Aerobasics – An Introduction to Aeronautics

14. Air Navigation Principles

S P Govinda Raju

S P Govinda Raju retired as professor from the Department of Aerospace Engineering, Indian Institute of Science in 2003. He is currently active as a consultant in wind tunnel testing and teaches short term courses in aerodynamics and flight mechanics.

Previous parts:
Resonance, Vol.13: p.836, p.971, p.1009, p.1107; Vol.14: p.19, p.191. p.272, p.328, p.650, p.916, p.1071; Vol15: p.64, p.206.

Keywords
Dead reckoning, celestial navigation, radio aids to navigation, instrument landing system (ILS), inertial navigation system (INS), global positioning system (GPS).

Air navigation has evolved over the last hundred years from an art into a mature science. Airplane accidents due to navigational errors are practically nonexistent. In large measure this advance has been due to radio aids to navigation which operate under all weather conditions. However in remote areas and in hostile airspaces, the radio aids may not be available and a fully self contained navigational device is required. Navigation by stars (celestial navigation) is feasible under such conditions. But the inertial navigation system (INS) is the preferred choice. In recent times, a navigational instrument based on radio communication with specially designed satellites in low Earth orbits called the global positioning system (GPS) has become popular. This article deals with the principles of these various methods of air navigation.

1. Introduction

Air navigation, the method of directing the flight of an airplane towards its destination, has evolved into a mature science over the last century. In the early days of aviation, the pilot visually recognized landmarks on the ground and directed the airplane from one landmark to another until he reached his destination. Mistakes in identifying landmarks due to poor visibility led to unscheduled landings or accidents. When landmarks were absent, as over large bodies of water or deserts, dead reckoning was used. In this method, the course of the airplane is plotted on a map of the region from time to time based on calculations using the heading of the aircraft obtained from a magnetic compass and distance traveled using the measured air speed and time of flight. Errors due to instruments or winds often led to serious navigational errors, particularly on long range flights. Navigation based on sky marks (stars in the sky) was used to supplement dead reckoning during such flights and this method is still available for use on very long flights particularly of military transports. This is called 'celestial navigation'. Over time, there have been other aids to navigation based on radio transmission from ground stations specifically designed for navigational purposes. A recent addition to navigational instruments is the Global Positioning System (GPS) which is based on radio communication with satellites specifically designed for this purpose. Navigation based on radio communication may be subject to hostile interference by the enemy. To counter this possibility, combat airplanes use a fully autonomous inertial navigation system (INS) which was originally developed for use in submarines. In what follows, we shall consider the various principles of air navigation in some detail.

2. Celestial Navigation

Stars are luminous objects which are located at very large distances from the Earth. The distances are so large compared with the diameter of the Earth or even its orbit around the Sun that the angular positions of the stars when viewed from anywhere on Earth appear fixed relative to each other. One may imagine the stars to be fixed to a celestial sphere of extremely large diameter with its axis coinciding with the Earth's polar axis. The angles corresponding to latitudes on Earth (*Box* 1) correspond to the declination on the celestial sphere.

If one imagines a taut string joining the centre of the Earth and a star as in *Figure* 1, the point at which the string passes through the Earth's surface is the substellar point or the geographical position (GP) of the

Box 1. Longitudes and Latitudes

The Earth is very nearly a sphere of about 12,700 km in diameter rotating about an axis passing through its centre. The Earth rotates once a day about this axis which has a fixed direction in space. The actual time per revolution (constituting a sidereal day) is shorter than a mean solar day of 24 hours by about 4 minutes. This difference is due to a complete rotation of the Earth around the Sun once a year. The Earth makes 366 revolutions (sidereal days) in a year of 365 solar days. But the difference is not normally important. In what follows, a day is a mean solar day unless it is specially qualified.

Any point on the Earth's surface can be specified in terms of two angles measured at the centre of the Earth – latitude and the longitude as in *Figure* A. The plane perpendicular to the axis of rotation passing through the centre is the equatorial plane. The radial line passing through a given point on Earth makes an angle with the equatorial plane which is called the latitude of that point. The latitude of any point lies in the range of 90 °N to 90 °S. A plane passing through the point and the axis of rotation is the meridianal plane. The meridianal plane through Greenwich in UK is the reference plane of zero longitude. The angle the meridianal plane at any point makes with the meridianal plane through Greenwich is the longitude of the point and varies over 180 °E to 180 °W. Thus any point on the surface of the Earth is specified by the two coordinates (latitude and longitude) pertaining to the location. This is a coordinate system fixed to the Earth.

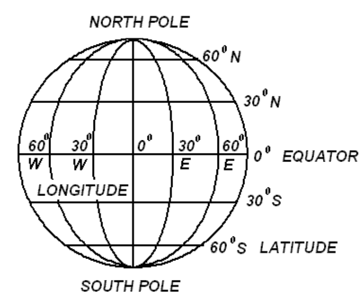


Figure A. Longitudes and latitudes.

star. As the Earth rotates, the GP of a star makes one complete rotation around the Earth in a sidereal day while maintaining a constant declination (latitude). The pole star which is on the axis of rotation of the Earth has its GP at the north pole of the Earth. The GP of any star (as well as the Sun, Moon and planets) at any given time can be estimated from the Air Almanac [1] published annually by the United States government.

