Getting Your Bearings

The Magnetic Compass and GPS

Richard B. Langley

The magnetic compass has been guiding travelers for a thousand years or more. It is one of the oldest navigation instruments and is still widely used by ship captains, pilots, Boy and Girl Scouts, and hikers. But thanks to modern microelectronic technology, the compass has been reborn. Electronic versions of the compass are available in standalone units, as a component in multisensor navigation systems, and as embedded modules in GPS receivers. Many cars and trucks now come equipped with electronic compasses. And despite their amazing ability to accurately determine a position, single-antenna GPS receivers cannot determine their heading — the direction in which the receiver, or the platform carrying it, is pointed. The compass comes to the rescue! When GPS signals are blocked by buildings, a GPS-based navigation system might have to resort to dead reckoning. Again, the compass pitches in by providing heading information to improve the procedure.

This month's column is a tutorial about these electronic magnetic sensors, how they measure the Earth's magnetic field, and how they complement the capabilities of a GPS receiver.

Humans have always needed to travel — to find food, to explore and conquer new lands, to trade. Navigating safely to a destination and back requires skilled use of knowledge and senses to determine current position and the direction in which to travel to reach one's destination. Navigators of old were guided by landmarks and by the stars or the direction of the Sun at midday. But in unfamiliar territory and under cloudy skies, how does one know in which direction to head off?

It seems that the Chinese first solved this problem. The magnetic properties of loadstone — magnetized rock — were known to the Chinese more than two millennia ago. They fashioned spoons out of loadstone which spun to indicate south — the imperial direction. The first recorded use of loadstone as an actual magnetic compass by the Chinese dates from around 1000 A.D. Arab traders probably brought the compass to the West, where it was first mentioned in the 1187 writings of the English monk Alexander Neckam. Europe's first compasses may have been used by Italian sailors ferrying crusaders to the Levant. Ever since, the compass has been a primary navigational tool for travels on land, on the sea, and in the air.

The basic compass remained relatively unchanged for a thousand years, but the advent of electronics launched new ways of direction sensing using the Earth's magnetic field. But before we look at some of these sensors, let's have a brief look at magnetism, Earth's magnetic field, and how a conventional compass works.

The Force

Magnetism constitutes one-half of the fundamental force known an *electromagnetism*. In 1873, the Scottish physicist James Clerk Maxwell published the theory that explains almost all electrical and magnetic phenomena. For example, it describes how a changing electrical force caused by moving electrical particles (an electric current) can generate a magnetic force (as occurs in an electric motor, for example), and a changing magnetic force can produce an electrical one (as occurs in a generator).

The electrical and magnetic forces acting on a moving charged particle typically vary in space and time. Instead of specifying how these forces vary from place to place, we can use auxiliary quantities independent of the charge and velocity of the particle so that we can describe the potential electrical and magnetic effects even when a charge is not present. These quantities are the electric field, $\boldsymbol{E}$, and the magnetic field, $\boldsymbol{B}$. A field is simply a physical quantity that takes on different values at different points in space. The magnetic field, like the electric field, is a vector field because it has both magnitude and direction.

A field can be represented by a family of field lines. At any point in space, the tangent to the field line gives the direction of $\boldsymbol{B}$ at that point, and the spacing of the field lines gives the magnitude of $\boldsymbol{B}$. The closer together the lines, the stronger the field. Back in high school, most of us visualized such field lines when examining the field of a bar magnet with iron filings.

Electrical devices, iron-bearing minerals, even the human body generate magnetic fields. But superimposed on all these fields is that of the Earth itself.

The Geomagnetic Field

The Earth's magnetic field, called the *geomagnetic field*, results primarily from the conducting fluid outer core. More than 90 percent of the geomagnetic field is generated in the Earth's outer core. This part of the field is known as the *main field*. Convective motions of the iron-rich molten material in the outer core generate the main field by a process known as the *self-exciting dynamo*.

Superimposed on the main field are additional fields caused by magnetized rocks in the Earth's crust, fields generated outside the Earth by electric currents in the ionosphere and magnetosphere caused by the flow of ions and electrons, electric currents flowing in the Earth's crust (usually induced by varying external magnetic fields), and ocean current effects. The magnitudes of these additional geomagnetic fields vary with time and place. The fields generated by the currents in the ionized upper atmosphere and magnetosphere, for example, can be as large as 10 percent of the main field. The time scales of the variations in the total magnetic field vary from a fraction of a second (micropulsations caused by magnetospheric effects and the solar wind) to millions of years (the time scale for the reversals of the main field).

