3

FUNDAMENTALS OF INERTIAL NAVIGATION

An inertial system does for geometry . . . what a watch does for time.[1]
—Charles Stark Draper

In Draper's analogy quoted above, watches keep track of time by being set to the correct time then incrementing that time according to the inputs from a "time sensor" (a frequency source) to update that initial value.

Inertial systems do something similar, only with different variables—and they increment doubly. They need to be set to the correct position and velocity. Thereafter, they use measured accelerations to increment that initial velocity and use the resulting velocities to increment position.

3.1 CHAPTER FOCUS

The overview of inertial navigation in Section 1.3 included some of the history of the technology and examples of the more popular sensor designs.

The focus here is on how these sensors are integrated into a navigation system, including the following:

[1]Quoted by author Tom Pickens in "Doc Gyro and His Wonderful 'Where Am I?' Machine," *American Way Magazine*, 1972.

Global Navigation Satellite Systems, Inertial Navigation, and Integration, Third Edition.
Mohinder S. Grewal, Angus P. Andrews, and Chris G. Bartone.

1. terminology for the phenomenology and apparatus of inertial navigation
2. mathematical models for calibrating and compensating sensor errors to improve accuracy
3. methods for determining the different calibration parameters used
4. models and methods used for calculating unsensed gravitational accelerations
5. methods for determining initial conditions of attitude, velocity, and position
6. methods for integrating sensed attitude rates and accelerations
7. computer requirements for implementing these methods.

How this all affects navigation performance is covered in Chapter 11.

3.2 BASIC TERMINOLOGY

The following is a breakdown of some terminolgy used throughout the book. An expanded standardized terminology for inertial sensors may be found in Ref. 4, and that for inertial systems in Ref. 6.

Inertia is the propensity of bodies to maintain constant translational and rotational velocity, unless disturbed by forces or torques, respectively (Newton's first law or motion).

Inertial reference frames are coordinate frames in which Newton's laws of motion are valid. They cannot be rotating or accelerating. They are not necessarily the same as the **navigation coordinates**, which are typically dictated by the navigation problem at hand. Our problem is that we live in a rotating and accelerating environment here on Earth, and that defines the coordinate system we are already familiar with. "Locally level" coordinates used for navigation near the surface of the earth are rotating (with the earth) and accelerating (to counter gravity). Such rotations and accelerations must be taken into account in the practical implementation of inertial navigation.

Inertial sensors measure inertial accelerations and rotations, both of which are vector-valued variables.

Accelerometers are sensors for measuring inertial acceleration, also called **specific force** to distinguish it from what we call "gravitational acceleration." The point is, **accelerometers do not measure gravitational acceleration**. What accelerometers measure is modeled by Newton's second law as $a = F/m$, where F is the physically applied force (not including gravity) and m is the mass it is applied to. The force per unit mass, F/m, is called **specific force**, and accelerometers are sometimes called **specific force receivers**.

Gyroscopes (usually shortened to "gyros") are sensors for measuring rotation.

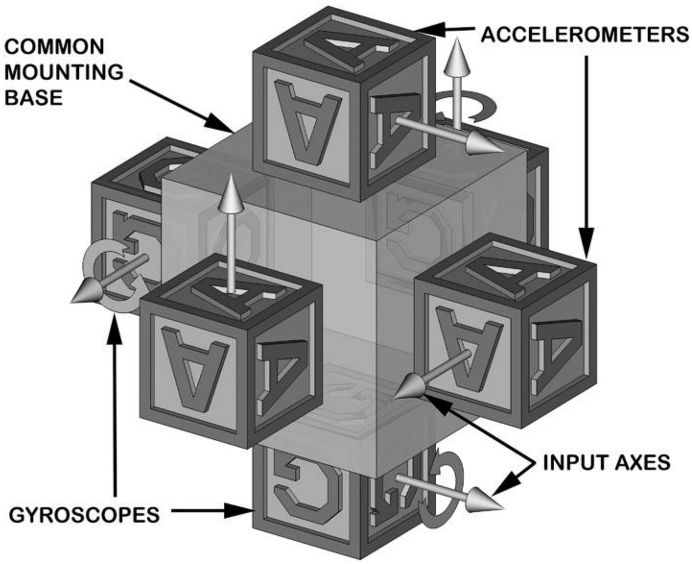

Fig. 3.1 Inertial sensor assembly (ISA) components.

Rate gyros measure rotation rates.

Displacement gyros (also called **whole-angle gyros**) measure accumulated rotation angles. Inertial navigation depends on gyros for maintaining knowledge of how the accelerometers are oriented in inertial and navigational coordinates.

Input axes of an inertial sensor define which vector components of acceleration or rotation rate it measures. These are illustrated by the arrows in Fig. 3.1, with rotation arrows wrapped around the input axes of gyroscopes to indicate the direction of rotation. Multiaxis sensors measure more than one component.

Calibration is a process for characterizing sensor behavior by observing input/output pairs, usually for the purpose of compensating sensor outputs to determine the sensor inputs.

Inertial sensor assemblies (ISAs) are ensembles of inertial sensors rigidly mounted to a common base to maintain the same relative orientations, as illustrated in Fig. 3.1. ISAs used in inertial navigation usually contain three accelerometers and three gyroscopes, represented in the figure by lettered blocks with arrows representing their respective input axes, or an equivalent configuration using multiaxis sensors. However, ISAs used for some other purposes (e.g., dynamic control applications such as autopilots or automotive steering augmentation) may not need as many sensors, and some designs use redundant sensors. Other terms used for the ISA are **instrument cluster** and (for gimbaled systems) **stable element** or **stable platform**.

Inertial reference unit (IRU) is a term commonly used for an inertial sensor system for attitude information only (i.e., using only gyroscopes). Space-based telescopes, for example, do not generally need accelerometers, but they do need gyroscopes to keep track of orientation.

Inertial measurement units (IMUs) include ISAs and associated support electronics for calibration and control of the ISA. Support electronics may also include thermal control or compensation, signal conditioning, and input/output control. An IMU may also include an IMU processor, and—for gimbaled systems—the gimbal control electronics.

Inertial navigation systems (INSs) measure rotation rates and accelerations, and calculate attitude, velocity, and position. Its subsystems include

IMUs, already mentioned above.

Navigation computers (one or more) to calculate the gravitational acceleration (not measured by accelerometers) and process the outputs of the accelerometers and gyroscopes from the IMU to maintain an estimate of the position of the IMU. Intermediate results of the implementation method usually include estimates of velocity, attitude, and attitude rates of the IMU.

User interfaces, such as display consoles for human operators and analog and/or digital data interfaces for vehicle guidance and control functions.

Power supplies and/or raw power conditioning for the complete INS.

Implementations of INSs include two general types:

Gimbaled systems use their gyroscopes for controlling ISA attitude. Commonly used ISA orientations include

Inertially stable (nonrotating), a common orientation for operations in space. In this case, the ISA may include one or more star trackers to correct for any gyroscope errors. However, locally level implementations may also use star trackers for the same purpose.

Locally level, a common orientation for terrestrial navigation. In this case, the ISA rotates with the earth and keeps two of its reference axes locally level during horizontal motion over the surface. Some early systems aligned the gyro and accelerometer input axes with the local directions of north, east, and down, because the gimbal angles could then represent the Euler angles for heading (yaw), pitch, and roll of the vehicle. However, there are also advantages in allowing the locally stabilized element to physically rotate about the local vertical direction.

Strapdown systems do nothing to physically control the orientation of the ISA, but they do process the gyroscope outputs to keep track of its orientation.

Gimbaled systems are generally more expensive than strapdown systems, but their performance is usually better. This is due, in part, to the fact that their gyroscopes and accelerometers are not required to endure high rotation rates.

Both gimbaled and strapdown systems commonly use some form of **shock and vibration isolation** to keep mechanical disturbances within the host vehicle frame from harming the sensors or the navigation implementation.

Because INSs perform integrals of acceleration and attitude rates, these integrals need initial values.

Alignment is a procedure for establishing the initial ISA attitude with respect to navigation coordinates. This can be done using external optical reference directions. However, systems with sufficiently accurate sensors can perform **self-alignment** when the system is stationary with respect to the earth. In that case, the implementation can be divided into two parts:

Leveling The implementation uses the accelerometers to measure the upward acceleration required to counter gravity, from which the system can determine the orientation of its ISA relative to local vertical. For gimbaled systems, the stable element (ISA) is physically leveled during this process (hence the name).

Gyrocompassing A procedure for estimating the direction of the earth's rotation axes with respect to ISA coordinates, using its gyroscopes. This and the direction of the local vertical then determine the north–south direction, so long as the stationary location is not in the vicinity of the poles. Given these two directions, the INS can orient itself relative to its location on the earth. The term *gyrocompassing* is a reference to the gyrocompass, an instrument introduced toward the end of the nineteenth century to replace the magnetic compass on iron ships. The gyrocompass uses only mechanical means to orient itself relative to north, whereas the INS requires a computer. For some gimbaled systems, gyrocompassing physically aligns the ISA with its level sensor axes pointing north and east.

Initialization is a procedure for establishing the initial position and velocity of the INS. Some of this can be done autonomously when the system is stationary with respect to the earth, in which case its relative velocity is zero in earth-fixed coordinates. The angle between the local vertical and the measured rotation axis of the earth, determined during alignment, can be used to estimate latitude. However, longitude and altitude must be determined by other means.

Host vehicle is a term used for the moving platforms on or in which an INS is mounted. It could be a spacecraft, aircraft, surface ship, submarine, land vehicle, or pack animal (including humans).

3.3 INERTIAL SENSOR ERROR MODELS

Inertial navigation has been called "navigation in a box" and "black-box navigation" because it is entirely self-contained. It infers what is going on outside by what it can sense inside.

Inertial sensors are also "black boxes" for the same reason. The problem is, there may be more going on outside the sensor than just accelerations and rotations,[2] as illustrated in Fig. 3.2. For the purpose of inertial navigation, one needs to know the sensor inputs, given the outputs. That is the art of inertial sensor modeling.

Mathematical models for how inertial sensors perform are used throughout the INS development cycle including the following:

1. In designing sensors to meet specified performance metrics.
2. To calibrate and compensate for fixed errors, such as scale factor and output bias. The extreme performance requirements for inertial sensors may not be attainable within manufacturing tolerances. Fortunately, the last few orders-of-magnitude improvement in performance can often be achieved through calibration. These calibration models are generally of three types:
   (a) Models based on engineering data and the principles of physics, such as the models carried over from the design trade-offs. These models generally have a known set of possible causes for each observed effect.

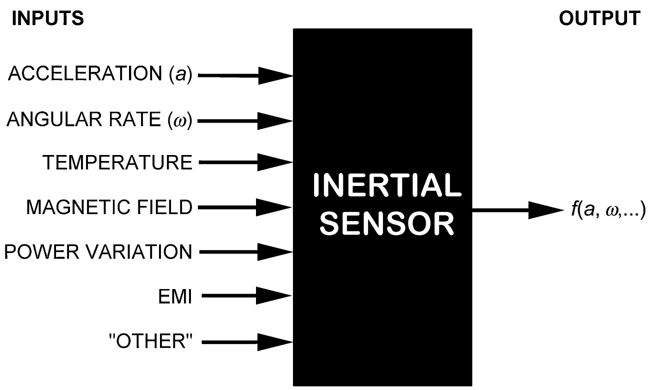

Fig. 3.2 Sensor black-box model.

[2]A comment often heard from inertial sensor designers is "No matter what sort of sensor we design, it always turns out to be a highly sensitive thermometer!"

(b) Abstract, general-purpose mathematical models such as polynomials, used to fit observed error data in such a way that the sensor output errors can be effectively corrected.

(c) Models for unpredictable variations in sensor output, used for predicting sensor and system performance.

3. Additional error models used in global navigation satellite system (GNSS)/INS integration for determining the optimal weighting (Kalman gain) in combining GNSS and INS navigation data.
4. Sensor models used in GNSS/INS integration for recalibrating the INS continuously while GNSS data are available. This approach gives the INS improved initial accuracy during periods of GNSS signal outage.

3.3.1 Zero-Mean Random Errors

These are the standard types of error models from Kalman filtering, used for modeling unpredictable outputs and described in Chapter 10.

3.3.1.1 White Sensor Noise This is usually lumped together under "electronic noise," which may come from power supplies, intrinsic noise in semiconductor devices, or from quantization errors in digitization.

3.3.1.2 Exponentially Correlated Noise Temperature sensitivity of sensor bias will often look like a time-varying additive noise source, driven by external ambient temperature variations or by internal heat distribution variations.

3.3.1.3 Random Walk Sensor Errors Random walk errors are characterized by variances that grow linearly with time and power spectral densities that fall off as $1/\text{frequency}^2$ (i.e., 20 dB per decade).