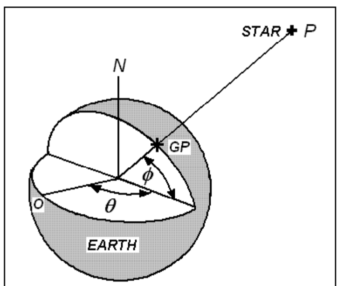
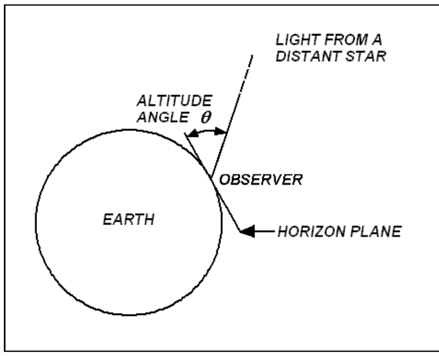


Figure 1. Geographical position of a star: A star is effectively at an infinite distance from the Earth. Its position in the sky can be indicated by its polar angles at the centre of the Earth using the system of longitudes and latitudes exactly like a point on the surface of the Earth. The geographical position (GP) of a star refers to a point on the Earth's surface having the same longitude and latitude as the star. As the Earth rotates, the longitude of the GP of a star varies with time.



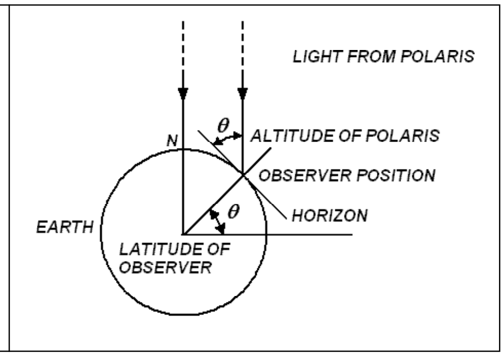


Figure 2 (left). The altitude of a star: The altitude of a star is the angle made by the line of sight of the star with the horizontal plane at an observer's location on Earth. This angle can be measured by an optical instrument called the sextant.

Figure 3 (right). Determination of the latitude of a place by sighting Polaris: Polaris, the pole star, is located on the axis of rotation of the Earth. Thus the altitude of Polaris at any point on Earth is equal to the latitude of the location.

For an observer located at any point on the Earth, the tangent plane at that point is called the horizon. The angle made by the line joining the star and the observer with the horizon is the altitude of the star and is illustrated in *Figure* 2. The altitude of a star at any location can be accurately measured by an optical instrument called a sextant which is essential for celestial navigation. All ships and some long-range airplanes still carry sextants with them.

For navigation using a star, it is necessary to identify the star. Individual stars can be identified easily by their relative location in a group of stars called constellations[1] which form a clear pattern.

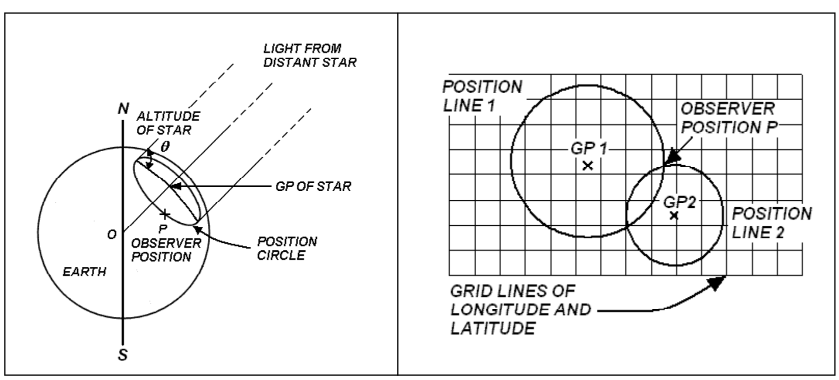
[1] They were recognized and named thousands of years ago. We recognize them by their Greek names even today. Orion, Pegasus, Ursa Major and Cassiopeia are well-known constellations. Polaris, Sirius, Aries and Arcturus are well-known names of stars which are easily located in a night sky. One may refer to a star atlas [2] for the names and positions of all the constellations and stars.

An observer on Earth can easily find his latitude by observing the stars in the sky. It is easily seen from *Figure* 3 that the altitude of the pole star is equal to the latitude of the observer's position on Earth. However, the pole star may not always be visible and it may be more practical to observe other stars. As an illustration, if one can recognize the star at the zenith (overhead, with an altitude of 90°), the GP of that star has the same latitude as the observer. As a further practical measure, observed altitude of any star corresponds to an observer's position on a position circle on the surface of the Earth. The position circle is constructed with its centre at the GP of the star and a radius corresponding to its altitude as illustrated in *Figure* 4. Two such observations lead to an intersection point which fixes the latitude (and longitude) of the observation point as in *Figure* 5.