TABLE 1 Approximate magnetic field strengths

Source	Magnitude
Strongest magnet ever made	16 T
Small bar magnet	10^{-2} T
Earth's main field	25,000–65,000 nT
Crustal rocks	$\geqslant$200 nT
Ionosphere/magnetosphere	20–1000 nT
Micropulsations	1–3 nT (occasionally 30 nT)

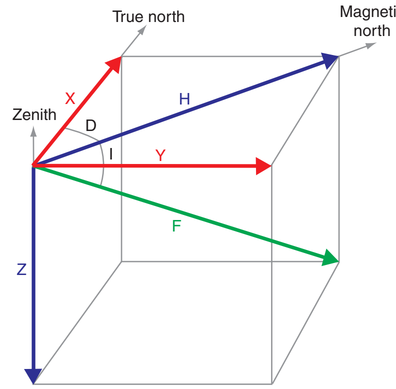
The main field has a high degree of symmetry. In fact, it is similar to the field that would be generated if there were a huge bar magnet centered in the Earth. This field — called a *dipole field* — has an axis of symmetry parallel to the magnet and intersects the Earth's surface in the Arctic at the geomagnetic north pole and in the Antarctic at the geomagnetic south pole. The magnetic field lines emerge from the southern hemisphere and reenter the Earth in the northern hemisphere so that the geomagnetic north pole is actually the south pole of the dipole field. The dipole axis is not parallel to the Earth's rotation axis but is currently inclined about 10 degrees away, so the geomagnetic poles and the geographic poles are not coincident. The geomagnetic north pole is currently near Ellesmere Island in the Canadian arctic archipelago, and the geomagnetic south pole is about 1,000 kilometers south of Australia. I say "currently" because the poles drift slowly with time as a result of the ongoing changes in the main field.

The actual main field is somewhat more complex than a dipole field. Consequently, the field lines of the main field are not perpendicular at the geomagnetic or dipole poles. The field lines are vertical at locations known as the *magnetic poles* or the *magnetic dip poles.* They are displaced some hundreds of kilometers from the geomagnetic poles.

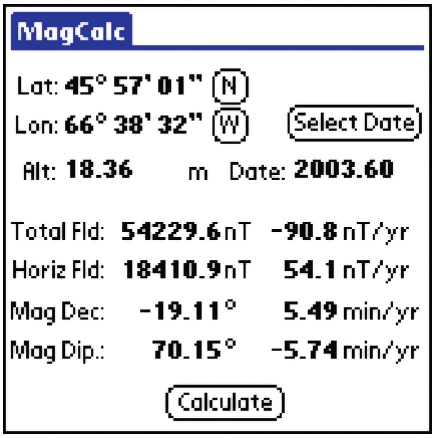
At any point in space, the geomagnetic field vector ***B*** can be described by its magnitude and direction, by three orthogonal components in a selected coordinate frame, or by a set of related quantities. It is convenient to describe the direction of the field in terms of horizontal and vertical components in a local reference frame whose x-axis points to astronomical or geographic north (commonly referred to as *true north*), whose z-axis points to the nadir (the local direction of gravity), and whose y-axis completes a right-handed system. Alternatively, one could use a geodetic reference frame and corresponding reference ellipsoid.

The angle between the horizontal component of ***B*** and true north is called the *(magnetic) declination* (D) or *(magnetic) variation.* The angle between ***B*** and the horizontal plane is called the *(magnetic) inclination* (I) or *(magnetic) dip.* D and I are measured in units of degrees, positive east for D and positive down for I. The magnitude or intensity of the total field (sometimes represented by (F) can be partitioned into a horizontal component (H) and a vertical component (Z), as shown in **Figure 1.** In addition, the horizontal component can be further partitioned into north (X) and east (Y) components. These components may be measured in units of oersted (Oe; equivalent to units of gauss in free air) but are generally reported in nanotesla (nT) (1 Oe = 10^5 nT). The Earth's magnetic field intensity is roughly between 25,000 and 65,000 nT (0.25 - 0.65 Oe). See Table 1 for the strengths of some other magnetic fields.

