is made to intersect the globe at two parallels of latitude, which are referred to as 'standard parallels'. The simple perspective conic projection is a special case of this, when the two standard parallels coincide. When developed, the Lambert projection yields a map in which the meridians are convergent lines (usually converging well outside the map) and the parallels of latitude appear as arcs of concentric circles (Fig. I.2). The advantages of the Lambert projection are that even over a fairly large area (as for example the U.S.A.), a line joining two points is very close to the great circle and the angular representation is preserved. The bearing of one station from another can, therefore, be read off with a protractor. Lambert projection is widely used in charts for aeronautical navigation.

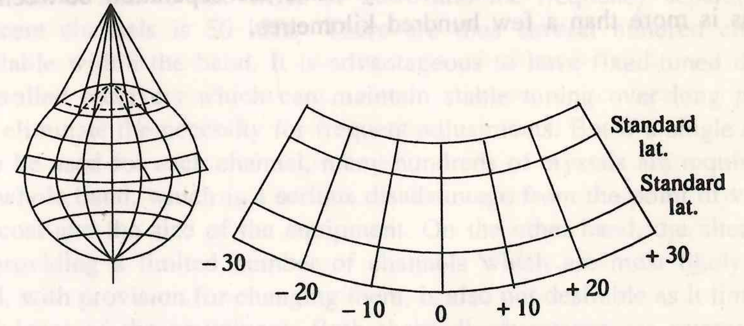


Fig. I.2 Lambert's projection

Gnomonic Projection

This is a perspective projection obtained by taking the centre of the globe as origin and projecting the points on it on to a plane tangential to the globe. All great circles will appear as straight lines and in particular, those passing through the point of contact form a set of radial lines (Fig. I.3). A common form of the gnomonic projection has the point of contact at the pole. The meridians, therefore, appear as radial lines and the parallels as concentric circles. A useful property of this gnomonic projection is that the bearing of all points from the point of contact are true. Though this projection is not used by itself for navigation, it is useful for determining the great circle route between any two points, which is simply the straight line joining the two points. The coordinates (latitudes and longitudes) of the points on this route can be found and transferred to any other desired projection.

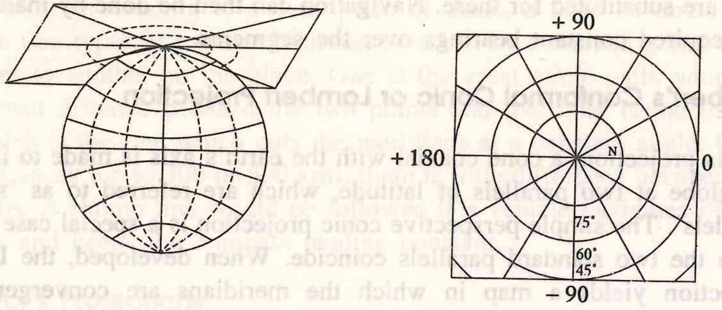


Fig. 1.3 Gnomonic projection

When obtaining the bearing of a fixed station or when obtaining a fix from two or more bearings, it must be remembered that the true bearings obtained at the fixed and mobile stations are not necessarily reciprocals. This is because the meridians are not parallel. Allowance has, therefore, to be made for the convergence of the meridians. These corrections become necessary, at the lower latitudes, only when the separation between the stations is more than a few hundred kilometres.

Appendix II

Multichannel Crystal Controlled Receivers

Aircraft navigation and communication equipment, particularly the ones operating in the VHF band, are crystal-controlled. The allocated band is from about 108 to 136 MHz or more and the frequency separation of adjacent channels is 50 kHz. There are thus several hundred channels available within the band. It is advantageous to have fixed-tuned crystal-controlled receivers which can maintain stable tuning over long periods and eliminate the necessity for frequent adjustments. But if a single crystal is to be used for each channel, many hundreds of crystals are required for the whole band, which is a serious disadvantage from the point of view of the cost and the size of the equipment. On the other hand, the alternative of providing a limited number of channels which are most likely to be used, with provision for changing them, is also not desirable as it limits the usefulness of the equipment. Both these disadvantages are overcome in modern equipment by making provision for all the available channels with a limited number of crystals. This is achieved by the technique of multiple-heterodyning.