The longitude of the observer's position can also be found by noting the solar time difference between a reference point on Earth (normally Greenwich whose longitude is zero), and the observer's location. This can be done by using an accurate reference clock set to Greenwich meridian time (reading zero when it is noontime

Figure 4 (left). Location circle of any star with known GP: On measuring the altitude of any star with a known GP, one can construct on the Earth's surface a position circle centered around the GP and with a radius related to the altitude as in the figure. The position circle is the locus of all points on the Earth's surface for which the altitude of the star is the measured value.

Figure 5 (right). Observer's location by the intersection of position lines: It is possible to sight two identified stars and construct two corresponding position circles on the Earth's surface to obtain their intersection point which is the observer's location. The figure illustrates the construction involved on a sheet of paper representing longitudes and latitudes around the observer's location. There are two intersection points and one of them is the correct location of the observer. The ambiguity is resolved by some other method. The grid lines of longitude and latitude and position circles for the two stars in the illustration will be somewhat distorted in the construction on a flat sheet of paper due to the projection of Earth's spherical surface on to a plane. This distortion is not shown in the figure.



at Greenwich). When the Sun reaches the meridian (position of maximum altitude) at the observer's position, it is local noontime. The time on the reference clock then indicates the difference in solar time between the observer and Greenwich. This difference can be related to the longitude of the observer. As a time difference of 24 hours corresponds to 360° of longitude, each hour of time difference corresponds to 15° of longitude. In the olden days, this method of estimating the longitude of the observer was very useful on ships. In fact the need for accurately estimating the longitudinal position of a ship at sea led to the development of accurate clocks or chronometers.

The above method of estimating the position of an observer on Earth was used by airplanes on long-range routes before the invention and perfection of the radio aids to navigation which are more convenient. Celestial navigation is also useful for navigating spacecraft.

3. Radio Aids to Navigation

The need for aids to navigation, particularly at night and during low visibility weather, became obvious during the years of development of civil aviation in the 1920s. Aircraft simply could not navigate under these conditions with any degree of safety. As communication from air to ground is only possible by radio waves, efforts were made to develop aids to navigation using these.

Radio waves are electromagnetic waves in the frequency range of about 100 kHz to 100 MHz. Radio waves at the lower end of this frequency range can travel far without much atmospheric absorption and follow the curvature of the Earth and are well suited for long range communication over thousands of kilometers. Radio waves at the upper end of the frequency band are called Very High Frequency (VHF) waves. They are good for line of sight communication over a few hundred kilometers. Within the above range, a choice of transmission frequency is possible for a specific radio station which can then be identified by its frequency. As radio waves are little affected by weather, all weather communication is feasible.

The earliest radio aid to navigation took the form of a radio beacon which is simply a radio transmitter sending waves in all directions. An aircraft within the range of the transmitter can find the direction from which it receives the waves by using an Automatic Direction Finding (ADF) antenna mounted on the aircraft. The radio beacon is thus analogous to the lighthouse at sea. Further developments led to navigation systems like the Long Range Navigation (LORAN) suitable for navigation over uninhabited areas like the North Polar Region. Systems suited for shorter distances also evolved during this time. The Very high frequency Omni-directional Radio range (VOR) which has been the mainstay of civil aviation over the last fifty years was developed and later enhanced by the addition of the Distance Measuring Equipment (DME). These systems are still in use. We shall briefly describe the principle of operation of these systems.

3.1 Long-Range Navigation

A radio navigation system using a pair of synchronized transmitters spaced several hundreds of miles apart is illustrated in *Figure* 6. It can approximately cover a region whose diameter is the line joining the transmitters. In this system, two transmitters S_1 and S_2 send out radio waves which can be received by the receiver at P (the airplane). The waves have to travel different distances d_1 and d_2 to reach the receiver. This path difference results in a phase difference between the signals received from S_1 and S_2. It is easily shown that lines of constant phase difference are hyperbolas as indicated

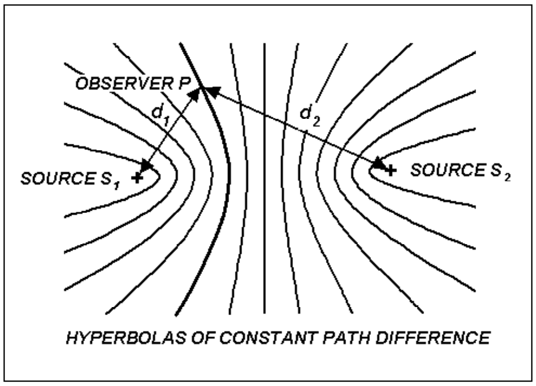


Figure 6. Principle of hyperbolic navigation: Two radio transmitters produce a field in which the lines of constant path difference are hyperbolas as shown. An observer at P receiving the time-marked signals from the two transmitters can measure the difference in arrival times of the two signals and calculate the path difference and fix his position on one of the hyperbolas. Two such systems (four transmitters) operating in the same region will provide sufficient information for uniquely identifying the observer's location.

in the figure. If one can measure the phase difference of the signals, it is possible to fix the position of the receiver on one of the hyperbolas. Two such systems in operation in the same region can fix the position of the receiver.