There are specifications for random walk noise in inertial sensors, but mostly for the integrals of their outputs and not in the outputs themselves. For example, the "angle random walk" from a rate gyroscope is equivalent to white noise in the angular rate outputs. In a similar fashion, the integral of white noise in accelerometer outputs would be equivalent to a "velocity random walk."

The random walk error model has the form

$$\varepsilon_k = \varepsilon_{k-1} + w_{k-1}$$

$$\begin{aligned}\sigma_k^2 &\stackrel{\text{def}}{=} \langle \varepsilon_k^2 \rangle \\ &= \sigma_{k-1}^2 + \langle w_{k-1}^2 \rangle \\ &= \sigma_0^2 + kQ_w \text{ for time-invariant systems}\end{aligned}$$

$$Q_w \stackrel{\text{def}}{=} \underset{k}{E} \langle w_k^2 \rangle.$$

The value of Q_w will be in units of squared error per discrete time step Δt. Random walk error sources are usually specified in terms of standard deviations, that is, error units per square root of time unit. Gyroscope angle random walk errors, for example, might be specified in $\text{deg}/\sqrt{\text{h}}$. Most navigation-grade gyroscopes (including RLG, HRG, IFOG) have angle random walk errors in the order of 10^{-3} $\text{deg}/\sqrt{\text{h}}$ or less.

3.3.1.4 Harmonic Noise Temperature control schemes (including building HVAC systems) often introduce cyclical errors due to thermal transport lags, and these can cause harmonic errors in sensor outputs, with harmonic periods that scale with device dimensions. Also, suspension and structural resonances of host vehicles introduce harmonic accelerations, which can excite acceleration-sensitive error sources in sensors.

3.3.1.5 "1/f" Noise This noise is characterized by power spectral densities that fall off as $1/f$, where f is the frequency. It is present in most electronic devices, its causes are not well understood, and it is usually modeled as some combination of white noise and random walk.

3.3.2 Fixed-Pattern Errors

These are repeatable sensor output errors, unlike the zero-mean random noise considered above. The same types of models apply to accelerometers and gyroscopes. Some of the more common types of sensor errors are illustrated in Fig. 3.3. These are

(a) bias, which is any nonzero sensor output when the input is zero;
(b) scale factor error, usually due to manufacturing tolerances;
(c) nonlinearity, which is present in most sensors to some degree;
(d) scale factor sign asymmetry (often from mismatched push–pull amplifiers);
(e) a dead zone, usually due to mechanical stiction or lock-in (for ring laser gyroscopes); and
(f) quantization error, inherent in all digitized systems; it may not be zero-mean when the input is held constant, as it could be under calibration conditions.

We can recover the sensor input from the sensor output so long as the input/output relationship is known and invertible. Dead-zone errors and quantization errors are the only ones shown with this problem. The cumulative effects of both types (dead zone and quantization) often benefit from zero-mean input noise or dithering. Also, not all digitization methods have equal cumulative effects. Cumulative quantization errors for sensors with frequency outputs are bounded by ± one-half least significant bit (LSB) of the digitized output,

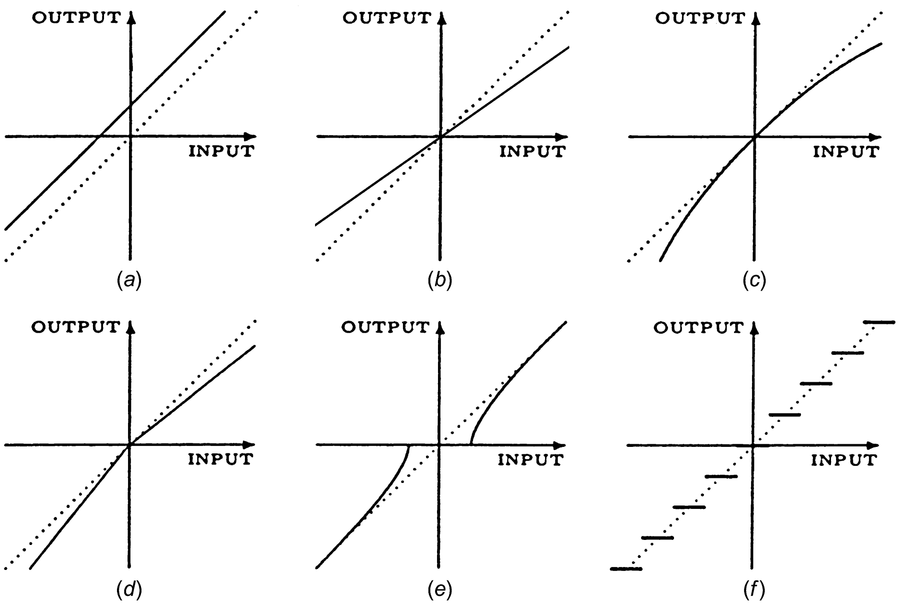

Fig. 3.3 Common input/output error types. (*a*) Bias. (*b*) Scale factor. (*c*) Nonlinearity. (*d*) ±Assymetry. (*e*) Dead zone. (*f*) Quantization.

but the variance of cumulative errors from independent sample-to-sample A/D conversion errors can grow linearly with time.

3.3.3 Sensor Error Stability

Outright sensor failure is a serious problem in inertial navigation, and sensor reliability is always addressed in system design. Also, much attention has been given to detecting and correcting sensor failures during operation. Slow sensor accuracy degradation is another problem.

In practice, fixed-pattern sensor errors do not necessarily remain fixed over long periods of time (hours to years). Some of this may be due to second-order sensitivities to ambient conditions (e.g., temperature, barometric pressure, humidity, power levels, and magnetic fields) or may be attributed to "aging." Truly second-order effects can be calibrated and compensated, but compensation requires additional sensors for the second variable, and calibration adds cost.

Navigation errors due to sensor instability can be compensated to some degree by integrating INS with other sensor systems, including GNSS. This requires models for the expected patterns of sensor degradation, a subject addressed in the next section.

3.4 SENSOR CALIBRATION AND COMPENSATION

Sensor compensation is the process of recovering the sensor inputs from the sensor outputs.

Sensor calibration is the process of determining the parameters of the compensation model.

3.4.1 Sensor Biases, Scale Factors, and Misalignments

This part of sensor compensation can be done using an affine (linear plus offset) model. The biases are offsets and the rest is linear.

3.4.1.1 Compensation Model Parameters The model used here is for ISA-level calibration. Calibration can also be done at the sensor level, but it is less expensive if it is done at the ISA level.

The changes in sensor input/output patterns due to biases and scale factors are illustrated in Fig. 3.3. Figure 3.4 illustrates how input axis misalignments and scale factors at the ISA level affect sensor ouputs, in terms of how they are related to the linear input/output model:

$$\mathbf{z}_{\text{output}} = \mathbf{M}\left(\mathbf{z}_{\text{input}} + \mathbf{b}_z\right) \tag{3.1}$$

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}, \tag{3.2}$$

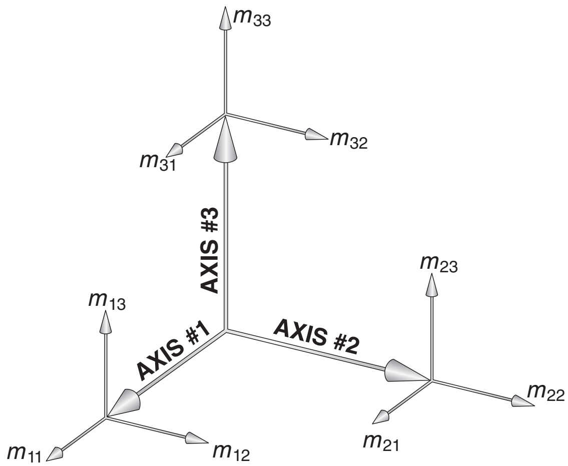

Fig. 3.4 Directions of modeled sensor cluster errors.

where $\mathbf{z}_{\text{input}}$ is a vector representing the inputs (accelerations or rotation rates) to three inertial sensors with nominally orthogonal input axes, $\mathbf{z}_{\text{output}}$ is a vector representing the corresponding outputs, $\mathbf{b}_z$ is a vector of sensor output biases, and the corresponding elements of $\mathbf{M}$ are labeled in Fig. 3.4.

The parameters m_{ij} and $\mathbf{b}_z$ of this model can be estimated from observations of sensor outputs when the inputs are known, the process called **calibration**.

The purpose of calibration is **sensor compensation**, which can be accomplished by inverting the "forward model" of Eq. 3.1 to obtain

$$\mathbf{z}_{\text{input}} = \mathbf{M}^{-1}\mathbf{z}_{\text{output}} - \mathbf{b}_z, \tag{3.3}$$

the sensor inputs compensated for scale factor, misalignment, and bias errors.

This result can be generalized for a cluster of $N \geq 3$ gyroscopes or accelerometers; the effects of individual **biases**, **scale factors**, and **input axis misalignments** can be modeled by an equation of the form

$$\underbrace{\mathbf{z}_{\text{input}}}_{3\times 1} = \underbrace{\mathbf{M}^{\dagger}_{\text{scale factor \& misalignment}}}_{3\times N} \underbrace{\mathbf{z}_{\text{output}}}_{N\times 1} - \underbrace{\mathbf{b}_z}_{3\times 1}, \tag{3.4}$$

where $\mathbf{M}^{\dagger}$ is the Moore–Penrose pseudoinverse of the corresponding $\mathbf{M}$, which can be determined by calibration.

3.4.1.2 Calibrating Sensor Biases, Scale Factors, and Misalignments In this case, calibration amounts to estimating the values of $\mathbf{M}^{\dagger}$ and $\mathbf{b}_z$, given input–output pairs $[\mathbf{z}_{\text{input},\ k}, \mathbf{z}_{\text{output},\ k}]$, where $\mathbf{z}_{\text{input},\ k}$ is known from controlled calibration conditions and $\mathbf{z}_{\text{output},\ k}$ is recorded under these conditions. For accelerometers, controlled conditions may include the direction and magnitude of gravity, conditions on a shake table, or those on a centrifuge. For gyroscopes, controlled conditions may include the relative direction of the rotation axis of Earth (e.g., with sensors mounted on a two-axis indexed rotary table), or controlled conditions on a rate table.

The full set of input/output pairs under K sets of calibration conditions yields a system of $3K$ linear equations

$$\underbrace{\begin{bmatrix} z_{1,\text{input},1} \\ z_{2,\text{input},1} \\ z_{3,\text{input},1} \\ \vdots \\ z_{3,\text{input},K} \end{bmatrix}}_{3K \text{ knowns}} = \underbrace{\begin{bmatrix} z_{1,\text{output},1} & z_{2,\text{output},1} & z_{3,\text{output},1} & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}}_{Z,\text{ a } 3K\times(3N+3) \text{ matrix of knowns}} \underbrace{\begin{bmatrix} m_{1,1} \\ m_{1,2} \\ m_{1,3} \\ \vdots \\ b_{3,z} \end{bmatrix}}_{3N+3 \text{ unknowns}} \tag{3.5}$$

in the $3N$ unknown parameters $m_{i,j}$ (the elements of the matrix $\mathbf{M}^{\dagger}$) and 3 unknown parameters $b_{i,z}$ (rows of the 3-vector $\mathbf{b}_z$), which will be overdetermined for $K > N + 1$. In that case, the system of linear equations may be solvable for the $3(N+1)$ calibration parameters by using the method of least squares:

$$\begin{bmatrix} m_{1,1} \\ m_{1,2} \\ m_{1,3} \\ \vdots \\ b_{3,z} \end{bmatrix} = \left[\mathbf{Z}^{\mathbf{T}}\mathbf{Z}\right]^{-1} \mathbf{Z}^{\mathbf{T}} \begin{bmatrix} z_{1,\text{input},1} \\ z_{2,\text{input},1} \\ z_{3,\text{input},1} \\ \vdots \\ z_{3,\text{input},K} \end{bmatrix}, \tag{3.6}$$

provided that the matrix $\mathbf{Z}^{\mathbf{T}}\mathbf{Z}$ is nonsingular.

The values of $\mathbf{M}^{\dagger}$ and $\mathbf{b}_z$ determined in this way are called **calibration parameters**.

Estimation of the calibration parameters can also be done using Kalman filtering, a by-product of which would be the covariance matrix of calibration parameter uncertainty. This covariance matrix is also useful in modeling system-level performance.

3.4.2 Other Calibration Parameters

3.4.2.1 Nonlinearities Sensor input–output nonlinearities are generally modeled by polynomials:

$$z_{\text{input}} = \sum_{i=0}^{N} a_i z_{\text{output}}^i, \tag{3.7}$$

where the first two parameters a_0 = bias and a_1 = scale factor. The polynomial input–output model of Eq. 3.7 is linear in the calibration parameters, so they can still be calibrated using a system of linear equations—as was used for scale factor and bias.