Models. The Earth's main field can be modeled mathematically. Two such models are the International Geomagnetic Reference Field (IGRF) and the World Magnetic Model (WMM). The IGRF is computed by a group of researchers associated with the International Association of Geomagnetism and Aeronomy. The IGRF models the field and its secular variation using a set of spherical harmonic coefficients (called *Gauss coefficients* in recognition of the development of this technique for use in studies of geomagnetism by the nineteenth-century German polymath Carl Friedrich Gauss) in a truncated series expansion of a geomagnetic potential function and its time derivative. The magnetic field is the gradient of this potential. The coefficients are computed from sets of measurements of field intensity from sites around the globe as well as from satellite observations. The IGRF is updated every five years, and the current model is known as *IGRF2000.* Field values for dates before 2000 can be computed by linear interpolation using the coefficient values from previous models or their updates. Values for dates between 2000 and 2005 are computed using the epoch 2000 coefficients and their secular variations.



▲ **FIGURE 1** The geomagnetic field parameters or elements are the declination (D), inclination (I), intensity of the total field (F), horizontal field component (H), and the north (X) and east (Y) field components.



Lat: 45° 57' 01" N

Lon: 66° 38' 32" W

Alt: 18.36 m Date: 2003.60

Total Fld: 54229.6 nT -90.8 nT/yr

Horiz Fld: 18410.9 nT 54.1 nT/yr

Mag Dec: -19.11° 5.49 min/yr

Mag Dip.: 70.15° -5.74 min/yr

▲ **FIGURE 2** World Magnetic Model values of the main field elements and their rates of change at Fredericton, New Brunswick, Canada.

The WMM coefficients for 2000 to 2005 were produced jointly by the British Geological Survey and the U.S. Geological Survey on behalf of the British Hydrographic Office and the National Imagery and Mapping Agency. The WMM mathematical model is similar to that used for the IGRF.

Figure 2 shows the output of an implementation of WMM running on a personal digital assistant. If declination is computed for the whole globe, a map of declination contours can be constructed as shown in **Figure 3.**

The field models account for only the long-wavelength spatial magnetic fluctuations that originate in the Earth's core. They do not account for intermediate and short-wavelength variations that originate in the Earth's mantle and crust. Consequently, isolated declination and inclination errors of

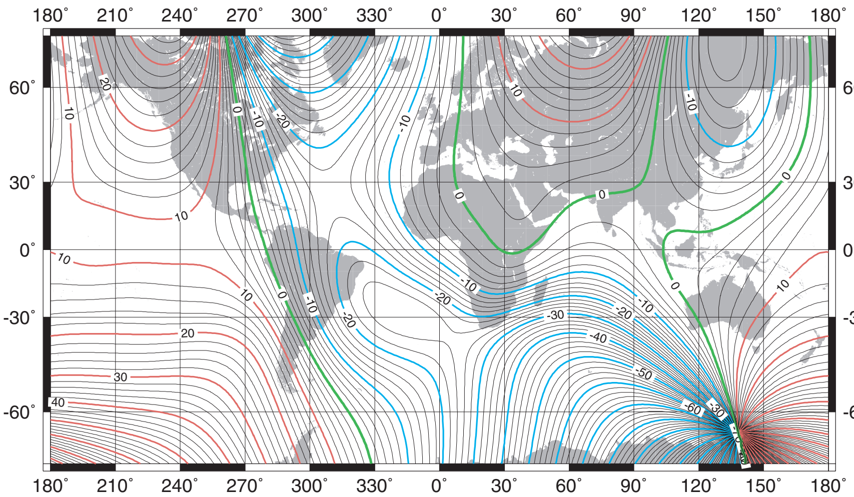


FIGURE 3 The contours on this Mercator-projection map indicate the declination at the beginning of year 2000 computed from the World Magnetic Model.

several degrees or more can exist at various positions on the Earth's surface (primarily over land, in continental margins, and over oceanic seamounts, ridges, and trenches). In fact in a few places, because of local geological formations, declination and inclination errors can reach 50 degrees or more! In oceanic areas, however, the root-mean-square declination and inclination errors of the models (within the applicable time span) are approximately 0.5 degrees.

Most GPS receivers include the IGRF or WMM (or tables derived from these models) in their firmware for converting true courses and bearings into magnetic ones (see sidebar, "Finding Your Way"). The accuracy of the conversions will depend on the inherent errors in the models and whether or not they are being used within the period of applicability. Eventually, the model values will become out of date, and larger errors will result unless the receiver is supplied with new firmware that includes updated model values or a user-supplied declination.