Various methods were used to implement the navigation principle indicated above. With the development of digital electronics, it became possible to measure the path difference $d_1 - d_2$ directly. In such a system called the LORAN, S_1 sends out a pulse pair with a repetition frequency between 20Hz and 33Hz on a carrier of about 2MHz. The second transmitter synchronized with the first radiates a similar pulse pair but delayed by a fixed time interval. On board an airplane, these signals are received and the difference in the arrival times of the two sets of pulse pairs is accurately measured. As radio waves travel at a constant speed, this time difference can be related to the difference in the relative distances of the two transmitters from the aircraft. This locates the airplane on a specific hyperbolic line. Two such systems operating in the same region are needed for a complete determination of the airplane position.

Radio aids to navigation have advanced in step with advances in electronics over time, the early analogue systems have been replaced by digitial systems of superior performance.

It may be pointed here that instruments based on measurement of small differences in the arrival times of two signals are in wide use. Apart from the LORAN, the DME and GPS systems to be described later in this article use the same technique. Accuracy of such an instrument depends on the accuracy of measurement of small time intervals. Radio waves travel a distance of a kilometer in about 3 microseconds. Thus an error in time measurement of the order of a microsecond leads to an error in distance calculation of 300 meters. With modern electronic techniques, time intervals as short as 10^{-12} seconds are measurable [4]. Thus instruments of good accuracy are easily constructed using the principle of measurement of transit time of radio waves.

3.2 *Short-Range Navigation*

Navigation over populated areas is assisted by VOR stations located generally near airports. A VOR station employs a carrier frequency in the range of 108–118 MHz amplitude modulated at 20–30Hz. The principle of operation of the instrument is given in *Box* 2.

A VOR station sends out an omni-directional reference signal R with a frequency of about 30 Hz. The station also sends a second signal S as in *Figure* 7. The phase of this signal is direction dependent and is synchronized to be in phase with R in the magnetic north direction. An airplane in flight within the range of the station thus receives two signals both of the same frequency but with a phase difference. The phase difference varies over $0° - 360°$ depending on the direction of the line joining

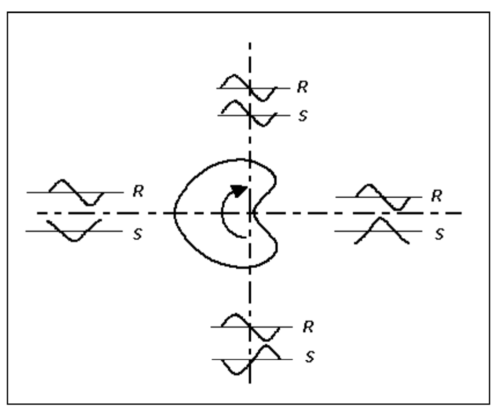


Figure 7. Operating principle of the VOR: A VOR station sends out two low frequency signals on a VHF (very high frequency) carrier. The reference signal R is omni-directional (same in all directions). The second signal S has a definite phase relationship with R depending on the direction of the receiver. A receiver on the airplane receives R and S and compares them to obtain the phase difference. This phase difference is easily related to the direction of the VOR station as viewed from the airplane.

the airplane to the VOR station and is illustrated in *Figure* 7. This phase difference is measured and interpreted by an instrument on the aircraft as the direction of the VOR station at the airplane in relation to magnetic north. This instrument is accurate to better than 0.5 degree. It may be remarked here that when the VOR concept was first introduced, the signal S was generated by a directional antenna rotating about a vertical axis as illustrated in *Figure* 7. Currently the same signal is generated by a pair of fixed antennas spaced apart suitably. One of the antennas radiates at the carrier frequency. The second antenna radiates at a slightly different frequency. The radiations from the two antennas interfere and produce the signal S at the beat frequency (difference between the two antenna feed frequencies). When properly synchronized, the signal so produced has exactly the same form as described earlier.

The VOR system described above provides information about the direction of the VOR station. This is currently complemented by the distance information obtained using the Distance Measuring Equipment (DME) collocated with the VOR station. The DME consists of a transmitter–receiver on the airplane and a matching receiver–transmitter on the airplane. The transmitter on the airplane sends a signal in the form of a pair of pulses. The equipment on the ground is a transponder. On receipt of the pulses, it re-transmits the same after a definite time delay. The receiver on board the airplane receives and compares this signal with the signal sent by it earlier. It measures the time delay between the transmitted and received signals and calculates the distance of the airplane from the DME station. The calculation is based on the known speed of the radio waves and the measured delay at the airplane. Thus the VOR–DME system completely defines the position of an aircraft relative to the ground station. Navigation can be easily effected using VOR–DME stations as ‘waypoints’ (radio