The generalization of Eq. 3.7 to vector-valued inputs and outputs includes all the cross-power terms between different sensors, but it also includes multidimensional data structures in place of the scalar parameters a_i. Such a model would, for example, include the acceleration sensitivities of gyroscopes and the rotation rate sensitivities of accelerometers.

3.4.2.2 Sensitivities to Other Measurable Conditions Most inertial sensors are also thermometers, and part of the art of sensor design is to minimize their temperature sensitivities. Other bothersome sensitivities include acceleration sensitivity of gyroscopes and rotation rate sensitivities of accelerometers (already mentioned above).

Compensating for temperature sensitivity requires adding one or more thermometers to the sensors and taking calibration data over the expected operational temperature range, but the other sensitivities can be "cross compensated" by using the outputs of the other inertial sensors. The accelerometer outputs can be used in compensating for acceleration sensitivities of gyroscopes, and the gyro outputs can be used in compensating for angular rate sensitivities of accelerometers.

3.4.2.3 Other Accelerometer Models *Centrifugal Acceleration Effects* Accelerometers have input axes defining the component(s) of acceleration that they measure. There is a not-uncommon superstition that these axes must intersect at a point to avoid some unspecified error source. That is generally not the case, but there can be some differential sensitivity to centrifugal accelerations due to high rotation rates and relative displacements between accelerometers. The effect is rather weak but not always negligible. It is modeled by the equation

$$a_{i,\text{centrifugal}} = \omega^2 r_i, \tag{3.8}$$

where ω is the rotation rate and r_i is the displacement component along the input axis from the axis of rotation to the effective center of the accelerometer. Even manned vehicles can rotate at $\omega \approx 3\,\text{rad/s}$, which creates centrifugal accelerations of about $1\,g$ at $r_i = 1\,\text{m}$ and $0.001\,g$ at $1\,\text{mm}$. The problem is less significant, if not insignificant, for microelectromechanical system (MEMS)-scale accelerometers that can be mounted within millimeters of one another.

Center of Percussion Because ω can be measured, sensed centrifugal accelerations can be compensated, if necessary. This requires designating some reference point within the instrument cluster and measuring the radial distances and directions to the accelerometers from that reference point. The point within the accelerometer required for this calculation is sometimes called its "center of percussion." It is effectively the point such that rotations about all axes through the point produce no sensible centrifugal accelerations, and that point can be located by testing the accelerometer at differential reference locations on a rate table.

Angular Acceleration Sensitivities Pendulous accelerometers are sensitive to angular acceleration about their hinge lines, with errors equal to $\dot{\omega}\Delta_{\text{hinge}}$, where $\dot{\omega}$ is the angular acceleration in radian per second squared and Δ_{hinge} is the displacement of the accelerometer proof mass (at its center of mass) from the hinge line. This effect can reach the $1\,g$ level for $\Delta_{\text{hinge}} \approx 1\,\text{cm}$ and $\dot{\omega} \approx 10^3\ \text{rad/s}^2$, but these extreme conditions are usually not persistent enough to matter in most applications.

3.4.3 Calibration Parameter Instabilities

INS calibration parameters are not always exactly constant. Their values can change over the operational life of the INS. Specifications for calibration stability generally divide these calibration parameter variations into two categories: (1) changes from one system turn-on to the next and (2) slow "parameter drift" during operating periods.

3.4.3.1 Calibration Parameter Changes between Turn-Ons These are changes that occur between a system shutdown and the next start-up. They may be caused by temperature transients during shutdowns and turn-ons, or by what is termed "aging." They are generally considered to be independent from turn-on to turn-on, so the model for the covariance of calibration errors for the kth turn-on would be of the form

$$\mathbf{P}_{\text{calib.},k} = \mathbf{P}_{\text{calib.},k-1} + \Delta\mathbf{P}_{\text{calib.}}, \tag{3.9}$$

where $\Delta\mathbf{P}_{\text{calib.}}$ is the covariance of turn-on-to-turn-on parameter changes. The initial value $\mathbf{P}_{\text{calib.},0}$ at the end of calibration is usually determinable from error covariance analysis of the calibration process. Note that this is the covariance model for a random walk, the covariance of which grows without bound.

3.4.3.2 Calibration Parameter Drift This term applies to changes that occur in the operational periods between start-ups and shutdowns. The calibration parameter uncertainty covariance equation has the same form as Eq. 3.9, but with $\Delta\mathbf{P}_{\text{calib.}}$ now representing the calibration parameter drift in the time interval $\Delta t = t_k - t_{k-1}$ between successive discrete times within an operational period.

Detecting Error Trends Incipient sensor failures can sometimes be predicted by observing the variations over time of the sensor calibration parameters. One of the advantages of tightly coupled GNSS/INS integration is that INS sensors can be continuously calibrated all the time that GNSS data are available. System health monitoring can then include tests for the trends of sensor calibration parameters, setting threshold conditions for failing the INS system, and isolating a likely set of causes for the observed trends.

3.4.4 Auxilliary Sensors before GNSS

3.4.4.1 Attitude Sensors Nongyroscopic attitude sensors can also be used as aids in inertial navigation. These include the following:

Magnetic sensors are used primarily for coarse heading initialization to speed up INS alignment.

Star trackers are used primarily for space-based or near-space applications. The Snark cruise missile and the U-2 spy plane used inertial-platform-mounted star trackers to maintain INS alignment on long flights, an idea attributed to Northrop.

Optical alignment systems have been used on some systems prior to launch. Some use Porro prisms mounted on the inertial platform to maintain

TABLE 3.1. Performance Grades for Gyroscopes

Performance Parameter	Performance Units	Performance Grades		
		Inertial	Intermediate	Moderate
Max. input	deg/h	10^2–10^6	10^2–10^6	10^2–10^6
	deg/s	10^{-2}–10^2	10^{-2}–10^2	10^{-2}–10^2
Scale factor	part/part	10^{-6}–10^{-4}	10^{-4}–10^{-3}	10^{-3}–10^{-2}
Bias stability	deg/h	10^{-4}–10^{-2}	10^{-2}–10	10–10^2
	deg/s	10^{-8}–10^{-6}	10^{-6}–10^{-3}	10^{-3}–10^{-2}
Bias drift	$\deg/\sqrt{h}$	10^{-4}–10^{-3}	10^{-2}–10^{-1}	1–10
	$\deg/\sqrt{s}$	10^{-6}–10^{-5}	10^{-5}–10^{-4}	10^{-4}–10^{-3}

optical line-of-sight reference through ground-based theodolites to reference directions at the launch complex.

3.4.4.2 Altitude Sensors These include barometric altimeters and radar altimeters. Without GNSS inputs, some sort of altitude sensor is required to stabilize INS vertical channel errors.

3.4.5 Sensor Performance Ranges

Table 3.1 lists some order-of-magnitude performance ranges for gyroscopes, in terms of the stabilities of their error characteristics. The ranges are labeled

Inertial for those acceptable in stand-alone INS applications.

Intermediate for those acceptable for some applications of integrated GNSS/INS navigators.

Moderate for those considered acceptable only for low-grade integrated GNSS/INS navigators.

These are only rough order-of-magnitude ranges for the different error characteristics. Sensor requirements are largely determined by the application. For example, gyroscopes for gimbaled systems can generally use much smaller input ranges than those for strapdown applications.

The requirements for a specific application are best determined through systems analysis, described in Chapter 11.

3.5 EARTH MODELS

For navigation in the terrestrial environment, both inertial navigation and satellite navigation require models for the shape, gravity, and rotation of the earth.

Both systems use a common set of navigation coordinates fitted to a model of the shape of the earth, and this coordinate system rotates with the earth.

Gravity modeling is important for INS because gravity cannot be sensed. It is necessary for filling in the unsensed acceleration of terrestrial navigation coordinates, and it is necessary for GNSS in determining precise ephemerides of the satellites.

3.5.1 Terrestrial Navigation Coordinates

Descriptions of the major coordinates used in inertial navigation and GNSS/INS integration are described in Appendix B. These include coordinate systems used for representing the trajectories of GNSS satellites and user vehicles in the near-earth environment and for representing the attitudes of host vehicles relative to locally level coordinates, including the following:

1. Inertial coordinates:
   (a) Earth-centered inertial (ECI), with origin at the center of mass of the earth and principal axes in the directions of the vernal equinox and the rotation axis of the earth.
   (b) Satellite orbital coordinates, used in GNSS ephemerides.
2. Earth-fixed coordinates:
   (a) Earth-centered, earth-fixed (ECEF), with origin at the center of mass of the earth and principal axes in the directions of the prime meridian at the equator and the rotation axis of the earth.
   (b) Geodetic coordinates, based on an ellipsoid model for the shape of the earth. Longitude in geodetic coordinates is the same as in ECEF coordinates, and geodetic latitude as defined as the angle between the equatorial plane and the normal to the reference ellipsoid surface. Geodetic latitude can differ from geocentric latitude by as much as 12 arc minutes, equivalent to about 20 km of northing distance.
   (c) Local tangent plane (LTP) coordinates, also called "locally level coordinates," essentially representing the earth as being locally flat. These coordinates are particularly useful from a human factor standpoint for representing the attitude of the host vehicle and for representing local directions. They include
      i. east–north–up (ENU), shown in Fig. B.7;
      ii. north–east–down (NED), which can be simpler to relate to vehicle coordinates; and
      iii. alpha wander, rotated from ENU coordinates through an angle α about the local vertical, as shown in Fig. B.8 and described in Section 3.6.3.1.
3. Vehicle-fixed coordinates:
   (a) roll–pitch–yaw (RPY), as shown in Fig. B.9.

Transformations between these different coordinate systems are important for representing vehicle attitudes, for resolving inertial sensor outputs into inertial navigation coordinates, and for GNSS/INS integration. Methods used for representing and implementing coordinate transformations are also presented in Appendix B, Section B.4.

3.5.2 Earth Rotation

Earth is the mother of all clocks. It has given us the time units of days, hours, minutes, and seconds we use to manage our lives. Not until the discovery of atomic clocks based on hyperfine quantum state transitions were we able to observe the imperfections in our earth clock. Despite these, we continue to use earth rotation as our primary time reference, adding or subtracting leap seconds to atomic clocks to keep them synchronized to the rotation of the earth. These time variations are significant for GNSS navigation but not for inertial navigation.

World Geodetic System 1984 (WGS84) earthrate model: A **geoid** is a model for the equipotential surface of the earth at sea level, referenced to the center of mass of Earth and its rotation axis. "WGS84" refers to the 1984 World Geodetic Survey, a cooperative international effort undertaken to determine a more precise model of the shape of the earth. This effort had begun in the 1950s. Member nations shared data and developed a common geoid model. A series of such efforts resulted in a series of such models, each based on more data. The 1984 values have become standard for most applications.

The value of earthrate in the WGS84 earth model used by the Global Positioning System (GPS) is $7{,}292{,}115{,}167 \times 10^{-14}$ rad/s or about 15.04109 deg/h. This is its **sidereal** rotation rate with respect to distant stars. Its mean rotation rate with respect to the nearest star (our sun), as viewed from the rotating earth, is 15 deg/h, averaged over 1 year. We also know that Earth is slowing down with age due to its transfer of energy and angular momentum to the moon through the effects of tides.[3]

3.5.3 Gravity Models

Gravity may be part of Newtonian mechanics, but it is more appropriately modeled as a warping of the space–time continuum. The important point is **inertial sensors cannot measure gravity**.

Newton could observe the apple falling, but he could not feel gravity pulling the apple (or himself) downward. He could only feel the surface of the earth pushing him upward to counter it.

[3]An effect discovered by George Darwin (1845–1912), son of Charles Darwin.

That leads to one of the problems that had to be solved for inertial navigation in the terrestrial environment: How can we account for gravitational acceleration if we cannot measure it?

The solution is to do what Newton did: Model it.

3.5.3.1 GNSS Gravity Models Accurate gravity modeling is important for maintaining ephemerides for GNSS satellites, and models developed for GNSS have been a boon to inertial navigation as well. However, spatial resolution of the earth gravitational field required for GNSS operation may be a bit coarse compared to that for precision inertial navigation because the GNSS satellites are not near the surface and the mass concentration anomalies that create surface gravity anomalies. GNSS orbits have very little sensitivity to equipotential surface-level undulations of the gravitational field with wavelengths on the order of 100 km or less, but these can be important for high-precision inertial systems.

3.5.3.2 INS Gravity Models Because an INS operates in a world with gravitational accelerations it is unable to sense and unable to ignore, it must use a reasonably faithful model of gravity.