This technique, as applied to a receiver, consists in translating the VHF signal frequency to a low intermediate frequency (approximately 500 kHz) in successive steps, using crystal oscillators in each step. This is illustrated in Fig.II.1. The several oscillator frequency groups and the succeeding IF's must be so chosen as to ensure that there is no spurious response from frequencies within the band or outside it. It is clear from the figure that the final number (N) of channels obtainable is given by

$$N = n_1 \times n_2 \times n_3 \ldots \times n_r$$

where r is the number of steps of frequency translation. The number of crystals required (M, say) is

$$M = n_1 + n_2 + n_3 + \ldots + n_r$$

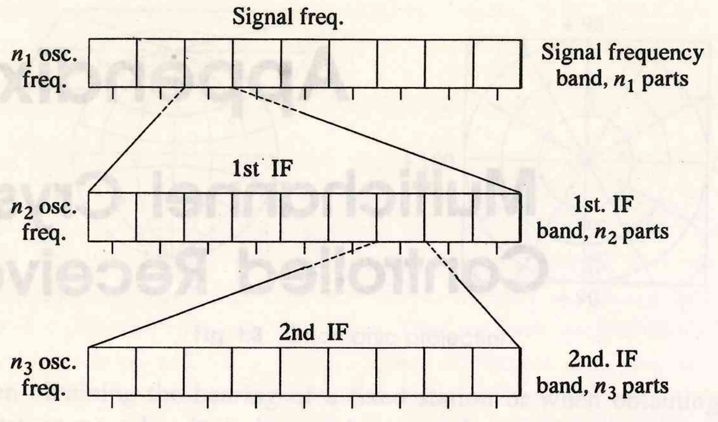


Fig. II.1 A scheme of multiple heterodyning

$$n_1 = n_2 = ... = n_r = n$$

which gives $N = n^r$ and $M = nr$.

For the minimum number of crystals, the value of n works out to be ε (= 2.7183...), but even taking the nearest whole number of $n = 3$ often gives an impracticably large number of heterodyning steps. Taking, for example, $N = 2000$, if $n = 3$, then $r = 7$ ($3^7 = 2187$), i.e. seven heterodyning stages are required. The number of crystals required is 21. Whereas, if three steps of ten and one of two are used, the number of crystals required is 32, which is still not very large, but is certainly very small compared with 2000. In practice, therefore, the number of stages of heterodyning is made two or three and the number of frequencies in each step is made fairly large, generally 10 to 20. The block diagram of a receiver (Marconi *AD* 160) shown in Fig. II.2 employs such a scheme. The diagram is self-explanatory. Apart from the frequency selection scheme, note that there are two demodulated outputs, one for navigational purposes, and the other for audio. The former uses the full bandwidth while the latter has a filter (in the squelch circuit block) which limits the banbdwidth to that required for radio telephony.

The principal disadvantage of the technique of multiple heterodyning is that it tends to increase the spurious frequency responses of the receiver. This has been sought to be overcome by an alternative scheme which employs a single stage of heterodyning followed (and proceeded) by highly selective filters.