Box 2. Principle of VOR

The principle of the VOR can be illustrated by an optical analogy. Imagine a distant tower on which two lights are installed. The first light flashes red (signal *R*) every 360 seconds and is non directional. Thus a distant observer anywhere around sees a red flash every 360 seconds. The second light produces a focused horizontal beam of light of green colour (signal *S*) but rotates around the tower at a speed of one revolution in 360 seconds. Thus an observer on the ground sees a green flash as the green beam sweeps across him. It can be arranged such that the red flash occurs when the green beam is facing north. Then the distant observer due north sees two flashes every 360 seconds, a red and a green. An observer to the east of the tower sees a red flash followed 90 seconds later by a green flash. An observer to the south of the tower sees a red flash followed by a green flash 180 seconds later. Thus it is seen that the time interval between the red flash and the green flash in seconds indicates the direction of the observer at the tower in degrees from north.

landmarks). They are used in this manner by all the general aviation aircraft.

3.3 *Ground Support for Navigation*

Ground radars are used to monitor well-defined air spaces for airplane traffic. Thus, the ground staff is in a position to advice an airplane by radio about its position as well as the presence of other aircraft in its neighbourhood. Areas near major airports are extensively monitored this way and ground controllers actually control the landing and take-off of aircraft. This ensures safety of flights near airports.

3.4 *Instrument Landing System* (*ILS*)

In performing a landing, an airplane descends slowly along the direction of the runway until it touches the runway at one end and rolls to a stop along the same. This is a critical operation and cannot be performed easily, particularly under conditions of bad weather or poor visibility. The high intensity runway lights which assist in landing are inadequate under such conditions. Additional guidance to the airplane to fly along the correct flight path is often provided by the Instrument Landing System (*Box* 3) available at many airports.

The ILS is a ground-based instrument system involving several radio signals and covers a region extending to about 10 miles along the runway direction. A reference glide path of 3° slope beginning from a point about 1,000 feet down the centre line of the runway is established by four radio beams of carefully designed radiation pattern as in *Figure* 8. Two of these beams provide lateral guidance and constitute the localizer. The other two provide vertical guidance along the glide slope. Guidance to the aircraft approaching the runway is provided by an ILS receiver on the aircraft.

In addition to the localizer and glide path indications,

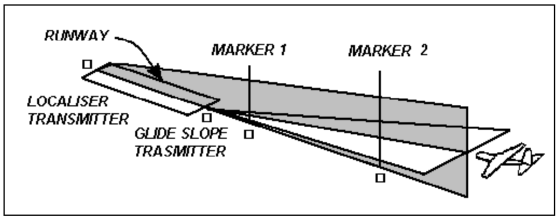


Figure 8. Operating principle of the ILS: The ILS system transmits a pair of signals for the indication of flight path deviation from the designed glide slope of 3º and another pair for the indication of the flight path deviation from the centre line of the runway. The pilot notes the deviations on the ILS instrument on board and controls the airplane to correct the flight path. He finally touches down at the beginning of the runway without error even in poor visibility conditions.

the ILS provides an indication about the airplane location along the glide path by means of three radio markers at distances of about 10 km, 1 km and 200 m from the landing threshold. On some installations, the markers are replaced by a DME system which gives a continuous indication of the distance of the airplane from the landing threshold.

Box 3. Principle of the ILS

The glide slope indicator consists of two strongly directional but overlapping radio beams transmitted in a vertical plane to define a line of 3º slope as in *Figure* A. On the glide slope of 3º the beams have equal intensity. The beams are modulated at 90Hz and 150 Hz over a carrier of about 330 MHz. An airplane flying down a 3º slope will receive the glide slope signal with equal modulation at 90 Hz and 150 Hz. Deviation of the airplane flight path will result in a relative change of modulation and will be indicated as the flight path being above or below the reference inclined plane of 3º slope.

The localizer consists of two overlapping radio beams transmitted from the far end of the runway. They are modulated to a 20% depth at 90 Hz and 150Hz respectively over a carrier of about 110MHz and have equal intensity on the runway centerline. Thus an airplane flying symmetrically with respect to the runway will receive a signal modulated to the same extent at 90Hz and 150Hz. Off the centre line, the signal will have a different modulation at 90Hz as against 150Hz and this difference is interpreted by the ILS receiver as a right or left deviation of the course of the airplane relative to the runway symmetry plane.

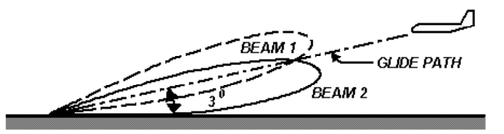


Figure A. The glide path indicator: The guidance of an airplane in the vertical plane during landing is provided by the glide path indicator. Two radio beams carefully designed to provide equal signal strength on the designed glide slope are transmitted from the beginning of the runway. An airplane approaching the runway for a landing receives the two signals and compares them. Error in the glide slope results in one of the signals being larger than the other. This is indicated on the ILS instrument on the airplane.