Gravity models for the earth include centrifugal acceleration due to the rotation of the earth as well as true gravitational accelerations due to the mass distribution of the earth, but they do not generally include oscillatory effects such as tidal variations.

Gravitational Potential Gravitational potential is defined to be zero at a point infinitely distant from all massive bodies and to decrease toward massive bodies such as the earth; that is, a point at infinity is the reference point for gravitational potential.

In effect, the gravitational potential at a point in or near the earth is defined by the potential energy lost per unit of mass falling to that point from infinite altitude. In falling from infinity, potential energy is converted to kinetic energy, $mv_{\text{escape}}^2/2$, where v_{escape} is the *escape velocity*. Escape velocity at the surface of the earth is about 11 km/s.

Gravitational Acceleration Gravitational acceleration is the negative gradient of gravitational potential. Potential is a scalar function, and its gradient is a vector. Because gravitational potential increases with altitude, its gradient points upward and the negative gradient points downward.

Equipotential Surfaces An equipotential surface is a surface of constant gravitational potential. If the ocean and atmosphere were not moving, then the surface of the ocean at static equilibrium would be an equipotential surface. **Mean sea level** is a theoretical equipotential surface obtained by time averaging the dynamic effects.

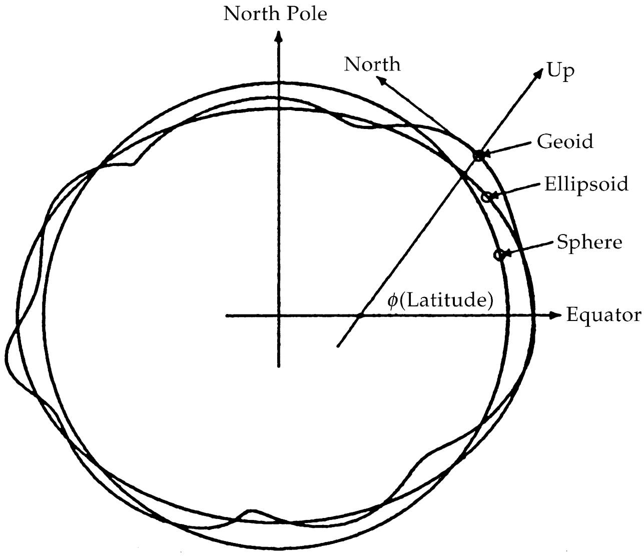

Fig. 3.5 Equipotential surface models for Earth.

Ellipsoid Models for Earth Geodesy is the process of determining the shape of the earth, often using ellipsoids as approximations of an equipotential surface (e.g., mean sea level), as illustrated in Fig. 3.5. The most common ones are ellipsoids of revolution, but there are many reference ellipsoids based on different survey data. Some are global approximations and some are local approximations. The global approximations deviate from a spherical surface by about ±10 km, and locations on the earth referenced to different ellipsoidal approximations can differ from one another by 10^2–10^3 m.

Geodetic latitude on a reference ellipsoid is measured in terms of the angle between the equator and the normal to the ellipsoid surface, as illustrated in Fig. 3.5.

Orthometric height is measured along the (curved) plumb line.

WGS84 Ellipsoid The WGS84 earth model approximates mean sea level (an equipotential surface) by an ellipsoid of revolution with its rotation axis coincident with the rotation axis of the earth, its center at the center of mass of the earth, and its prime meridian through Greenwich. Its semimajor axis (equatorial radius) is defined to be 6,378,137 m, and its semiminor axis (polar radius) is defined to be 6,356,752.3142 m.

Geoid Models Geoids are approximations of mean sea-level orthometric height with respect to a reference ellipsoid. Geoids are defined by additional higher-order shapes, commonly modeled by spherical harmonics of height deviations from an ellipsoid, as illustrated in Fig. 3.5. There are many geoid

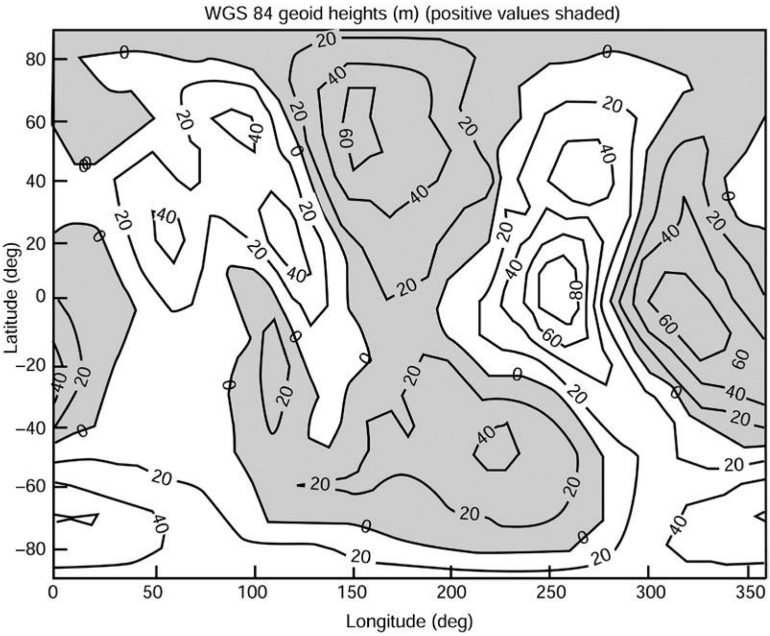

Fig. 3.6 WGS84 geoid heights.

models based on different data, but the more recent, most accurate models depend heavily on GPS data. Geoid heights deviate from reference ellipsoids by tens of meters, typically.

The WGS84 geoid heights vary about ±100 m from the reference ellipsoid. As a rule, oceans tend to have lower geoid heights and continents tend to have higher geoid heights. Coarse 20-m contour intervals are plotted versus longitude and latitude in Fig. 3.6, with geoid regions above the ellipsoid shaded gray.

3.5.3.3 Longitude and Latitude Rates The second integral of acceleration in locally level coordinates should result in the estimated vehicle position. This integral is somewhat less than straightforward when longitude and latitude are the preferred horizontal location variables.

The rate of change of vehicle altitude equals its vertical velocity, which is the first integral of net (i.e., including gravity) vertical acceleration. The rates of change of vehicle longitude and latitude depend on the horizontal components of vehicle velocity, but in a less direct manner. The relationship between longitude and latitude rates and east and north velocities is further complicated by the oblate shape of the earth.

The rates at which these angular coordinates change as the vehicle moves tangent to the surface will depend upon the radius of curvature of the reference surface model, which is an ellipsoid of revolution for the WGS84 model.

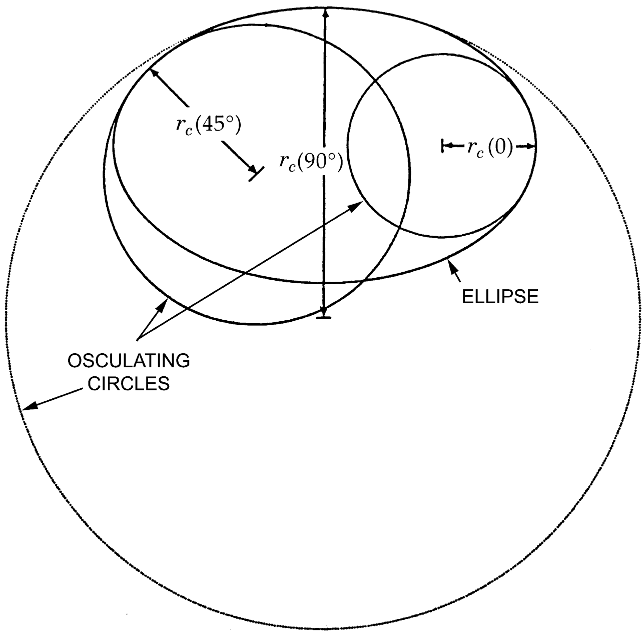

Fig. 3.7 Ellipse and osculating circles.

Radius of curvature can depend on the direction of travel, and for an ellipsoidal model, there is one radius of curvature for north–south motion and another radius of curvature for east–west motion.

Meridional Radius of Curvature The radius of curvature for north–south motion is called the "meridional" radius of curvature, because north–south travel is along a meridian (i.e., line of constant longitude). For an ellipsoid of revolution (the WGS84 model), all meridians have the same shape, which is that of the ellipse that was rotated to produce the ellipsoidal surface model. The tangent circle with the same radius of curvature as the ellipse is called the *osculating circle* (*osculating* means "kissing"). As illustrated in Fig. 3.7 for an oblate earth model, the radius of the meridional osculating circle is smallest where the geocentric radius is largest (at the equator), and the radius of the osculating circle is largest where the geocentric radius is smallest (at the poles). The osculating circle lies inside or on the ellipsoid at the equator and outside or on the ellipsoid at the poles and passes through the ellipsoid surface for latitudes in between.

The formula for meridional radius of curvature as a function of geodetic latitude (ϕ_{geodetic}) is

$$r_M = \frac{b^2}{a[1 - e^2 \sin^2(\phi_{\text{geodetic}})]^{3/2}} \tag{3.10}$$

$$= \frac{a(1-e^2)}{\left[1-e^2\sin^2(\phi_{\text{geodetic}})\right]^{3/2}}, \tag{3.11}$$

where a is the semimajor axis of the ellipse, b is the semiminor axis, and $e^2 = (a^2 - b^2)/a^2$ is the eccentricity squared.

Geodetic Latitude Rate The rate of change of geodetic latitude as a function of north velocity v_N is then

$$\frac{d\phi_{\text{geodetic}}}{dt} = \frac{v_N}{r_M + h}, \tag{3.12}$$

and geodetic latitude can be maintained as the integral

$$\phi_{\text{geodetic}}(t_{\text{now}}) = \phi_{\text{geodetic}}(t_{\text{start}}) + \int_{t_{\text{start}}}^{t_{\text{now}}} \frac{v_{N(t)dt}}{a(1-e^2)/\{1-e^2\sin^2[\phi_{\text{geodetic}}(t)]\}^{3/2} + h(t)]}, \tag{3.13}$$

where $h(t)$ is height above (+) or below (–) the ellipsoid surface and $\phi_{\text{geodetic}}(t)$ will be in radians if $v_N(t)$ is in meters per second and $r_M(t)$ and $h(t)$ are in meters.

Transverse Radius of Curvature The radius of curvature of the reference ellipsoid surface in the east–west direction (i.e., orthogonal to the direction in which the meridional radius of curvature is measured) is called the *transverse radius of curvature*. It is the radius of the osculating circle in the local east–up plane, as illustrated in Fig. 3.8, where the arrows at the point of tangency of the

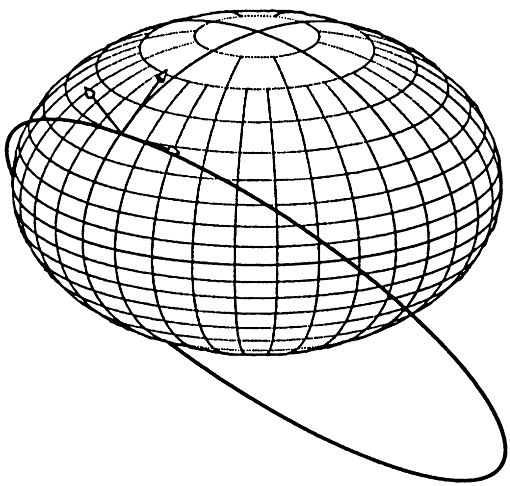

Fig. 3.8 Transverse osculating circle.

transverse osculating circle are in the local ENU coordinate directions. As this figure illustrates, on an oblate earth, the plane of a transverse osculating circle does not pass through the center of the earth except when the point of osculation is at the equator. (All osculating circles at the poles are in meridional planes.) Also, unlike meridional osculating circles, transverse osculating circles generally lie outside the ellipsoidal surface, except at the point of tangency and at the equator, where the transverse osculating circle *is* the equator.

The formula for the transverse radius of curvature on an ellipsoid of revolution is

$$r_T = \frac{a}{\sqrt{1-e^2\sin^2(\phi_{\text{geodetic}})}}, \tag{3.14}$$

where a is the semimajor axis of the generating ellipse and e is its eccentricity.

LONGITUDE RATE The rate of change of longitude as a function of east velocity is then

$$\frac{d\theta}{dt} = \frac{v_E}{\cos(\phi_{\text{geodetic}})(r_T + h)}, \tag{3.15}$$

and longitude can be maintained by the integral

$$\theta(t_{\text{now}}) = \theta(t_{\text{start}}) + \int_{t_{\text{start}}}^{t_{\text{now}}} \frac{v_E(t)\,dt}{\cos[\phi_{\text{geodetic}}(t)]\left(a/\sqrt{1-e^2\sin^2(\phi_{\text{geodetic}}(t))} + h(t)\right)}, \tag{3.16}$$

where $h(t)$ is height above (+) or below (−) the ellipsoid surface and θ will be in radians if $v_E(t)$ is in meters per second and $r_T(t)$ and $h(t)$ are in meters. Note that this formula has a singularity at the poles, where $\cos(\phi_{\text{geodetic}}) = 0$, a consequence of using latitude and longitude as location variables.