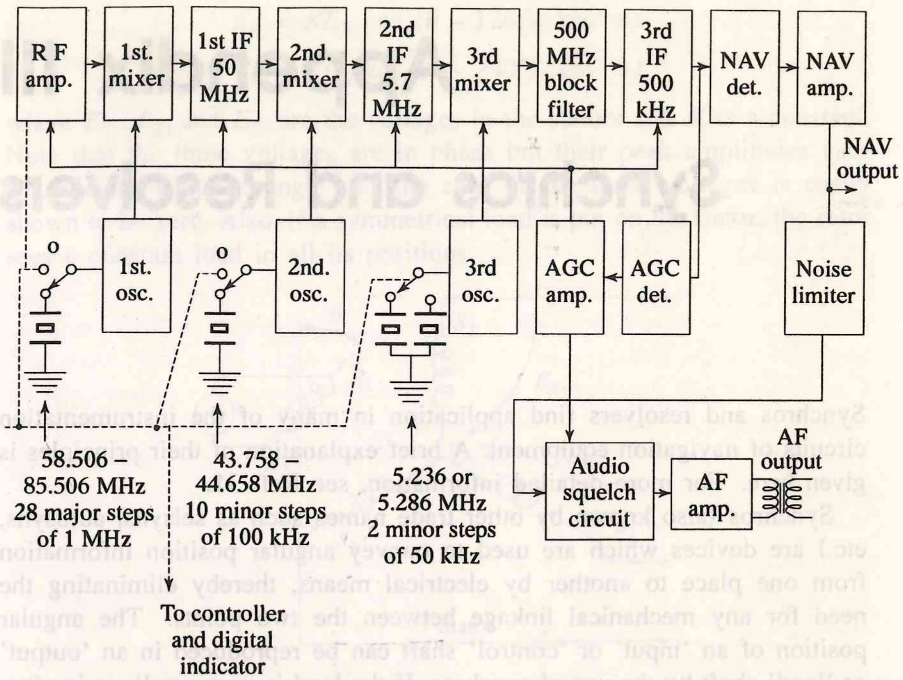


Fig. II.2 A 560-channel receiver (Marconi) using three stages of heterodyning

Appendix III

Synchros and Resolvers

Synchros and resolvers find application in many of the instrumentation circuits of navigation equipment. A brief explanation of their principles is given here. For more detailed information, see Ref. 31.

Synchros (also known by other trade names such as selsyns, autosyns, etc.) are devices which are used to convey angular position information from one place to another by electrical means, thereby eliminating the need for any mechanical linkage between the two points. The angular position of an 'input' or 'control' shaft can be reproduced in an 'output' or 'load' shaft by the use of synchros. If the load is very small, as is often the case with indicating instruments, the synchro can directly operate the load shaft, but if considerable power is required, a feedback control system may be used in which the synchro is employed to convey the output position data to the input point for comparison and generation of error signals.

The synchro is a small ac machine which resembles a motor in construction. It has a stator with three identical windings symmetrically positioned and a rotor with a single winding. Several different types of synchros are used in instrumentation, namely, synchro-generators, synchro-motors, and control-transformers. Electrically, they are identical but there are some differences in the construction of the motors and generators. Synchro differentials (both generators and motors) are also used. These differ from the former electrically in having three rotor windings, symmetrically positioned, instead of a single one of the former.

The windings of a synchro generators are shown schematically in Fig. III.I. An ac voltage E_R applied to the rotor winding R_1R_2 produces an alternating magnetic field which induces voltages in the stator windings OS_1, OS_2 and OS_3. By virtue of the construction of the synchro, each of these voltages will be proportional to the cosine of the angle between the plane of the rotor and of the corresponding stator. If θ is the angle between the plane of the rotor loop and that of OS_2 which we will take as reference, then the three voltage will be

$$E_{S_1} = KE_R \cos (\theta - 120°) . \cos \omega_s t$$

$$E_{S_3} = KE_R \cos (\theta - 240°) . \cos \omega_s t$$

where E_{S_1}, E_{S_2} and E_{S_3} are the voltages in the stators and K is a constant. Note that the three voltages are in phase but their peak amplitudes vary sinusoidally with the angle θ. The sum of the three voltages is easily shown to be zero. Also, if a symmetrical load is put on the stator, the rotor sees a constant load in all its positions.

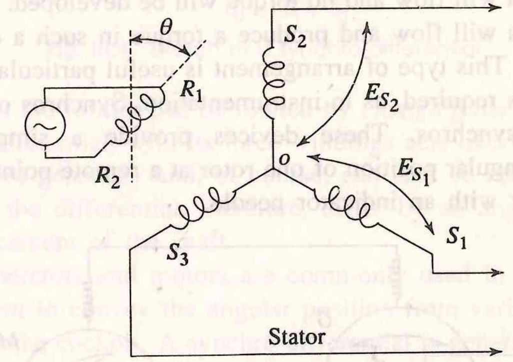


Fig. III.1 Schematic arrangement of a synchro

One form in which two synchros can be used is shown in Fig. III.2. Here, the stator of synchro 2 is the load for the stator of synchro 1. The ac voltage is applied to the rotor of No. 1, while the output is taken from the rotor of 2. In this configuration, the first one is called a synchro-generator and the second one a control-transformer. The currents in the stator coils of the control transformer are identical with the currents in those of the generator and, therefore, a field is produced in the control transformer which has the same orientation as the field in the transmitter,