The ILS system thus effectively guides the pilot to a point along the runway direction until he can see the same and land safely. ILS systems of various accuracies (indicated by supplementary suffixes like Cat1, Cat2, Cat3A, Cat3B, and Cat3C) provide guidance to various distances from the runway threshold. The highest category ILS, Cat3C guides the aircraft all the way to the runway surface and is suitable for very poor visibility weather conditions. It is possible to couple this system to the autopilot on board an airplane to perform a fully automatic landing.

4. Global Positioning System (GPS)

Global Positioning System (GPS) is a complete navigational system based on radio transmission from satellites in medium Earth orbits of about 20,000 km from Earth's centre. The system consists of over 24 satellites in six orbits inclined at about 55° to the equatorial plane of the Earth. Each satellite completes about two orbits in 24 hours. Typically, six satellites are in the line of sight contact with any point on Earth. The navigational system is based on using the satellites as sky marks. Each satellite transmits digital radio signals carrying information about the current position of the satellite and time on the onboard precision clock. The accuracy of the position and time information is monitored and maintained by ground stations in USA, which owns and operates the system.

The GPS receiver is a small, inexpensive device which brings accurate navigation within the reach of everyone. It is used extensively on motor vehicles, unmanned airplanes and missiles.

A GPS receiver is required to use the GPS system. The receiver receives a signal from a satellite and compares the time information in the signal with the time on its internal clock. The difference is interpreted as the travel time for the signal from the satellite to the receiver. This travel time is converted into the travel distance by multiplication with the speed of radio wave propagation. Thus the GPS receiver computes the distance of the receiver from a satellite. The receiver similarly receives information from two other GPS satellites and calculates the distances to these. Knowing the position of the three satellites and their distances, the receiver can obtain a fix on its current location by trigonometric calculations. An additional signal from a fourth satellite is useful in calculating and correcting for any error in the local clock in relation to that on the satellite system.

The accuracy of the GPS position fix is dependent on errors due to refraction of radio signals in the atmosphere. Typical positioning accuracy of a GPS is around 10 meters. Various methods of enhancing the accuracy of GPS locally or over selected areas exist. These go by the names of Local Area Augmentation Systems and are based on ground stations transmitting correction signals to the GPS receivers.

The GPS is useable practically everywhere on Earth and the GPS receiver is small and inexpensive. The system is being introduced for navigation in civil aviation. It can also be used for route finding within cities by motor vehicles and over oceans by ships. The success of the GPS has induced Europe to develop its own GPS by the name Galileo. However the system can be easily jammed by radio disturbances and cannot be depended upon for use on combat airplanes for combat missions.

5. Autonomous Navigation

During conflicts, combat airplanes have to penetrate deep inside hostile territory to attack ground targets. During such missions radio aids to navigation may not be available and the airplanes have to depend entirely on onboard instruments for navigation. A navigation system using inertial sensors with or without Doppler radar (an instrument which measures ground velocity vector of an aircraft) has been developed for use in such situations. A navigational system using only inertial sensors is called an Inertial Navigation System (INS). The

The INS is a totally self-contained navigation system. It is immune to interference from external radio sources.

operating principle of INS follows directly from Newton's laws of classical physics. A body in a state of uniform motion in a straight line continues in that state unless disturbed by an external force which causes acceleration. Thus a uniform motion cannot be detected by any instrument on the body. But the external force which causes acceleration can be measured. Thus change in velocity (magnitude or direction) due to the external force can be calculated by a process of integration of the acceleration. By a further integration of the velocity, the displacement of a body starting from known initial conditions can be calculated.

To implement the above principle on an airplane, one can measure the net external force acting on the airplane. But this is not practicable. Instead one measures the force acting on a small mass moving with the airplane. Such an instrument is called an accelerometer. The instrument consists of a small mass free to move along one direction. Its motion relative to the airplane along this axis is detected by an electrical sensor and nulled (reduced to zero) by applying a force electromagnetically. The current required for this is measured and is an indication of the acceleration of the mass and hence of the airplane itself. Three such units mounted along mutually perpendicular axes are required for completely defining the acceleration vector.

To derive velocities and displacements from the above measurements, the integrations have to be performed with respect to a suitable reference axis system and it is necessary to know in this system the directions of the three accelerometers at every instant. The classical solution to the problem is to mount all the three accelerometers on a stable platform which remains horizontal at all times. Further, a reference line on the platform is made to point north as in *Figure* 9. Thus, integrations with respect to this axis system provide the displacements of the airplane in the Earth-fixed coordinate system of

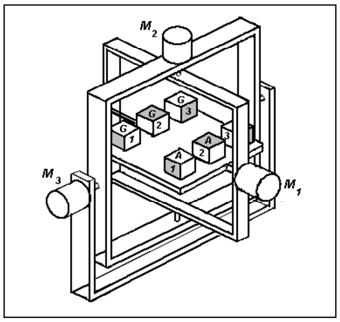


Figure 9. The stabilized platform for inertial navigation: The accelerometers A_1, A_2 and A_3 which form the heart of the instrument are mounted on a platform mounted on gimbals which provide three angular degrees of freedom. The gimbals can be rotated about their respective axes by servo motors M_1, M_2 and M_3. The three gyros G_1, G_2 and G_3 mounted on the platform detect any slight angular displacements of the platform and feed this information to a computer. The computer commands the servo motors to rotate the gimbals so as to keep the platform horizontal at all times.

latitudes and longitudes. This is schematically indicated in *Figure* 10.