WGS84 Reference Surface Curvatures The apparent variations in meridional radius of curvature in Fig. 3.7 are rather large because the ellipse used in generating Fig. 3.7 has an eccentricity of about 0.75. The WGS84 ellipse has an eccentricity of about 0.08, with geocentric, meridional, and transverse radius of curvature as plotted in Fig. 3.9 versus geodetic latitude. For the WGS84 model,

- mean geocentric radius is about 6371 km, from which it varies by −14.3 km (−0.22%) to +7.1 km (+0.11%);
- mean meridional radius of curvature is about 6357 km, from which it varies by −21.3 km (−0.33%) to 42.8 km (+0.67%); and
- mean transverse radius of curvature is about 6385 km, from which it varies by −7.1 km (−0.11%) to +14.3 km (+0.22%).

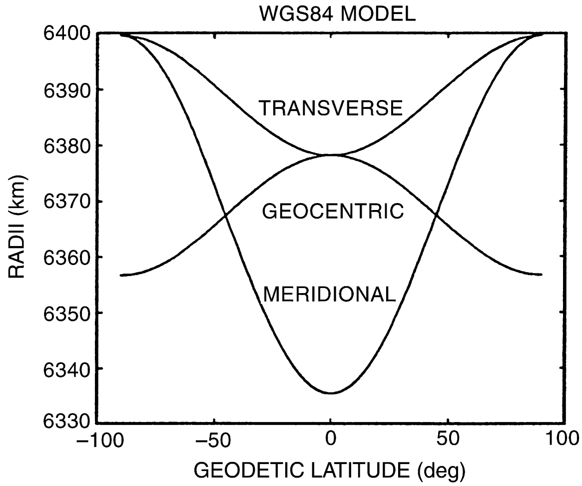

Fig. 3.9 Radii of WGS84 reference ellipsoid.

Because these vary by several parts per thousand, one must take the radius of curvature into account when integrating horizontal velocity increments to obtain the longitude and the latitude.

3.6 HARDWARE IMPLEMENTATIONS

INSs generally fall into two categories depending on the hardware configuration:

1. **Gimbaled** or **floated** systems, in which the ISA is isolated from rotations of the host vehicle, as illustrated in Fig. 3.10(*a*–*c*). This shows three alternative structures that have been tried at different times:
   (a) Gimbals, also called a Cardan[4] suspension. This is the most popular implementation using hardware to solve the attitude problem.
   (b) Ball joint, which Fritz Mueller called "inverted gimbals" [7]. It has not become popular, perhaps because of the difficulties of applying controlled torques about the spherical bearing to stabilize the ISA. This configuration is not discussed here.
   (c) A floated sphere, a configuration also called "FLIMBAL," an acronym for floated inertial measurement ball.[5] Despite the difficul-

[4]Named after the Italian physician, inventor, and polymath Girolamo Cardano (1501–1576), who also invented what Americans call a "universal joint" and Europeans call a "Cardan shaft."
[5]A name used at the MIT Instrumentation Laboratory around 1957, when work on its development started.

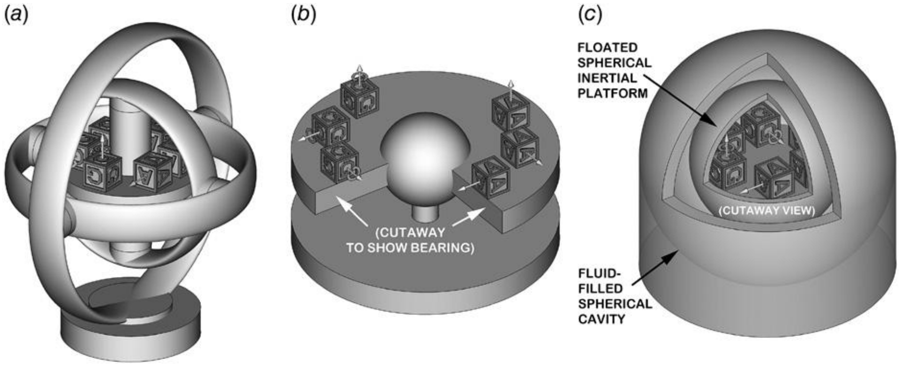

Fig. 3.10 Gimbaled IMU alternatives. (*a*) Gimbal (Cardan suspension). (*b*) Ball joint ("inverted gimbals"). (*c*) Floated sphere ("FLIMBAL").

ties of transferring power, signals, heat, torque, and relative attitude between the housing and the inner spherical ISA, it is probably the most accurate (and expensive) implementation for high-g rocket booster applications.

In all cases, the rotation-isolated ISA is also called an inertial platform, stable platform, or stable element. The IMU includes the ISA, the gimbal/float structure, and all associated electronics (e.g., gimbal wiring, rotary slip rings, gimbal bearing angle encoders, signal conditioning, gimbal bearing torque motors, and thermal control).

2. **Strapdown** systems, essentially as illustrated in Fig. 3.1. In this case, the ISA is not isolated from rotations but is "quasi-rigidly" mounted to the frame structure of the host vehicle.

We use the term "quasi-rigid" for IMU mountings that can provide some isolation of the IMU from shock and vibration transmitted through the host vehicle frame. These vibrational accelerations do not significantly alter the navigation solution, but they can damage the ISA and its sensors. Strapdown, gimbaled, and floated systems may require **shock and vibration isolators** to dampen the vibrational torques and forces transmitted to the inertial sensors. These isolators are commonly made from "lossy" elastomers that provide some amount of damping, as well.

3.6.1 Gimbaled Implementations

The use of gimbals for isolation from rotation has been documented as far back as the third century BCE. The term generally applies to the entire structure, usually consisting of a set of two or three (or four) nested rings with orthogonal pivots (also called "gimbal bearings"). As illustrated in Fig. 3.11(*a*), three sets of gimbal bearings are sufficient for complete rotational isolation in applications with limited attitude mobility (e.g., surface ships), but applications in fully maneuverable host vehicles require an additional gimbal bearing to avoid the condition shown in Fig. 3.11(*b*), known as “gimbal lock,” in which the gimbal configuration no longer provides isolation from outside rotations about all three axes.

For inertial navigation, gyroscopes inside the gimbals detect any incipient rotation of that frame due to torques from any source (e.g., bearing friction or mass imbalance) and apply feedback to torquing motors in the gimbal bearings to keep the rotation rates inside the gimbals at zero. For navigation with respect to the rotating earth, the gimbals can also be servoed to maintain the sensor axes fixed in locally level coordinates.

Design of gimbal torquing servos is complicated by the motions of the gimbals during operation, which changes how the torquing to correct for sensed rotation must be applied to the different gimbal bearings. This requires a bearing angle sensor for each gimbal axis.

The gimbal arrangement shown in Fig. 3.11(*a*), with the outer gimbal axis aligned to the roll (longitudinal) axis of the host vehicle and the inner gimbal axis maintained in the vertical direction, is a popular one. If the ISA is kept aligned with locally level ENU directions, the gimbal bearing angles will equal the heading (yaw), pitch, and roll Euler angles defining the host vehicle attitude relative to north, east, and down directions. These are the same Euler angles used to drive **attitude and heading reference systems** (AHRSs) (e.g., compass card and artificial horizon displays) in aircraft cockpits.

Advantages The principal advantage of both gimbaled and floated systems is the isolation of the inertial sensors from high angular rates, which eliminates many rate-dependent sensor errors (including gyro scale factor sensitivity) and generally allows for higher accuracy sensors. Also, gimbaled systems can be self-calibrated by orienting the ISA with respect to gravity (for calibrating the accelerometers) and with respect to the earth rotation axis (for

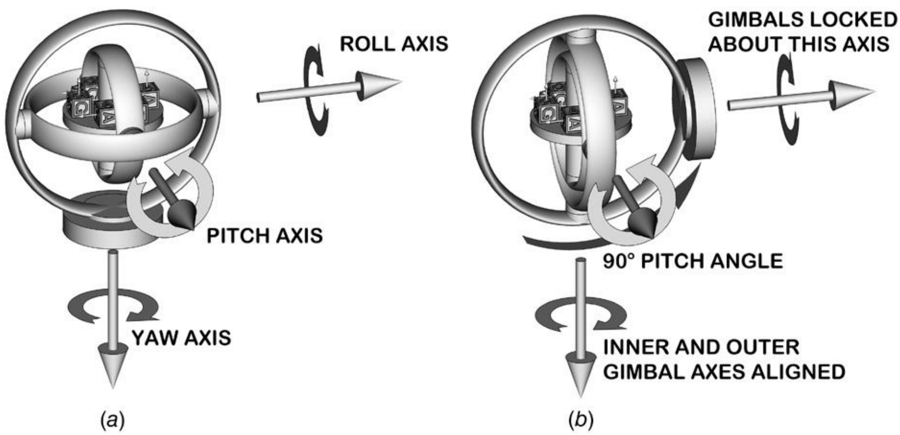

Fig. 3.11 Three-axis gimbal lock. (*a*) Straight and level. (*b*) Pitched up 90°.

calibrating the gyros), and by using external optical autocollimators with mirrors on the ISA to independently measure its orientation with respect to its environment.

The most demanding INS applications for "cruise" applications (i.e., at ≈1 g) are probably for nuclear missile-carrying submarines, which must navigate submerged for months. The gimbaled Electrically Supported Gyro Navigation (ESGN, DoD designation AN/WSN-3 [2]) system developed in the 1970s for USN Trident-class submarines was probably the most accurate INS of that era [6].

Disadvantages The principal drawbacks of gimbals are cost, weight, volume, and gimbal flexure in high-*g* environments. In traditional designs, electrical pathways are required through the gimbal structure to provide power to the IMU and to carry power and the encoder, torquer, and sensor signals. These require slip rings (which can introduce noise) or cable wraps at the gimbal bearings. More recent designs, however, have used wireless signal transmission. The gimbals can alter air circulation used to maintain uniform temperatures within the IMU, and they can hamper access to the sensors during test and operation (e.g., optical measurements using mirrors on the ISA to check its attitude).

3.6.2 Floated Implementation

Gimbals and gimbal lock can be eliminated by floating the ISA in a liquid and operating it like a robotic submersible, using liquid thrusters to maintain its orientation and to keep itself centered within the flotation cavity—as illustrated on the right in Fig. 3.10. The floated assembly must also be neutrally buoyant and balanced to eliminate acceleration-dependent disturbances.

Advantages Floated systems have the advantage over gimbaled systems that there are no gimbal structures to flex or vibrate under dynamic loading, and no gimbals interfering with heat transfer from the ISA. Floated systems have the same advantages as gimbaled systems over strapdown systems: isolation of the inertial sensors from high angular rates, which eliminates many rate-dependent error effects and generally allows for higher accuracy sensors. They also have the ability to orient the sphere for self-calibration, although it is more difficult to verify the orientation of the ISA from the outside. The floated Advanced Inertial Reference Sphere (AIRS) designed at the Charles Stark Draper Laboratory for MX/Peacekeeper and Minuteman III missiles is probably the most accurate (and most expensive) high-g INS ever developed [6].

Disadvantages A major disadvantage of floated systems is the difficulty of accessing the ISA for diagnostic testing, maintenance, or repair. The flotation system must be disassembled and the fluid drained for access, and then reassembled for operation. Floated systems also require some means for determining the attitude of the floated assembly relative to the host vehicle and providing power to the floated assembly, and passing commands and sensor signals through the fluid, and for maintaining precisely controlled temperatures within the floated IMU.

Also, for applications in which the host vehicle attitude must be controlled, gimbaled or floated systems provide only vehicle attitude information, whereas strapdown systems provide attitude rates for vehicle attitude control loops.

3.6.3 Carouseling and Indexing

3.6.3.1 Alpha Wander and Carouseling *Alpha Wander* Near the poles, there is a problem with locally level gimbal orientations in which the level axes are constrained to point north and east. At the poles, there is no north or east direction; all directions are either south (at the North Pole) or north (at the South Pole). Also, near the poles the slewing rates required to keep the north/east orientations can be unreasonably high. The solution is to allow the locally level axes to wander, with the angle α (alpha) designating the angle between one of the level axes and north (except near the poles, where it is referenced to earth-fixed polar stereographic coordinates). This is called an "alpha wander" implementation.