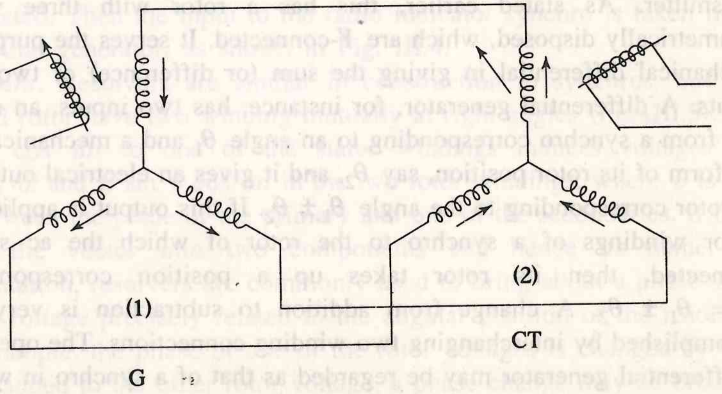


Fig. III.2 Sketch of a synchro-generator and control transformer but is in the opposite direction. The voltage picked up by the rotor (open circuit) of the control transformer then varies sinusoidally with its orientation, being zero when its plane is parallel to the resultant magnetic field. This property is made use of in servo systems by using the rotor output as the error signal.

Another common synchro configuration is shown in Fig. III.3. In this case, power is fed to the rotors of both the synchros. If the rotors are aligned and if the voltages developed in the stators of the two synchros are equal, no current will flow and no torque will be developed. If they are not aligned, currents will flow and produce a torque in such a direction as to align the rotors. This type of arrangement is useful particularly when only a small torque is required, as in instrumentation. Synchros of this type are called torque synchros. These devices provide a simple means of displaying the angular position of one rotor at a remote point, by installing a synchro-motor with an indicator needle.

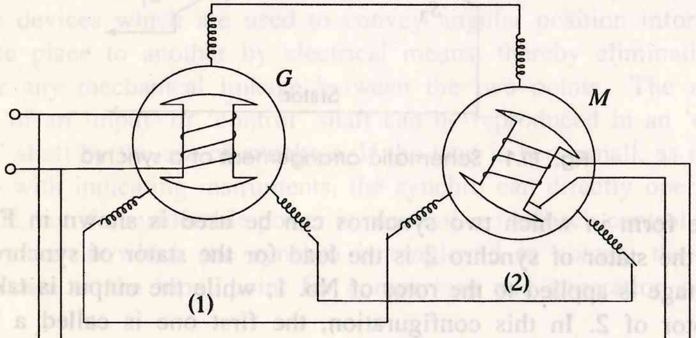
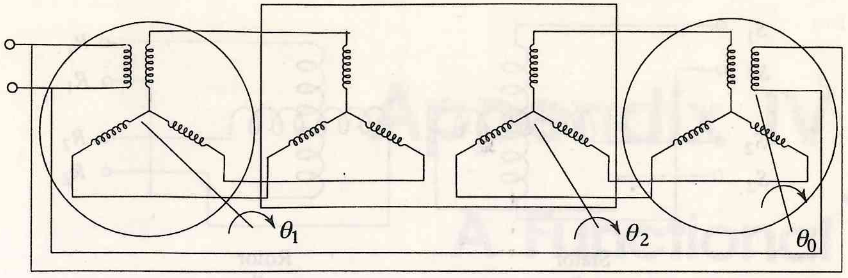