The Needle Compass

So how does a conventional compass work? A compass is nothing more than a magnetized needle supported by a low-friction pivot that allows it to freely rotate. The needle aligns itself so that it is parallel to the magnetic field lines in its immediate vicinity. If the needle is supported in such a fashion that it is free to rotate in both the horizontal and the vertical plane, the direction of the needle will indicate both the declination and the inclination of the local geomagnetic field. To keep the needle horizontal so that it will accurately indicate magnetic north, it is usually weight-balanced for the geographical region in which it is to be used. Some manufacturers balance their compasses for one of five different regions. Models with special global balancing can be used all over the world. A needle compass is usually filled with a liquid such as a mixture of water and alcohol or clear oil to damp needle motion and prevent oscillation as the compass is moved.

Electronic Compasses

In many applications, electronic devices are now replacing the ancient magnetized-needle technology, which is subject to errors from external sources such as vibration, tilt, acceleration, and extraneous magnetic fields as we have already discussed. Also, the traditional compass is not readily adaptable to digital readout or computer interface and hence is difficult to integrate into a navigation system.

Most electronic compasses are based on sensors that are magnetometers. A magnetometer is a device for measuring the intensity of one or more components of the Earth's magnetic field. It is inherently an analog device, responding linearly or nonlinearly to changes in the magnetic field it senses. Its response can be sampled for input into a microprocessor. Magnetometers and their associated electronics were originally very bulky devices, but thanks to advances in semiconductor technology, they have shrunk in size — some are even incorporated into integrated circuits.

A basic electronic compass may be fabricated by mounting two magnetometers at right angles on a flat horizontal support. Each sensor measures one component of the horizontal field — one along the support's x-axis and one along its y-axis. If we represent the components as B_x and B_y, then the angle between the x-axis and the direction of the horizontal field (indicating magnetic north) is given by

$$\psi = \arctan\left(\frac{B_y}{B_x}\right) \qquad (1)$$

(see **Figure 4**).

Several different types of magnetometers operate on different principles, and I will briefly describe four of the most common types: the fluxgate, Hall-effect sensor, magnetoinductive sensor, and the magnetoresistive sensor.

Fluxgate. A fluxgate is essentially a transformer with a core made of an alloy whose magnetic domains are easily aligned (saturated), such as nickel-iron (permalloy). If the current in the primary winding (known as the *bias* or *drive coil*) is varied, the current in the secondary or *sense coil* will vary in a manner that depends on the ambient magnetic field. The basic principle for measuring this field is to compare the drive-coil

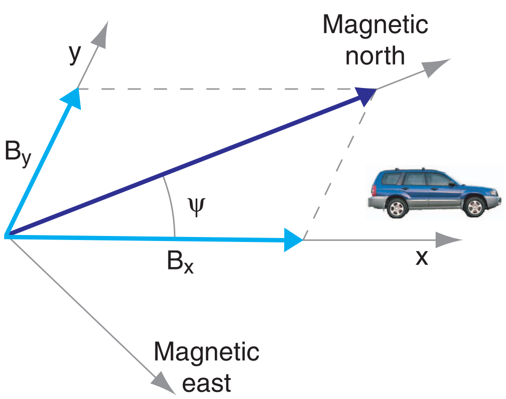


FIGURE 4 If the direction in which a vehicle is pointing is represented by the x-axis, then its heading with respect to magnetic north, ψ, is given by the arctangent of the horizontal component of the geomagnetic field measured along the y-axis divided by that measured along the x-axis.

current needed to saturate the core in one direction with that in the opposite direction. The difference results from the ambient field. Full saturation is not necessary; any nonlinearity will suffice.

The fluxgate core can be a rod or a toroid. If a rod is used, the drive and sense coils are wound side by side or on top of each other, and the fluxgate responds to the component of the ambient field along the rod. One can use a toroidal core to sense two orthogonal field components (dual axis) by winding two sense coils around the outside of the core at right angles to each other.

Victor Vacquier invented the fluxgate magnetometer in the 1930s while he was employed at Gulf Research Laboratories. His pioneering device led to the development of magnetometers for a variety of applications, including geophysical exploration for minerals and hydrocarbons, detecting submarines, and monitoring and mapping the Earth's magnetic field from sensors on the ground, in aircraft, and in satellites.