Carouseling A *carousel* is an amusement ride using continuous rotation of a circular platform about a vertical axis. The term "carouseling" has been applied to an implementation for gimbaled or floated systems in which the ISA revolves slowly around the local vertical axis—at rates in the order of a revolution per minute. The three-gimbal configuration shown on the left in Fig. 3.11(*a*) can implement carouseling using only the inner (vertical) gimbal axis. Carouseling significantly reduces long-term navigation errors due to certain types of sensor errors (uncompensated biases of nominally level accelerometers and gyroscopes, in particular). This effect was discovered and exploited at the AC[6] Spark Plug (later Delco Electronics) Division of General Motors in the early 1960s. Delco Carousel systems became phenomenally successful and popular for years.

3.6.3.2 Indexing Alternative implementations called **indexing** or **gimbal flipping** use discrete rotations (usually by multiples of 90°) to the same effect.

Indexing can also be used at the sensor level. The electrostatic gyroscopes in the U.S. Navy's ESGN (DoD designation AN/WSN-3 [2]) use independent gimbals to keep the spin axis of each rotor in the same direction relative to its suspension cavity but index those gimbals[7] to provide periodic cavity rotations. The resulting system accuracy is classified, but good enough that nothing was able to surpass it for decades.

[6]The "AC" in the name is the initials of Albert Champion (1878–1927), a former world-class cyclist who founded Champion Spark Plug Company in Boston in the early 1900s. When he lost control of that company in 1908, Albert founded the Champion Ignition Company in Flint, Michigan. It was renamed the AC Spark Plug Company in 1909, when it was purchased by General Motors.

[7]Using a pattern designed by Kenneth P. Gow (1917–2001), and called the "Gow flip."

3.6.4 Strapdown Systems

Strapdown systems use an IMU that is not isolated from rotations of its host vehicle—except possibly by shock and vibration isolators. The gimbals are effectively replaced by software that uses the gyroscope outputs to calculate the equivalent accelerometer outputs in an attitude-stabilized coordinate frame, and integrates them to provide updates of velocity and position. This requires more computation (which is cheap) than the gimbaled implementation, but it eliminates the gimbal system (which may not be cheap). It also exposes the accelerometers and gyroscopes to relatively high rotation rates, which can cause attitude rate-dependent sensor errors.

Advantages The principal advantage of strapdown systems over gimbaled or floated systems is cost. The cost or replicating software is vanishingly small compared to the cost of replicating a gimbal system for each IMU. For applications requiring attitude control of the host vehicle, strapdown gyroscopes generally provide more accurate rotation rate data than the attitude readouts of gimbaled or floated systems.

Disadvantages Strapdown sensors must operate at much higher rotation rates, which limits the possible design choices. Pendulous integrating gyroscopic accelerometers, for example, are very sensitive to rotation. The dynamic ranges of the inputs to strapdown gyroscopes may be orders of magnitude greater than those for gyroscopes in gimbaled systems. To achieve comparable navigation performance, this generally requires orders of magnitude better scale factor stability for the strapdown gyroscopes. Strapdown systems generally require much shorter integration intervals—especially for integrating gyroscope outputs, which increases computer costs relative to gimbaled systems. Another disadvantage for strapdown is the cost of gyroscope calibration and testing, which requires a precision rate table. Precision rate table testing is not required for whole-angle gyroscopes—including electrostatic gyroscopes—or for any gyroscopes used in gimbaled systems.

3.6.5 Strapdown Carouseling and Indexing

For host vehicles that are nominally upright during operation (e.g., ships), a strapdown system can be rotated about the host vehicle yaw axis. So long as the vehicle yaw axis remains close to the local vehicle, slow rotation (carouseling) or indexing about this axis can significantly reduce the effects of uncompensated biases of the nominally level accelerometers and gyroscopes. The rotation is normally oscillatory, with reversal of direction after a full rotation, so that the connectors can be wrapped to avoid using slip rings (an option not generally available for gimbaled systems).

Carouseling or indexing of strapdown systems requires the addition of a rotation bearing and associated motor drive, wiring and control electronics. However, the improvement in navigational performance may justify the additional cost. The U.S. Air Force N–73 inertial navigator was an early strapdown system using carouseling, and its navigation accuracy was better than one nautical mile per hour CEP rate.

3.7 SOFTWARE IMPLEMENTATIONS

3.7.1 Example in One Dimension

Inertial navigation would be much simpler if we all lived in a one-dimensional "line land." For one thing, there would be no rotation and no need for gyroscopes or gimbals. In that case, an INS would need only one accelerometer and navigation computer (all one-dimensional line segments, of course), and its implementation would be about as illustrated in Fig. 3.12 (in two dimensions), where the dependent variable x denotes position on the line in one dimension and the independent variable t is time.

This implementation for one dimension still has some features in common with implementations for three dimensions:

1. Accelerometers cannot measure gravitational acceleration.
2. Accelerometers have *scale factors*, which are the ratios of input acceleration units to output signal magnitude units (e.g., meters per second squared per volt). The signal must be rescaled in the navigation computer by multiplying by this scale factor.
3. Accelerometers have *output errors*, including
   (a) unknown constant *offsets*, also called *biases*;
   (b) unknown constant *scale factor errors*;
   (c) unknown *nonconstant variations* in bias and scale factor; and
   (d) unknown zero-mean additive *noise* on the sensor outputs, including quantization noise and electronic noise. The noise itself is not predictable, but its statistical properties may be used in Kalman filtering to estimate drifting scale factor and biases.

   In one dimension, there is no such thing as input axis misalignment.
4. Gravitational accelerations must be modeled and calculated in the navigational computer, then added to the sensed acceleration (after

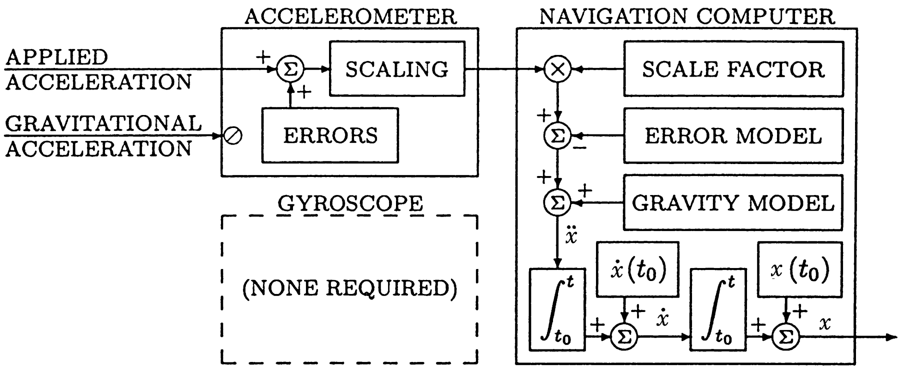

Fig. 3.12 INS functional implementation for a one-dimensional world.

error and scale compensation) to obtain the net acceleration $\ddot{x}$ of the INS.

5. The navigation computer must integrate acceleration to obtain velocity. This is a definite integral and it requires an initial value, $\dot{x}(t_0)$; that is, the INS implementation in the navigation computer must start with a known initial velocity.
6. The navigation computer must also integrate velocity $(\dot{x})$ to obtain position (x). This is also a definite integral and it also requires an initial value, $x(t_0)$. The INS implementation in the navigation computer must start with a known initial location too.

Inertial navigation in three dimensions requires more sensors and more signal processing than in one dimension, and it also introduces more possibilities for implementation (e.g., gimbaled or strapdown).

3.7.2 Initialization in Nine Dimensions

In the three-dimensional world of inertial navigation, the navigation solution requires initial values for

1. position (three-dimensional),
2. velocity (also three-dimensional), and
3. attitude (also three-dimensional).

3.7.2.1 Navigation Initialization INS initialization is the process of determining initial values for system position, velocity, and attitude in navigation coordinates. INS position initialization ordinarily relies on external sources such as GNSS, local wireless service, or manual entry. INS velocity initialization can be accomplished by starting when it is zero (i.e., the host vehicle is not moving) or (for vehicles carried in or on other vehicles) by reference to the carrier velocity. (See alignment method 3 below.) INS attitude initialization is called **alignment**.

3.7.2.2 INS Alignment Methods INS alignment is the process of determining the ISA orientation relative to navigation coordinates.

There are four basic alignment methods:

1. **Optical alignment**, using either of the following:
   (a) Optical line-of-sight reference to a ground-based direction (e.g., using a ground-based theodolite and a mirror on the platform). Some space boosters have used this type of optical alignment, which is much faster and more accurate than gyrocompass alignment. Because it requires a stable platform for mounting the mirror, it is only applicable to gimbaled systems.

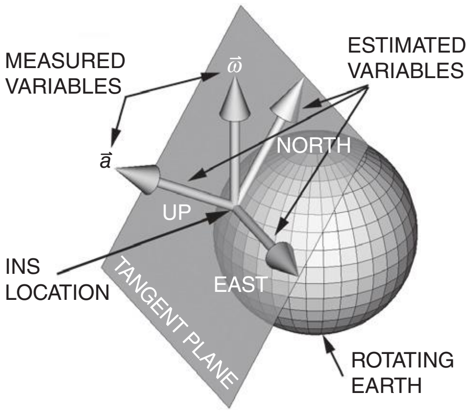

Fig. 3.13 Gyrocompassing determines sensor orientations with respect to east, north, and up.

(b) An onboard star tracker, used primarily for alignment of gimbaled or strapdown systems in space or near space (e.g., above the clouds).

2. **Gyrocompass alignment** of stationary vehicles, using the sensed direction of acceleration to determine the local vertical and the sensed direction of rotation to determine north, as illustrated in Fig. 3.13. Latitude can be determined by the angle between the earth rotation vector and the horizontal, but longitude must be determined by other means and entered manually or electronically. This method is inexpensive but the most time-consuming (several minutes, typically).
3. **Transfer alignment** in a moving host vehicle, using velocity matching with an aligned and operating INS. This method is generally faster than gyrocompass alignment, but it requires another INS on the host vehicle and it may require special maneuvering of the host vehicle to attain observability of the alignment variables. It is commonly used for in-air INS alignment for missiles launched from aircraft and for on-deck INS alignment for aircraft launched from carriers. Alignment of carrier-launched aircraft may also use the direction of the velocity impulse imparted by the steam catapult.
4. **GNSS-aided alignment**, using position matching with GNSS to estimate the alignment variables. It is an integral part of integrated GNSS/INS implementations. It does not require the host vehicle to remain stationary during alignment, but there will be some period of time after turn-on (a few minutes, typically) before system navigation errors settle to acceptable levels.

3.7.2.3 Gyrocompass Alignment Gyrocompass alignment is the only one of those listed above requiring no external aiding. Gyrocompass alignment is not necessary for integrated GNSS/INS, although many INSs may already be configured for it.

Accuracy A rough rule of thumb for gyrocompass alignment accuracy is

$$\sigma^2_{\text{gyrocompass}} > \sigma^2_{\text{acc}} + \frac{\sigma^2_{\text{gyro}}}{15^2 \cos^2(\phi_{\text{geodetic}})}, \tag{3.17}$$

where

$\sigma_{\text{gyrocompass}}$ is the minimum achievable RMS alignment error in radians,
σ_{acc} is the RMS accelerometer accuracy in g's,
σ_{gyro} is the RMS gyroscope accuracy in degrees per hour,
15 deg/h is the rotation rate of the earth, and
ϕ_{geodetic} is the latitude at which gyrocompassing is performed.

Alignment accuracy is also a function of the time allotted for it, and the time required to achieve a specified accuracy is generally a function of sensor error magnitudes (including noise) and the degree to which the vehicle remains stationary.

Gimbaled Implementation Gyrocompass alignment for gimbaled systems is a process for aligning the inertial platform axes with the navigation coordinates using only the sensor outputs, while the host vehicle is essentially stationary. For systems using ENU navigation coordinates, for example, the platform can be tilted until two of its accelerometer inputs are zero, at which time both input axes will be horizontal. In this locally leveled orientation, the sensed rotation axis will be in the north–up plane, and the platform can be slewed about the vertical axis to null the input of one of its horizontal gyroscopes, at which time that gyroscope input axis will point east–west. That is the basic concept used for gyrocompass alignment, but practical implementation requires filtering[8] to reduce the effects of sensor noise and unpredictable zero-mean vehicle disturbances due to loading activities and/or wind gusts.

Strapdown Implementation Gyrocompass alignment for strapdown systems is a process for "virtual alignment" by determining the sensor cluster attitude with respect to navigation coordinates using only the sensor outputs while the system is essentially stationary.

Error-Free Implementation If the sensor cluster could be firmly affixed to the earth and there were no sensor errors, then the sensed acceleration vector

[8]The vehicle dynamic model used for gyrocompass alignment filtering can be "tuned" to include the major resonance modes of the vehicle suspension.