Fig. III.3 Sketch of torque-system

The third configuration shown in Fig. III.4 employs a differential transmitter. As stated earlier, this has a rotor with three windings symmetrically disposed, which are Y-connected. It serves the purpose of a mechanical differential in giving the sum (or difference) of two angular inputs. A differential generator, for instance, has two inputs, an electrical one from a synchro corresponding to an angle θ_1 and a mechanical one in the form of its rotor position, say θ_2, and it gives an electrical output from its rotor corresponding to the angle $\theta_1 \pm \theta_2$. If this output is applied to the stator windings of a synchro to the rotor of which the ac supply is connected, then its rotor takes up a position corresponding to $\theta_0 = \theta_1 \pm \theta_2$. A change from addition to subtraction is very simply accomplished by interchanging two winding connections. The operation of a differential generator may be regarded as that of a synchro in which the rotor and stator can both be rotated. The field produced by the stator of the differential is similar to the one produced by the rotor of a synchro except that the field in the former can be rotated by changing the electrical input. The rotor of differential, with its three windings acts like the stator of the ordinary synchro-generator and, of course, can also be rotated. The input and output of the differential, therefore, differ by an angle equal to the angular displacement of the shaft.



$$\theta_0 = \theta_1 + \theta_2$$

Fig. III.4 Sketch of a synchro-differential

Synchro-generators and motors are commonly used in aircraft navigational equipment to convey the angular position from various instruments to indicators in the cockpit. A synchro differential is generally met with in radio magnetic indicators which are instruments that indicate the magnetic North and the direction of a radio facility with respect to the heading of the aircraft, which is indicated by a fixed line ('Lubber line'). The instrument has two rotors, one of which carries the compass card and the other a pointer. The stator of the magnetic indicator gets its input from a synchro coupled to the gyro-compass. The stator of the radio indicator comes from a synchro associated with the appropriate radio facility. If the latter is a radio-compass, the input can come directly from the synchro coupled to the antenna system. But if a VOR is being used, the equipment gives the bearing of the facility with respect to North and to obtain the bearing with respect to the heading of the aircraft, to heading angle must be subtracted. Then the input to the radio indicator synchro is taken from a differential generator, as shown in Fig. III.4.

Resolvers. Resolvers are similar in construction to synchros, but the stator and rotor have two winding mutually at right angles (Fig. III.5). An input $E \cos \omega t$ to one of the stator windings induces voltages $E \cos \theta \cos \omega t$ and $E \sin \theta \cos \omega t$ in the two rotor windings, where θ is the angle between the plane of the primary and one of the secondaries. It thus resolves the vector into two components and hence its name. In instrumentation, resolvers are commonly used to bring about a phase shift of an ac voltage precisely related to the angular position of the resolver. If, for example, the phase of one of the rotor voltages is changed by 90° and it is added to the other rotor voltage, a phase change may be brought about, as is evident from the following relations.

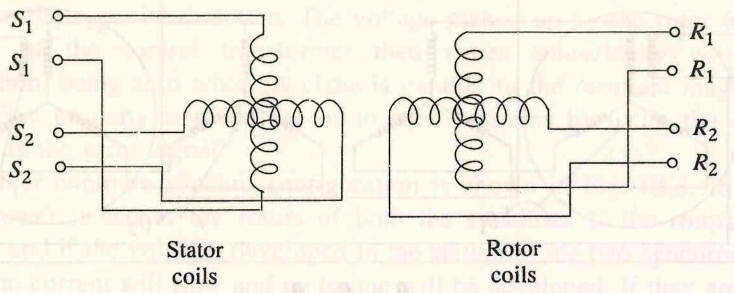


Fig. III.5 Schematic arrangement of a resolver

Let $e_1 = E_s \cos \theta \,.\, \cos \omega_t$ and $e_2 = E_s \sin \theta \,.\, \cos \omega t$ be the outputs of the two stator windings. Let the phase of the former be advanced by $\pi/4$ and that of the latter be retarded by $\pi/4$ by R–C networks, the relative amplitudes being maintained. If the resulting voltages are added, we get a voltage e_0 given by

$$e_0 = E_s \cos \theta \cdot \cos (\omega t + \pi/4) + E_s \sin \theta \cos (\omega t - \pi/4)$$

But $\cos (\omega t - \pi/4) = \cos \left(\overline{\omega t + \pi/4} - \pi/2\right)$

$$= \sin (\omega t + \pi/4)$$

substituting, $e = E \cdot \cos (\omega t + \pi/4 - \theta)$

A phase change of $\pi/4 - \theta$ is thus brought about, which is equal to the angular displacement of the rotor plus a constant.