Construction of a compass requires a dual-axis fluxgate. One sense coil provides the component of the Earth's field along one axis (let's call it the x-axis), and the other provides the component along the y-axis. The angle between magnetic north and the x-axis is just the arctangent of the y value divided by the x-value (as in Equation 1). Clearly the fluxgate must be horizontal to accurately determine direction. Any tilting of the sensor such as would occur on a rolling vessel will reduce accuracy. To avoid such loss of accuracy, the sensor should be tilt compensated, for example with a gimbal, so that it always remains horizontal. Both fixed and gimbaled fluxgate sensors are readily available, as are complete compass modules, and some have industry-standard National Marine Electronics Association–formatted output.

A tilt-insensitive compass can be constructed from a three-axis fluxgate sensor coupled with an electronic tilt meter. Such a compass has no moving parts and is known as a *strapdown compass.*

Hall-Effect Sensor. In 1879, Edwin Hall, then a graduate student at Johns Hopkins University, discovered that if a current is passed lengthwise through a thin conductor in the presence of a magnetic field, a small voltage develops across the width of the conductor. The effect was little more than a scientific curiosity until the development of semiconductors that permitted the construction of Hall-effect integrated circuits with transistors to amplify the weak

Finding Your Way

A *bearing* is the horizontal angle between a line from one point to another measured from a reference direction, usually north, clockwise from 0 through 360 degrees. The reference direction may be true, compass, magnetic, or grid north, or an arbitrary specified direction. A *true bearing* is referenced to true north, which is the direction of the north geographic pole. A *compass bearing* is a bearing relative to north as indicated by an uncorrected compass as affected by deviation. If a compass bearing is corrected for deviation, the result is a *magnetic bearing.* Magnetic and true bearings differ due to declination, and one can be converted to the other if the declination is known. A bearing can also be referenced to the grid system used to portray maps or charts. Grid north is indicated by the direction of the vertical grid lines (for a map oriented with north "up").

A commonly used map grid system is Universal Transverse Mercator (UTM). Grid north is the same as true north at the center of each six-degree UTM zone. Its direction changes to each side of the central meridian because the rectangular grid does not follow the convergence of meridians toward the pole. A *grid bearing* is simply a bearing referenced to grid north. Maps often indicate the directions and relationships of true, magnetic, and grid north allowing conversion of bearings of one type to another. Most GPS receivers can portray bearings as true, magnetic, or grid using stored information on declination and stored grid and datum parameter values.

Sometimes bearings are given in a range of 0 to 90 or 0 to 180 degrees, in which case the appropriate quadrant or semicircle must be indicated. For example, the bearing N 40° W is 40 degrees west of north or equivalently 320 degrees. Occasionally the term *bearing angle* is used to designate such quadrant bearings. Bearings can also be given in mils, a common military practice. The mil is based on the milliradian. However, there are approximately 6,283 milliradians in a circle and as this is a rather unwieldy number, a mil is usually defined as 1/6400 of a circle. Bearings can also be measured in grads, a common European practice. A grad is 1/100 of a right angle or, in other words, there are 400 grads in a circle.

Frequently the term *azimuth* is used synonymously with bearing. However, some navigation purists prefer to reserve the term azimuth to describe the position of an astronomical object on the celestial sphere and use the word bearing to refer to terrestrial objects.

The horizontal direction in which a platform is actually moving (or its intended direction) expressed as an angular distance measured clockwise from north is called the *course.* In other words, it is the bearing or azimuth of a line along which a platform travels. In marine use, the term strictly refers to the direction through the water, ignoring the effects of currents and other sea motion. The direction with respect to the ground is called the *track.* However, the terms course and track are often used interchangeably, particularly in land and air navigation. A course may be designated as true, compass, magnetic, or grid according to the reference direction.

In navigation, the terms course and track can be applied to either an actual or desired path. To remove this ambiguity, the terms *course made good* and *track made good* are used to describe the direction from a previous position to the actual current position. Once again, these terms are often used interchangeably, ignoring the motion of the air or water through which the platform may be traveling. A third term, course over ground, is also sometimes used to refer to the current direction of the platform. Note that these terms are not necessarily the same as the *heading* or direction in which the platform is pointing. Due to winds or currents, a platform may have to be pointed at an angle with respect to the desired track.

signal voltage. The integrated circuits also included circuitry to reduce signal drift from temperature and supply voltage fluctuations and other factors.