$\mathbf{a}_{\text{output}}$ in sensor coordinates would be in the direction of the local vertical, the sensed rotation vector $\boldsymbol{\omega}_{\text{output}}$ would be in the direction of the earth rotation axis, and the unit column vectors

$$1_U = \frac{\mathbf{a}_{\text{output}}}{|\mathbf{a}_{\text{output}}|}, \tag{3.18}$$

$$1_N = \frac{\boldsymbol{w}_{\text{output}} - (1_U^T \boldsymbol{w}_{\text{output}})1_U}{|\boldsymbol{w}_{\text{output}} - (1_U^T \boldsymbol{w}_{\text{output}})1_U|}, \tag{3.19}$$

$$1_E = 1_N \otimes 1_U \tag{3.20}$$

would define the initial value of the coordinate transformation matrix from sensor-fixed coordinates to ENU coordinates:

$$\mathbf{C}_{\text{ENU}}^{\text{sensor}} = [1_E | 1_N | 1_U]^T. \tag{3.21}$$

Practical Implementation In practice, the sensor cluster is usually mounted in a vehicle that is not moving over the surface of the earth but may be buffeted by wind gusts or disturbed during fueling and loading operations. Gyrocompassing then requires some amount of filtering (Kalman filtering, as a rule) to reduce the effects of vehicle buffeting and sensor noise. The gyrocompass filtering period is typically on the order of several minutes for a medium-accuracy INS but may continue for hours, days, or continuously for high-accuracy systems.

3.7.3 Gimbal Attitude Implementations

The primary function of gimbals is to isolate the ISA from vehicle rotations, but they are also used for other INS functions.

3.7.3.1 Accelerometer Recalibration Navigation accuracy is very sensitive to accelerometer biases, which can shift due to thermal transients in turn-on/turn-off cycles, and can also drift randomly over time. Fortunately, the gimbals can be used to calibrate accelerometer biases in a stationary 1-g environment. In fact, both bias and scale factor can be determined by using the gimbals to point the accelerometer input axis straight up and straight down, and recording the respective accelerometer outputs a_{up} and a_{down}. Then the bias $a_{\text{bias}} = (a_{\text{up}} + a_{\text{down}})/2$ and scale factor $s = (a_{\text{up}} - a_{\text{down}})/2g_{\text{local}}$, where g_{local} is the local gravitational acceleration.

3.7.3.2 Vehicle Attitude Determination The gimbal angles determine the vehicle attitude with respect to the ISA, which has a controlled orientation with respect to navigation coordinates. Each gimbal angle encoder output determines the relative rotation of the structure outside the gimbal axis relative to the structure inside the gimbal axis; the effect of each rotation can be

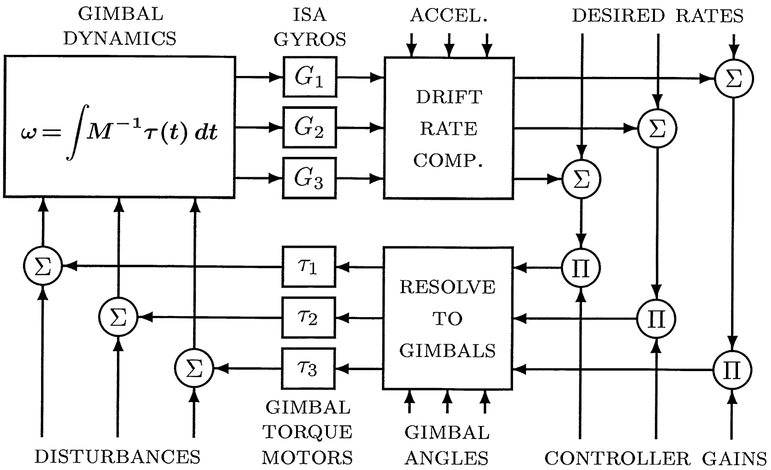

Fig. 3.14 Simplified control flow diagram for three gimbals.

represented by a 3×3 rotation matrix, and the coordinate transformation matrix representing the attitude of the vehicle with respect to the ISA will be the ordered product of these matrices.

For example, in the gimbal structure shown in Fig. 3.11(a), each gimbal angle represents an Euler angle for vehicle rotations about the vehicle roll, pitch, and yaw axes.

3.7.3.3 ISA Attitude Control Gimbals control ISA orientation. This is a 3-degree-of-freedom problem, and the solution is unique for three gimbals; that is, there are three attitude control loops with (at least) three sensors (the gyroscopes) and three torquers. Each control loop can use a proportional, integral, and differential (PID) controller, with the commanded torque distributed to the three torquers according to the direction of the torquer/gimbal axis with respect to the gyro input axis, somewhat as illustrated in Fig. 3.14, where

DISTURBANCES includes the sum of all torque disturbances on the individual gimbals and the ISA, including those due to ISA mass unbalance and acceleration, rotations of the host vehicle, air currents, and torque motor errors.

GIMBAL DYNAMICS is actually quite a bit more complicated than the rigid-body torque equation

$$\boldsymbol{\tau} = \mathbf{M}_{\text{inertia}}\dot{\boldsymbol{\omega}},$$

which is the torque analog of $\mathbf{F} = m\mathbf{a}$, where $\mathbf{M}_{\text{inertia}}$ is the moment of inertia matrix. The IMU is not a rigid body, and the gimbal torque motors apply torques *between* the gimbal elements (i.e., ISA, gimbal rings, and host vehicle).

DESIRED RATES refer to the rates required to keep the ISA aligned to a moving coordinate frame (e.g., locally level).

RESOLVE TO GIMBALS is where the required torques are apportioned among the individual torquer motors on the gimbal axes. The actual control loop is more complicated than that shown in the figure, but it does illustrate in general terms how the sensors and actuators are used.

For systems using four gimbals to avoid gimbal lock, the added gimbal adds another degree of freedom to be controlled. In this case, the control law usually adds a fourth constraint (e.g., maximize the minimum angle between gimbal axes) to avoid gimbal lock.

3.7.4 Gimbaled Navigation Implementation

The signal flowchart in Fig. 3.15 shows the essential navigation signal processing functions for a gimbaled INS with inertial sensor axes aligned to locally level coordinates, where

f_{specific} is the **specific force** (i.e., the sensible acceleration, exclusive of gravitational acceleration) applied to the host vehicle.

Ω_{inertial} is the instantaneous inertial rotation rate vector of the host vehicle.

A denotes a specific force sensor (accelerometer).

θ_j denotes the ensemble of gimbal angle encoders, one for each gimbal angle. There are several possible formats for the gimbal angles, including digitized angles, three-wire synchros signals, or `sin/cos` pairs.

G denotes an inertial rotation rate sensor (gyroscope).

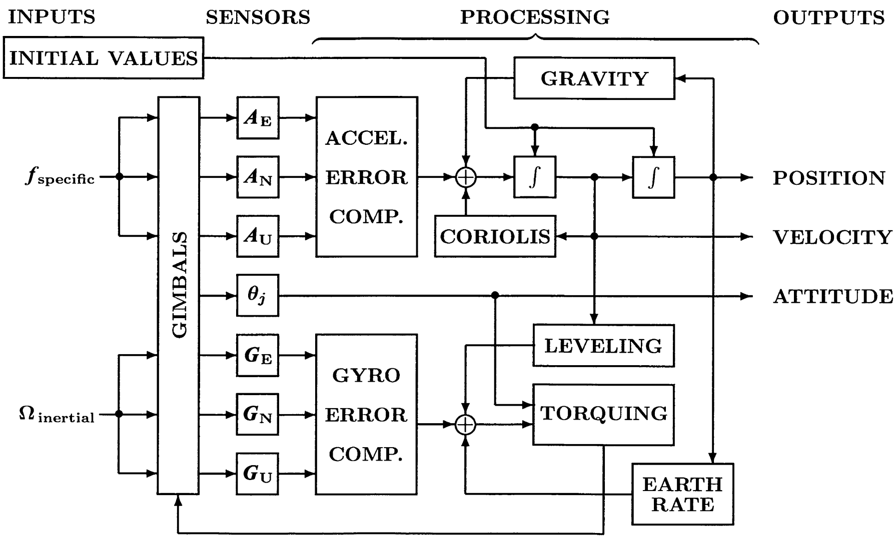

Fig. 3.15 Essential navigation signal processing for gimbaled INS.

POSITION is the estimated position of the host vehicle in navigation coordinates (e.g., longitude, latitude, and altitude relative to sea level).

VELOCITY is the estimated velocity of the host vehicle in navigation coordinates (e.g., east, north, and vertical).

ATTITUDE is the estimated attitude of the host vehicle relative to locally level coordinates. For some three-gimbal systems, the gimbal angles are the Euler angles representing vehicle heading (with respect to north), pitch, and roll. Output attitude may also be used to drive cockpit displays such as compass cards or artificial horizon indicators.

ACCELEROMETER ERROR COMPENSATION and GYROSCOPE ERROR COMPENSATION denote the calibrated corrections for sensor errors. These generally include corrections for scale factor variations, output biases and input axis misalignments for both types of sensors, and acceleration-dependent errors for gyroscopes.

GRAVITY denotes the gravity model used to compute the acceleration due to gravity as a function of position.

CORIOLIS denotes the acceleration correction for coriolis effect in rotating coordinates.

LEVELING denotes the rotation rate correction to maintain locally level coordinates while moving over the surface of the earth.

EARTHRATE denotes the model used to calculate the earth rotation rate in locally level INS coordinates.

TORQUING denotes the servo loop gain computations used in stabilizing the INS in locally level coordinates.

Not shown in the figure is the input altitude reference (e.g., barometic altimeter or GPS) required for vertical channel (altitude) stabilization.

3.7.5 Strapdown Attitude Implementations

3.7.5.1 Strapdown Attitude Problems Early on, strapdown systems technology had an "attitude problem," which was the problem of representing attitude rate in a format amenable to accurate computer integration. The eventual solution was to represent attitude in different mathematical formats as it is processed from raw gyro outputs to the matrices used for transforming sensed acceleration to inertial coordinates for integration.

Figure 3.16 illustrates the resulting major gyro signal processing operations and the formats of the data used for representing attitude information. The processing starts with gyro outputs and ends with a coordinate transformation matrix from sensor coordinates to the coordinates used for integrating the sensed accelerations.

3.7.5.2 Coning Motion This type of motion is a problem for attitude integration when the frequency of motion is near or above the sampling frequency.

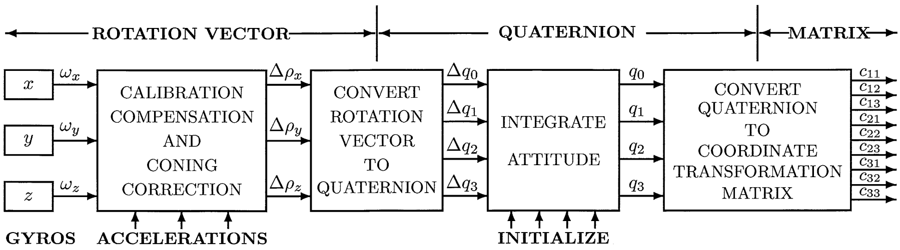

Fig. 3.16 Strapdown attitude representations.

It is usually a consequence of host vehicle frame vibration modes where the INS is mounted, and INS shock and vibration isolation is often designed to eliminate or substantially reduce this type of rotational vibration.

Coning motion is an example of an attitude trajectory (i.e., attitude as a function of time) for which the integral of attitude rates does *not* equal the attitude change. An example trajectory would be

$$\boldsymbol{\rho}(t) = \theta_{\text{cone}} \begin{bmatrix} \cos(\Omega_{\text{coning}} t) \\ \sin(\Omega_{\text{coning}} t) \\ 0 \end{bmatrix} \tag{3.22}$$

$$\dot{\boldsymbol{\rho}}(t) = \theta_{\text{cone}} \Omega_{\text{coning}} \begin{bmatrix} -\sin(\Omega_{\text{coning}} t) \\ \cos(\Omega_{\text{coning}} t) \\ 0 \end{bmatrix}, \tag{3.23}$$

where

θ_{cone} is called the *cone angle* of the motion,
Ω_{coning} is the *coning frequency* of the motion, as illustrated in Fig. 3.17.