Appendix IV

A Functional Description of Navigational Facilities

There are many navigational facilites that are either in use or in the process of evaluation which it has not been possible to include in the main part. There are some facilities which are either obsolete or obsolescent but to which references are frequently made in literature. A brief description of some of these are included here. The facilities listed are arranged alphabetically.

BABS This stands for 'beam approach beacon system'. This is a landing aid operating in conjunction with Rebecca interrogator in aircraft. It is a pulsed secondary radar beacon operating in a band at about 220 MHz. By lobeswitching techniques, it gives lateral guidance to aircraft (like the localizer) but has no facility for vertical guidance. But as distance information is available, the pilot can use it, along with the altimeter, to keep to the desired glide-slope. The earlier versions of Rebecca had a cathode-ray indication but this has been replaced by a right-left meter indication in the later models. BABS is not very accurate as a landing aid but is a useful addition to other facilities.

BLEU The letters stand for 'blind landing experimental unit'. This is a blind landing system in which the lateral position of the aircraft is determined by sensing the magnetic fields set up by cables buried in the ground in the approach zone. Position in the vertical plane is determined by a radio altimeter.

CONSOL This is the same as the system which was developed in Germany during the last war called Sonne. It is a rotating beacon operating in the LF/MF band which employs a system of three antennas producing a multi-lobed pattern which is switched to produce a number of equi-signal courses, as in a radio range. Off these courses, either a dot or by counting the number of dots or dashes he receives before obtaining equal signals from both the patterns. This still leaves an ambiguity as there are about 20 lobes in the pattern but this can be resolved. Only a radio receiver is required in the craft to utilize this facility. Useful ranges of 1000 to 2000 km have been obtained with consol.

CONSOLAN This is the same as consol except that a two antenna system is used instead of three-antenna one as in consol.

DECTRA (Decca tracking and ranging) This is a long range hyperbolic nagivational system working at a frequency of about 70 kHz. The system is designed to provide navigation information over a long route, particularly over the sea, as for example across the Atlantic. A master-slave pair is provided at each end of the route and the base line is so short that the hyperbolic lines are almost straight lines. The two pairs of stations help to give the lateral position. In addition, the frequencies of the two pairs are simply related and are phase-locked and comparison of the phase at the sub-harmonic gives a family of hyperbolae with a very long base line intersecting the former family and helps to fix the position of the craft along the route. A feature of Dectra is that the master and slave transmit alternately at the same frequency and a high stability oscillator synchronized to the master station is used for phase comparison.

DELRAC (Decca long range area coverage) This is a very low frequency (10-14 kHz) hyperbolic system which is intended to give world-wide coverage. It is a CW system like Decca, but the transmissions of master and slave stations are made in time sequence and ambiguities are reduced by phase comparison at subharmonic frequencies.

DME (Distance measuring equipment) This is a secondary radar system operating in the UHF band. See Chapter 5.

DMET (Distance measuring equipment TACAN) This refers to the distance measuring part (secondary radar) of the TACAN equipment. The ground installation of DMET provides only distance information to airborne TACAN receiving equipment or DMET receiving equipment (i.e. bearing information is not available). The DMET ground beacon is generally installed with VOR to provide bearing to suitably equipped aircraft. The combination is called VOR/DMET.

GEE This is a hyperbolic system developed during the war in Britain, at the same time as Loran was developed in the United States. It is a pulse system operating in the VHF band and was intended for short distance precision air-navigation.

GCA (Ground controlled approach) See Chapter 6.

ILS (Instrument landing system) See Chapter 6.

LORAN (Standard, SS, LF and C) Loran (standard) or Loran-A Loran-C have been dealt with in Chapter 4. S.S. Loran (skywave synchronized

frequency but had a base line of 1800 to 2500 km. Synchronization by ground wave is not possible for such a separation between stations and so skywave was used for this purpose and hence the name. The chain consisted of two pairs of stations, with the base lines intersecting at right angles. Errors up to 5 nm in fix were encountered. Due to the nature of the propagation, it could be used only at night.