The orientation and stabilization of the measurement platform is achieved by mounting the platform on a set of three gimbals (rigid frames each of which is free to rotate about one axis) provided with torque motors as in *Figure* 9. The motors are controlled by a computer using signals from three angular position sensors (called the attitude gyroscopes (*Box* 4) or gyros for short) on the platform. The gyros detect any slight angular displacement of the platform induced by aircraft motion due to maneuvers or other causes. The signals from the gyros are received by a computer which commands the torque motors to torque the gimbals about their respective axes

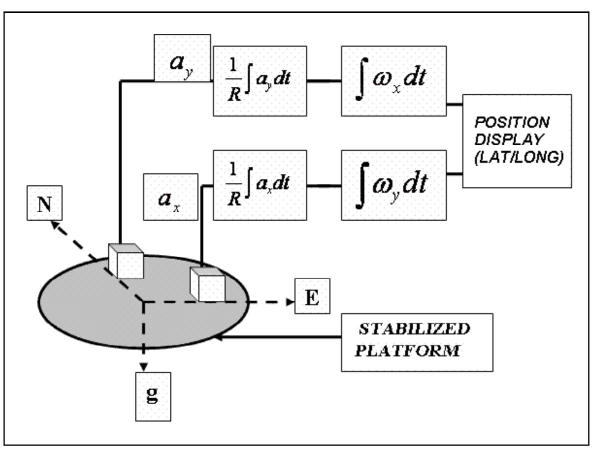


Figure 10. Calculation of the airplane position: The calculation of the airplane position (longitude and latitude) is performed by the computer. The acceleration due east integrated once gives the velocity component in that direction. This quantity divided by the distance of the airplane from the Earth's centre (approximately equal to Earth's radius at sea level) gives the rate of change of longitude. A further integration gives the longitude of the airplane. A similar calculation can be done using the acceleration due north. This gives the latitude of the airplane. A more accurate calculation can add the altitude of the airplane to the Earth's radius at sea level to get the true distance of the airplane from the center of the Earth.

Box 4. The Gyroscope

A rapidly spinning wheel electrically driven to maintain constant speed mounted so that its axis is free to tilt about two axes perpendicular to the axis of rotation as in *Figure* A is called a gyroscope. In this figure, OZ is the axis of rotation and the gyroscope axis can tilt about the OX and OY axes. If such a device is mounted on a vehicle in motion, due to the inertia of the rotating wheel, the axis of rotation of the gyroscope points to a fixed direction in space regardless of the angular motions of the vehicle. Thus a gyroscope can be used for providing a reference direction anywhere including in outer space and is useful for navigation. In such applications, the random drift of the axis of the gyroscope due to its imperfections must be small. Well-constructed gyroscopes have a random drift rate of less than 0.01 degrees per hour. This corresponds to an error of less than 90 degrees in a year and is adequate for airplane navigation where the flights are generally of less than ten hours duration. It may be noted here that the Earth is a giant-sized gyroscope with its axis of rotation pointing towards the Polaris with negligible drift over thousands of years. Celestial navigation in effect uses the Earth as a gyroscope.

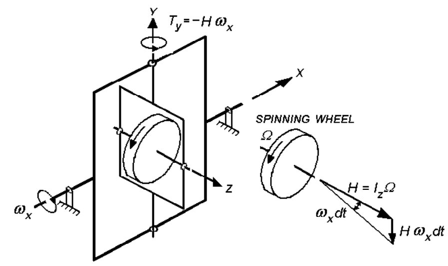


Figure A. The gyroscope: A rapidly spinning wheel whose axis is free to tilt about X and Y axes as in the figure, is a gyroscope. The wheel has an angular momentum H about the wheel axis. When the gyroscope wheel tilts with an angular velocity ω_x about the X axis, there is a rate of change in the angular momentum vector to the extent of $H\omega_x$ about the Y axis. This demands a torque T_y about the Y axis. Gyroscopic instruments are based on the measurement of this reaction torque.