Several electronic devices based on the Hall effect are available, including Hall-effect switches of various kinds and linear Hall-effect sensors. When a Hall-effect switch is exposed to a magnetic field of the correct polarity greater than the operating threshold, the output transistor is switched on, permitting current to flow through it. When the field intensity drops below the threshold, the transistor is switched off.

One can construct a simple electronic compass indicating the nearest cardinal or intercardinal point by arranging four Hall-effect switches in a circle around a miniature freely rotating magnetized needle. When the needle points to one of the switches, its output transistor is switched on, permitting the flow of current and illuminating, for example, a light-emitting diode (LED) that indicates the direction or heading of the compass. If the needle points midway between two switches, then both switches will turn on. This action can illuminate the LED at the intercardinal point.

More-precise compasses require linear Hall-effect sensors with a voltage output that accurately tracks the changes in the ambient magnetic field. In the quiescent state (no magnetic field), the output is ideally equal to one-half of the supply voltage over the operating voltage and temperature ranges of the device. An increasing south-pole magnetic field will increase the output voltage from its quiescent voltage. Conversely, an increasing north-pole field will decrease the voltage from its quiescent value. A pair of sensors orthogonally mounted in the horizontal plane can determine, to a precision of a few degrees or better, the direction in which a magnetized needle points.

Magnetoinductive Sensor. The magnetoinductive magnetometer measures a magnetic field by sensing its effect on the inductance of a wire coil or solenoid. The coil is used as the inductive element in an inductor/resistor relaxation oscillator. As the ambient field changes, so does the coil's inductance. This response changes the oscillator's frequency, which can be measured and interpreted in terms of the strength of the field component parallel to the coil's axis. Magnetoinductive sensors are relatively new, with the first patent issued in 1989.

As with the other kinds of magnetic sensors, two orthogonally mounted sensors in the horizontal plane are needed to determine the direction of the horizontal field and hence magnetic north. A two-axis device could be gimbaled to keep it horizontal, or a three-axis device coupled with a tilt sensor could be used. Many vehicle compasses now contain magnetoinductive sensors.

Magnetoresistive Sensor. Anisotropic magnetoresistive (AMR) sensors are special resistors made of a permalloy thin film deposited on a silicon wafer. During manufacture, a strong magnetic field is applied to the film to orient its magnetic domains in the same direction, establishing a magnetization vector. Subsequently, an external magnetic field applied perpendicularly to the side of the film causes the magnetization vector to rotate and change angle. This in turn causes the film's resistance to vary. If the AMR device is included in an electrical circuit such as a Wheatstone bridge, then the change in resistance can be detected as a voltage change and the strength of the applied magnetic field inferred. William Thompson, Lord Kelvin, first described the magnetoresistive effect in 1856.

Magnetoresistive sensors, which come in one-, two-, and three-axis designs, can be made very small. For example, a three-axis sensor is available in a package with a footprint of only 2.8 millimeters by 8.1 millimeters, with a height of 4.0 millimeters. These miniature solid-state low-power sensors can be assembled into a standalone electronic compass or embedded into other products. Properly calibrated, these electronic compasses can have an accuracy of better than one degree. The compass in some GPS receivers is based on magnetoresistive sensors.

Calibration

A compass responds to the vector sum of the Earth's field plus all disturbing fields. Depending on their strengths, these fields can significantly reduce the accuracy of a compass. Induced magnetism in ferrous objects such as iron and steel ("soft iron") in the vicinity of the compass will distort the ambient magnetic field, as will objects that may have acquired permanent magnetism ("hard iron"). Even car speakers and the electrostatic discharge from nylon clothing can affect a compass. Consequently, the direction in which the magnet of a compass actually points, called *compass north*, will in general be different from magnetic north. The angular difference is called *(magnetic) deviation*. If the disturbing field is constant, the compass can be adjusted, or calibrated, to account for the disturbing field.

A conventional compass in a fixed location on a platform such as a ship can be adjusted to offset the deviation by placing small magnets and/or pieces of ferrous material at specific locations around the compass. Such compass adjustments can be difficult and time consuming. Alternatively, the compass can be simply calibrated by noting the compass errors at a series of known headings — a technique known as *swinging the compass*.