LF Loran (low frequency Loran) was conceived during the war but its development was not completed. It was to operate at about 100 kHz and envisaged cycle matching. It is actually the precursor of Loran-C.

MARKER BEACONS As the name indicates, these are radio beacons which are intended to mark some salient points. The three markers associated with the ILS have been mentioned in Chapter 6. These beacons operate in the VHF band (at 75 MHz). The outer marker has a fan-shaped beam and is sometimes called 'fan-marker'. The others have essentially conical beams. The inner marker is also called 'boundary marker'.

A marker called the Z-marker is used in conjunction with radio ranges. It has also a conical beam and is intended to give an indication to the aircraft when it passes over the radio range.

NAVAGLOBE This is a long-range navigational aid operating in the band centred about 100 kHz which provides bearing information. The ground station has three antennas arranged in an equilateral triangle and power is fed to two of them at a time, cyclically, for 1/4 sec. During the last quarter-second a pair of antennas is energized, one with the normal carrier and the other with a frequency differing by 100 Hz. This last transmission helps to synchronize the receiver and identify the three other transmissions. The bearing is computed from the amplitude of the signals from the three antenna tower pairs.

NAVARHO An integrated rho-theta system consisting of a Navaglobe and Facom. The former gives the bearing and the latter the distance. The distance information is obtained by phase-comparison of two signals, one from the ground stations and the other generated in the aircraft. The phase difference gives the distance that the em wave has travelled from the transmitter, if the phases of two oscillators are suitably set initially. This method calls for a very high degree of stability of the oscillators both in the aircraft and on the ground (at least 1 in 10^9 for 12 hrs) but this can be achieved by modern techniques, particularly the atomic clock. The frequency of the facility being about 100 kHz, phase comparison leads to ambiguity, but this can be resolved by phase comparison at the modulation frequency of 100 Hz.

POPI (Post office position indicator) This is a hyperbolic system with a very short base line which employs phase comparison to obtain bearing. The development of this was undertaken by the British Post Office but it does not appear to have been put to extensive operational evaluation.

RADIO-MAILLES OR RADIO MESH This is navigation system developed in France. It is a hyperbolic system in which the frequency of one station differs from that of the other slightly, so that the space pattern moves with a velocity which depends upon the difference in frequency. Position lines are derived from the number of equi-phase lines that pass the receiver in a given time.

RADUX This is a low frequency hyperbolic system operating at 40 kHz. The Omega system is a development of Radux, the main difference being in the frequencies used. In Radux, modulation frequencies of 200 Hz and 10 kHz are used, the first to reduce ambiguities and the second to provide a vernier.

RADIO SEXTANT This is a sextant operating on the radio frequency emission of heavenly bodies, like a radio telescope. Some equipment working in the microwave band designed to receive the emission from the sun seems to have been built and tested. It could also be possibly used in conjunction with artificial satellites.

REBECCA-EUREKA This is a war time development of secondary radar navigational aid, operating in the VHF band. This is a precursor of DME. The direction of a transponder beacon could be determined in the craft by the use of two antennas and lobe switching technique. It is now obsolete. Eureka was the ground beacon equipment and Rebecca the airborne interrogator.

SHORAN (Short range navigation) This is a secondary radar system in which fix is obtained by the craft, which carries the interrogator, by simultaneously interrogating two ground beacons (transponders).

TACAN (Tactical air navigation) See Chapter 5.

VOR (VHF Omnirange) See Chapter 3.

VOR—DME (See DME)

VORTAC An installation consisting of VOR and TACAN at the same site.

Z-MARKER See under Markers.