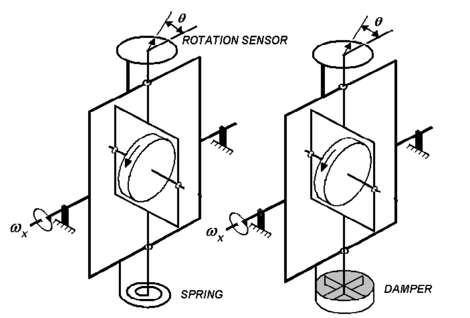
To construct an instrument for navigation using the gyroscope, use is made of an important dynamical property of the gyroscope which is illustrated in *Figure* A. When subjected to a torque T_y about OY, the gyroscope rotates (precesses) about the OX axis with an angular velocity ω_x related to the torque by the relation:

$$T_y = -H\omega_x . \qquad \text{(i)}$$

In (i), H represents the angular momentum of the gyroscope. In a practical gyroscope, H is maintained constant for all time by an electrical drive system which overcomes the frictional resistance to the rotation of the wheel. Equation (i) can be easily derived from first principles. In a time interval δt, the direction of angular momentum vector changes by $\omega_x \delta t$ in the Y direction. This corresponds to a rate of change of angular momentumof $H\omega_x$ which is the torque and thus corresponds to the above equation.

Equation (i) can be looked upon as the input-output relation of a gyroscope. For an input torque T_y, the gyroscope produces an output precession, ω_x. Equivalently, for an input angular velocity ω_x the gyroscope produces an output torque T_y about the Y axis. The second interpretation is useful in constructing gyroscopic instruments.

Box 4. Continued...



(a) Rate gyro (b) Rate integrating gyro

Figure B. Rate and rate integrating gyroscopes: In a gyroscopic instrument, the outer gimbal is attached to the case of the instrument and moves with it. The rotation about the inner gimbal due to the reaction torque T_y is resisted by a spring as in (a) or a damper as in (b). In (a), the rotation of the inner gimbal about its initial position θ is proportional to ω_x. In (b), θ is proportional to the integral over time of ω_z. The rotation θ is sensed by an electrical transducer and is a measure of ω_x in (a) and its integral over time in (b).

For constructing a rate gyroscope to measure the angular velocity ω_x, the output torque is resisted by a spring of rate K as in *Figure* B. The angular displacement about OY is sensed by a suitable non-contact electrical device. This device is schematically indicated as a pointer in the figure. The movement of the pointer is proportional to ω_x. Rate gyroscopes are used in flight control systems of airplanes.

For constructing an altitude sensing instrument called the rate integrating gyroscope, the torque is resisted by a damper of damping constant C. Thus there is an angular velocity about the Y axis of magnitude T_y/C. Integration over time then indicates that a rotation about the input axis OX results in proportional rotation about the output axis OY. This rotation is sensed by an electrical non-contact device as in a rate gyroscope. Rate integrating gyroscopes are used for platform attitude sensing in inertial navigation systems.

so as to null the output of the gyros. This stabilizes the platform against disturbances due to aircraft motion. Further, the computer commands the torque motors to rotate the platform to compensate for the Earth's angular velocity (corresponding to 360° rotation in a sidereal day) as well as rotation of the horizontal plane as the airplane moves around the Earth's spherical surface due to its own motion. The computer thus effectively keeps the platform horizontal and pointing towards geographic north at all times.

All the calculations required for commanding the platform and for obtaining navigational information are performed by the computer about 20 times per second. In addition, the computer also calculates other navigational information like desired heading, wind vectors, distance and time to next way point, etc. On combat airplanes, the computer also calculates the attitude of the airplane which is required for bomb dropping, etc.

To provide inputs for these additional calculations, sensors are provided on the gimbals to sense their angular positions at all times. Any other information required during the calculations is obtained from the air data system of the airplane itself.

The system as shown above is mechanically complex. Successful efforts have been made to eliminate the gimbals and mount the platform directly on the airplane. Such a system is called a 'strap down INS system'. This system is mechanically simple. But this increases the complexity of calculations. Further, the accuracy and range requirements of sensors are altered. Newer types of sensors are called for. One such is the laser gyro which senses the angular velocity by comparing the propagation times of two laser beams going in opposite directions around an optical path in a solid piece of glass.

Gyroscopic instruments as described above have been in use for a few decades now. While they are accurate, they are bulky and mechanically complex. Their position is currently challenged by devices based on nanotechnology (miniature tuning forks) and optics (laser gyros). These promise superior performance with less complexity, weight and volume.

Suggested reading

[1] **_The Air Almanac 2009_, published by the Nautical Almanac Office of the United States Naval Observatory.**

[2] **R M G Inglis, _A new popular star atlas_, Gall and Inglis, 11 Newington Road, Edinburgh, 6th impression, 1967.**

[3] **John Hulbert, _All about Navigating and Route Finding_, Transworld Publishers Ltd, 1974.**

[4] **Vasant Natarajan, V Balkrishnan and N Mukunda, Space and Time in Life and Science, _Resonance_, Vol.13, No.9, pp.843–865, 2008.**

[5] **Earnest O Doebelin, Measurement Systems – Application and Design, Fifth edition, McGraw-Hill Publishing Company, 1990.**

[6] **E H J Pallett, _Automatic Flight Control_, Third Edition, BSP Professional Books, Oxford, 1987.**

Address for Correspondence
S P Govinda Raju
Department of Aerospace Engineering
Indian Institute of Science
Bangalore 560 012, India.
Email: spg@aero.iisc.ernet.in