Electronic compasses must also be calibrated to correct for deviation and other possible errors such as scale factor and misalignment errors. Although one could use the compass-swinging procedure to calibrate electronic compasses, it has some major disadvantages. First, the technique requires known headings, and such headings might not be readily available. Second, the deviation depends in part on the local geomagnetic field strength. If a calibration is performed at one location, it might not be valid at another. Third, this procedure is not appropriate if the electronic compass consists of three orthogonal sensors.

Researchers at Stanford University have developed an alternative procedure that does not rely on reference headings and is not location dependent. The calibration method is based on the fact that the locus of error-free measurements from two perpendicularly mounted sensors is a circle. That is, as the sensor unit is rotated in a circle, the output of each sensor should be a sinewave with the same magnitude and with one output 90 degrees out of phase with the other. If B_H is the magnitude of the horizontal component of the geomagnetic field, B_x and B_y are the horizontal components in a frame attached to the platform, say a vehicle, carrying the unit, and ψ is the vehicle's heading, then

$$B_x^2 + B_y^2 = B_H^2 \cos^2\psi + B_H^2 \sin^2\psi = B_H^2 \quad (2)$$

This is the equation of a circle with its center at the origin of the frame. The radius of the circle is equal to the local horizontal component of the geomagnetic field. It can usually be well approximated by one of the global field models.

The effect of deviations and other errors will be to distort the circle, changing its shape and/or the location of its center. Hard-iron effects add a constant magnitude field component along each axis of the sensor output, which results in a shift of the circle's center, whereas soft-iron effects and sensor scale errors change the circle into an ellipse. The net result is an off-centered ellipse represented by the equation

$$\left(\frac{\hat{B}_x - \delta B_x}{1 + s_x}\right)^2 + \left(\frac{\hat{B}_y - \delta B_y}{1 + s_y}\right)^2 = B_H^2 \qquad (3)$$

in which $\hat{B}_x$ and $\hat{B}_y$ are the sensor outputs, δB_x and δB_y are the offsets to the geomagnetic field caused by hard-iron effects, and s_x and s_y are scale factors to account for sensor scale errors and soft-iron effects. It is assumed that there is no sensor alignment error (the sensors are orthogonal in the horizontal plane). If such an error exists, then the ellipse will be rotated so that its axes are not parallel to the body-fixed x and y axes.

Equation 3 contains four unknown quantities or parameters. These parameters can be estimated with a nonlinear mathematical technique from measurements obtained while the sensor module is rotated in the horizontal plane. Although in principle only four measurements are needed to determine the values of the four unknowns, more observations will result in more-accurate calibration. It is not necessary to rotate the module through a full 360 degrees, but the larger the section of the ellipse sampled, the smaller is the effect of measurement noise. In fact, depending on the sensitivity of the module, one might need to rotate it through a complete turn or even two turns to adequately calibrate it. Once determined, the parameter values can be stored in memory and used to correct future compass measurements.

This calibration technique, or a variant of it, is typically incorporated into the microprocessor controller of commercial compass modules as well as in GPS receivers sporting a compass. The recommended calibration procedure for a particular device may suggest carrying out the calibration far away from metal objects, in which case only scale variation and the effect of its housing are accounted for.

The calibration procedure can be extended to a triad of orthogonally mounted sensors. In this case, the parameters of an ellipsoid are estimated, and the sensor module must be rotated in both the horizontal and vertical planes.

Some electronic compasses feature a self-calibration capability. A tiny magnetic field is generated within the device and varied under microprocessor control. The response to the varying field is used to calibrate the compass.

GPS

A GPS receiver can provide accurate information about its position and velocity, and from the velocity vector it can determine the direction in which it is moving — called the *course* or *track*. The accuracy with which a receiver can compute this direction depends on its speed (the velocity magnitude) but is usually better than one degree for speeds greater than about 10 kilometers per hour. The course is not necessarily the same as the heading or direction in which the GPS receiver, or the platform on which it is mounted, is pointing. A single-antenna GPS receiver cannot determine heading. However, a compass can provide this information and as mentioned earlier, some GPS receivers incorporate an electronic compass, usually a two-axis sensor. Some receivers have three-axis sensors that give relatively accurate bearings even if they are slightly tilted.