References

1. Bowditch, N., *The New American Practical Navigator*, U.S. Hydrographic Office, Washington,1964.
2. Bond, D.S. *Radio Direction Finders*, McGraw-Hill Book Co., New York, 1944.
3. Keen, K., *Wireless Direction Finding*, Iliffe and Sons, London, 1947.
4. Hopkins, H.G. and Pressey, B.G., *"Current direction-finding practice"*, Proc. IEE, Vol. 105(B), supp. 9, 307–316, 1958.
5. *Radio Research Special Report No. 22*, "Siting of direction-finding stations," Dept. of Scientific and Industrial Research, U.K., H.M.S.O. London.
6. Terman, F.E.. *Radio Engineers Handbook*, McGraw-Hill Book Co., New York, 1950.
7. Cleaver, R.F., "Development of single receiver automatic Adcock direction-finger for use in the frequency band 100-150 Mc/s, *J. Inst. Elect. Engr.* London, Vol. 94 (III), 783–797 (1947).
8. Joliffe, S., A.W. "Some factors in the design of VHF automatic direction finders," *The Marconi Review*, Vol. 22 (135), 168–198, 1959.
9. Earp, C.W. and Godfrey, R.M., "Radio direction finding by cyclical differential measurement of phase," *J. Inst. Elec. Eng.* (London), Vol. 94 (IIIA), 705–721, 1947.
10. Earp, C.W. and Cooper-Jones, D.L., "The practical evolution of the commutated aerial direction finding system," *Proc. IEE (London)*, Pt B, Supp. 9, 317–325, March 1958.
11. Lundberg, F.T. and Bucher, F.X., "The cage-type VHF phase-comparison radio range antenna," *Electrical Commun.*, Vol. 29(2), 108–116, June 1952.
12. Hurley, H.C., Anderson, S.R., and Keary, M.E., "The Civil Aeronautics Administration VHF omnirange," *Proc. IRE*, Vol. 39(12), 1506–1520, Dec. 1951.
13. Sandretto, P.C., Electronic Aviation Engineering, *IT & T Corp*, New York, 1958.
14. Bauss, W., *Radio Navigation Systems for Aviation and Maritime use*, Pergammon Press, Oxford, 1963.
15. Strong, C.E., "General aspects of short-range Rho-Theta systems," *Proc. IEE*, Pt B, Supp. 9, 284–306, 1959.
16. Thome, T.G., *Navigation Systems for Aircraft and Space Vehicles*, Pergammon Press, Oxford, 1962.

18. Powell, C., "The Decca navigation system for ship and aircraft use," *Proc. IEE,* Pt B, Supp. 9, 225–234, 1958.
19. Roberts, G.E., "The design and development of Decca flightlog," *J. Brit. IRE,* Vol. 12, 117–131, 1952.
20. Keyton, M. and Fried, W., *Avionic Navigations Systems*, John Wiley, New York, 1969.
21. Moorcraft, G.J., "Precision approach radar," *J. Inst. Elect. Engr*., 15(B), Supp. 9, 344–350, 1958.
22. Skolnik, *M.I., Introduction to Radar Systems*, McGraw-Hill, New York, 1962.
23. Teman, F.E., *Electronic and Radio Engineering*, McGraw-Hill, New York, 1955.
24. McMahon, F.A., "The AN/APN-81 Doppler navigation system," *IRE Trans*., ANE-4, 202–211, Dec. 1957.
25. Pitman, G.R. (Ed.), *Inertial Guidance*, John Wiley, New York, 1962.
26. Slater, *J.M., Inertial Guidance Sensors*. Reinhold Publishing Co., New York, 1964.
27. Scarborough, J.B., *The Gyroscope—Theory and Applications*, Interscience Publishing Co., New York, 1958.
28. Wrighley, W., Hollister, W.M., and Denhard, W.G, *Gyroscopic Theory, Design and Instrumentation*, The MIT Press, Cambridge (Mass), 1969.
29. Parvin. R.H., Inertial Navigation, D.Van Nostrand, New York, 1962.
30. Kalman, R.E., and Declaris, N. (Ed), Aspects of Network and Systems *Theory*, Holt, Reinhart and Winston, 1971.
31. Reintjes, J.F. and Coate, G.T., *Principles of Radar*, McGraw-Hill, New York, 1952.
32. Sonnenberg G.T., *"Radar and Electronic Navigation"*, Butterworth & Co. (1988).
33. Tom Logsdon, *"The Navstar Global Positioning System* (GPO)," Van Nostrand-Reinhold, 1992.
34. Pandharipande, V.M., *"Microwave Landing System"* Telematics (India) April 1990.
35. Howell, J.M., Phased Arrays for microwave landing system, *Microwave journal*, Jan. 1987.

Index