As previously mentioned, the firmware in a GPS receiver contains a model of the main field — whether or not it has an embedded electronic compass — from which it can compute the declination value at the receiver's current position. So the receiver can convert a course referenced to true north to one referenced to magnetic north. Likewise, it can convert true bearings to magnetic ones.

FURTHER READING

For a popular-level introduction to magnetism and the discovery of the compass, see

- *The Riddle of the Compass: The Invention That Changed the World,* by A.D. Aczel, published by Harcourt, Inc., New York, 2001.
- *Latitude and the Magnetic Earth,* by S. Pumfrey, published by Icon Books Ltd., Duxford, England, 2002.

For brief discussions of the basics of geomagnetism and geomagnetic models, see

Geomagnetic Field Frequently Asked Questions, <http://www.ngdc.noaa.gov/seg/potfld/faqgeom.shtml>.

For online calculation of geomagnetic field values and maps of field elements, see

- National Geomagnetic Information Center, <http://geomag.usgs.gov/frames/ngic.htm>.

For a review of magnetic field sensors, see

- "A New Perspective on Magnetic Field Sensing," by M.J. Caruso, T. Bratland, C.H. Smith, and R. Schneider in *Sensors,* Vol. 15, No. 12, December 1998, pp. 34–46. An on-line version is available at <http://www.sensorsmag.com/articles/1298/mag1298/main.shtml>.

For a discussion of low-cost magnetic sensors for compass applications, see

- "Applications of Magnetic Sensors for Low Cost Compass Systems," by M.J. Caruso <http://www.ssec.honeywell.com/magnetic/datasheets/lowcost.pdf>.

For information about Stanford University's electronic compass calibration procedure, see

- "A Non-Linear, Two-Step Estimation Algorithm for Calibrating Solid-State Strapdown Magnetometers," by D. Gebre-Egziabher, G.H. Elkaim, J.D. Powell, and B.W. Parkinson, published in *Proceedings of the International Conference on Integrated Navigation Systems,* St. Petersburg, Russia, May 28–30, 2001, pp. 290–297. An on-line version is available at <http://waas.stanford.edu/~wwu/papers/gps/PDF/demozins201.pdf>.

For an introduction to navigation with a GPS receiver, see

- "Navigation 101: Basic Navigation with a GPS Receiver," by R.B. Langley in *GPS World,* Vol. 11, No. 10, October 2000, pp. 50–54.

A GPS receiver's compass allows the use of a navigation technique known as *sight and go*. The receiver displays a compass ring with a pointer. Holding the receiver horizontally near eye level, the user lines up two sighting marks on the receiver case and the pointer with a distant object and instructs the receiver to "lock in" the direction to the object. The receiver then continuously updates the object's bearing as the user follows an arbitrary path toward the object.

The compass module in a GPS receiver usually can be switched off to conserve power. The receiver can also be configured to switch its display from compass heading to GPS-derived course once a certain speed, 5 or 10 kilometers per hour for example, is achieved. The receiver can be set to revert to the compass when the speed drops below the threshold for a preset number of seconds.

A single-frequency GPS receiver also must know something about the geomagnetic field to compute an estimate of the delay experienced by GPS signals traversing the ionosphere. The so-called broadcast model requires the geomagnetic latitude of the point of intersection of the range vector to a GPS satellite and the shell assumed to contain all the ionosphere's electrons. The geomagnetic latitude is computed to sufficient accuracy as

$$\phi_m = \phi_i + 11.6°\cos(\lambda_i - 291°) \qquad (4)$$

in which ϕ_i and λ_i are the geodetic latitude and longitude of the ionospheric intersection point. This conversion is based on the dipole field as it existed in the mid-1900s.

Conclusion

With millions of GPS receivers in use around the globe, the venerable compass is still an important navigational tool. Whether relying on a simple needle compass or an electronic compass embedded in the receiver, GPS users can always determine their bearings whether they are standing still or are on the move.

Acknowledgment

Figure 3 was produced by the United States Geological Survey (USGS) Geomagnetism Program and, together with other information about geomagnetism, is available from the USGS Web site <http://geomag.usgs.gov/>.

"Innovation" is a regular column featuring discussions about recent advances in GPS technology and its applications as well as the fundamentals of GPS positioning. The column is coordinated by **RICHARD LANGLEY** of the Department of Geodesy and Geomatics Engineering at the University of New Brunswick, who appreciates receiving your comments and topic suggestions. To contact him, see the "Columnists" section on page 2 of this issue.