The coordinate transformation matrix from body coordinates to inertial coordinates (Eq. B.112 of Appendix B) will be

$$\begin{aligned}\mathbf{C}^{\text{body}}_{\text{inertial}}(\boldsymbol{\rho}) &= \cos\theta \mathbf{I} \\ &+ (1-\cos\theta)\begin{bmatrix} \cos(\Omega_{\text{coning}} t)^2 & \sin(\Omega_{\text{coning}} t)\cos(\Omega_{\text{coning}} t) & 0 \\ \sin(\Omega_{\text{coning}} t)\cos(\Omega_{\text{coning}} t) & \sin(\Omega_{\text{coning}} t)^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &+ \sin\theta \begin{bmatrix} 0 & 0 & \sin(\Omega_{\text{coning}} t) \\ 0 & 0 & -\cos(\Omega_{\text{coning}} t) \\ -\sin(\Omega_{\text{coning}} t) & \cos(\Omega_{\text{coning}} t) & 0 \end{bmatrix},\end{aligned} \tag{3.24}$$

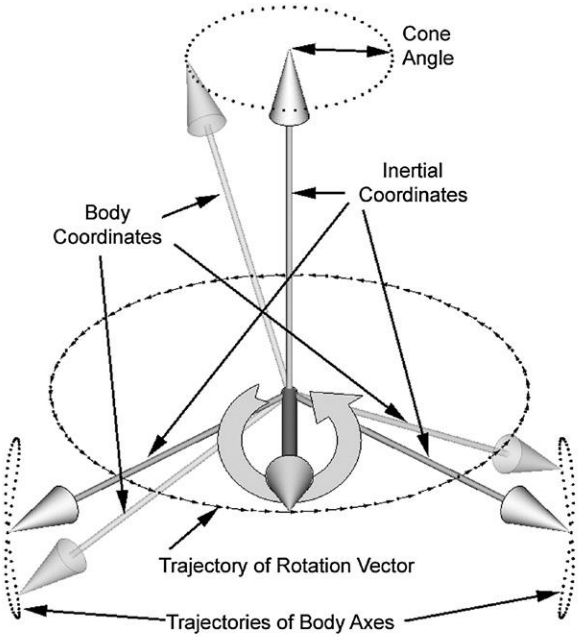

Fig. 3.17 Coning motion.

and the measured inertial rotation rates in body coordinates will be

$$\boldsymbol{\omega}_{\text{body}} = \mathbf{C}_{\text{body}}^{\text{inertial}} \dot{\boldsymbol{\rho}}_{\text{inertial}} \tag{3.25}$$

$$= \theta_{\text{cone}} \Omega_{\text{coning}} \left[\mathbf{C}_{\text{inertial}}^{\text{body}} \right]^T \begin{bmatrix} -\sin(\Omega_{\text{coning}} t) \\ \cos(\Omega_{\text{coning}} t) \\ 0 \end{bmatrix} \tag{3.26}$$

$$= \begin{bmatrix} -\theta_{\text{cone}} \Omega_{\text{coning}} \sin(\Omega_{\text{coning}} t) \cos(\theta_{\text{cone}}) \\ \theta_{\text{cone}} \Omega_{\text{coning}} \cos(\Omega_{\text{coning}} t) \cos(\theta_{\text{cone}}) \\ -\sin(\theta_{\text{cone}}) \theta_{\text{cone}} \Omega_{\text{coning}} \end{bmatrix}. \tag{3.27}$$

The integral of $\boldsymbol{\omega}_{\text{body}}$

$$\int_{s=0}^{t} \boldsymbol{\omega}_{\text{body}}(s) ds = \begin{bmatrix} -\theta_{\text{cone}} \cos(\theta_{\text{cone}}) [1 - \cos(\Omega_{\text{coning}} t)] \\ \theta_{\text{cone}} \cos(\theta_{\text{cone}}) \sin(\Omega_{\text{coning}} t) \\ -\sin(\theta_{\text{cone}}) \theta_{\text{cone}} \Omega_{\text{coning}} t \end{bmatrix}, \tag{3.28}$$

which is what a rate integrating gyroscope would measure.

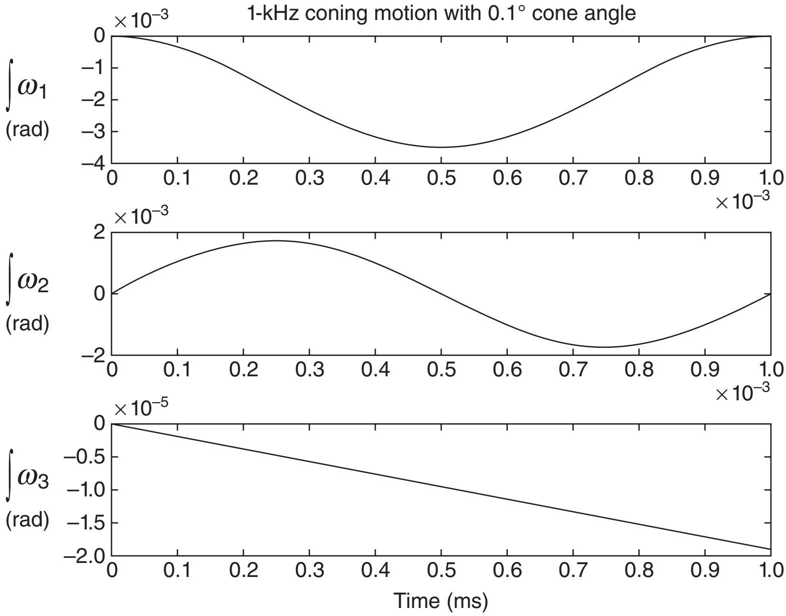

Fig. 3.18 Coning error for 0.1° cone angle, 1-kHz coning rate.

The solutions for $\theta_{\text{cone}} = 0.1°$ and $\Omega_{\text{coning}} = 1\,\text{kHz}$ are plotted over one cycle (1 ms) in Fig. 3.18. The first two components are cyclical, but the third component accumulates linearly over time at about -1.9×10^{-5} rad in 10^{-3} s, which is a bit more than −1 deg/s. *This is why coning error compensation is important.*

3.7.5.3 Rotation Vector Implementation This implementation is primarily used at a faster sampling rate than the nominal sampling rate (i.e., that required for resolving measured accelerations into navigation coordinates). It is used to remove the nonlinear effects of coning and skulling motion that would otherwise corrupt the accumulated angle rates over the nominal intersample period. This implementation is also called a "coning correction."

Bortz Model for Attitude Dynamics This exact model for attitude integration based on measured rotation rates and rotation vectors was developed by John Bortz [1]. It represents ISA attitude with respect to the reference inertial coordinate frame in terms of the rotation vector $\boldsymbol{\rho}$ required to rotate the reference inertial coordinate frame into coincidence with the sensor-fixed coordinate frame, as illustrated in Fig. 3.19.

The Bortz dynamic model for attitude then has the form

$$\dot{\boldsymbol{\rho}} = \boldsymbol{\omega} + \mathbf{f}_{\text{Bortz}}(\boldsymbol{\omega}, \boldsymbol{\rho}), \tag{3.29}$$

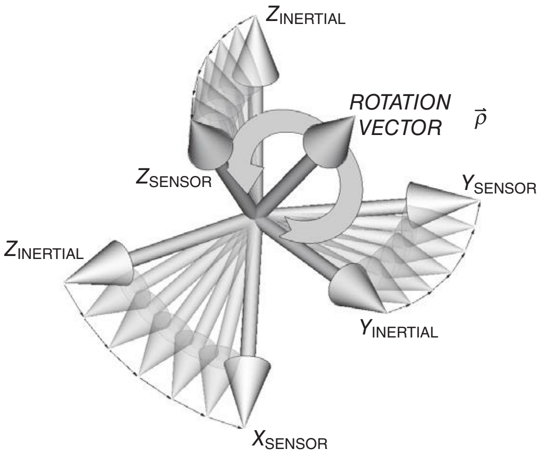

Fig. 3.19 Rotation vector representing coordinate transformation.

where $\boldsymbol{\omega}$ is the vector of measured rotation rates. The Bortz "noncommutative rate vector"

$$\mathbf{f}_{\text{Bortz}}(\boldsymbol{\omega}, \boldsymbol{\rho}) = \frac{1}{2}\boldsymbol{\rho} \otimes \boldsymbol{\omega} + \frac{1}{\|\boldsymbol{\rho}\|^2}\left\{1 - \frac{\|\boldsymbol{\rho}\|\sin(\|\boldsymbol{\rho}\|)}{2[1-\cos(\|\boldsymbol{\rho}\|)]}\right\}\boldsymbol{\rho} \otimes (\boldsymbol{\rho} \otimes \boldsymbol{\omega}) \tag{3.30}$$

$$|\boldsymbol{\rho}| < \frac{\pi}{2}. \tag{3.31}$$

Equation 3.29 represents the rate of change of attitude as a nonlinear differential equation that is linear in the measured instantaneous body rates $\boldsymbol{\omega}$. Therefore, by integrating this equation over the nominal intersample period $[0, \Delta t]$ with initial value $\boldsymbol{\rho}(0) = 0$, an exact solution of the body attitude change over that period can be obtained in terms of the net rotation vector

$$\Delta\boldsymbol{\rho}(\Delta t) = \int_0^{\Delta t} \dot{\boldsymbol{\rho}}(\boldsymbol{\rho}(s), \boldsymbol{\omega}(s))\, ds, \tag{3.32}$$

which avoids all the noncommutativity errors and satisfies the constraint of Eq. 3.31 so long as the body cannot turn 180° in one sample interval Δt. In practice, the integral is done numerically with the gyro outputs ω_1, ω_2, ω_3 sampled at intervals $\delta t = \Delta t$. The choice of δt is usually made by analyzing the gyro outputs under operating conditions (including vibration isolation), and selecting a sampling freqency $1/\delta t$ well above the Nyquist frequency for the observed attitude rate spectrum. The frequency response of the gyros also enters into this design analysis.

The MATLAB® function `fBortz.m` on the accompanying website calculates $\mathbf{f}_{\text{Bortz}}(\boldsymbol{\omega})$ defined by Eq. 3.30.

3.7.5.4 Quaternion Implementation The quaternion representation of vehicle attitude is the most reliable, and it is used as the "holy point" of attitude representation. Its value is maintained using the incremental rotations $\Delta\boldsymbol{\rho}$ from the rotation vector representation, and the resulting values are used to generate the coordinate transformation matrix for accumulating velocity changes in inertial coordinates.

Quaternions represent three-dimensional attitude on the three-dimensional surface of the four-dimensional sphere, much like two-dimensional directions can be represented on the two-dimensional surface of the three-dimensional sphere.

Converting Incremental Rotations to Incremental Quaternions An incremental rotation vector $\Delta\boldsymbol{\rho}$ from the Bortz coning correction implementation of Eq. 3.32 can be converted to an equivalent incremental quaternion $\Delta\mathbf{q}$ by the operations

$$\Delta\theta = |\Delta\boldsymbol{\rho}| \text{ (rotation angle in radians)} \tag{3.33}$$

$$\mathbf{u} = \frac{1}{\theta}\Delta\boldsymbol{\rho} \tag{3.34}$$

$$= \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \text{(unit vector)} \tag{3.35}$$

$$\Delta\mathbf{q} = \begin{bmatrix} \cos\left(\frac{\theta}{2}\right) \\ u_1 \sin\left(\frac{\theta}{2}\right) \\ u_2 \sin\left(\frac{\theta}{2}\right) \\ u_3 \sin\left(\frac{\theta}{2}\right) \end{bmatrix} \tag{3.36}$$

$$= \begin{bmatrix} \Delta q_0 \\ \Delta q_1 \\ \Delta q_2 \\ \Delta q_3 \end{bmatrix} \text{(unit quaternion).} \tag{3.37}$$

Quaternion Implementation of Attitude Integration If

$\mathbf{q}_{k-1}$ is the quaternion representing the prior value of attitude,
$\Delta\mathbf{q}$ is the quaternion representing the change in attitude, and
$\mathbf{q}_k$ is the quaternion representing the updated value of attitude,

then the update equation for quaternion representation of attitude is

$$\mathbf{q}_k = \Delta\mathbf{q} \times \mathbf{q}_{k-1} \times \Delta\mathbf{q}^*, \tag{3.38}$$

where the post superscript * represents the conjugate of a quaternion,

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}^\star \stackrel{\text{def}}{=} \begin{bmatrix} q_1 \\ -q_2 \\ -q_3 \\ -q_4 \end{bmatrix}. \tag{3.39}$$

3.7.5.5 Direction Cosines Implementation The coordinate transformation matrix $\mathbf{C}^{\text{body}}_{\text{inertial}}$ from body-fixed coordinates to inertial coordinates is needed for transforming discretized velocity changes measured by accelerometers into inertial coordinates for integration. The quaternion representation of attitude is used for computing $\mathbf{C}^{\text{body}}_{\text{inertial}}$.

Quaternions to Direction Cosines Matrices The direction cosines matrix $\mathbf{C}^{\text{body}}_{\text{inertial}}$ from body-fixed coordinates to inertial coordinates can be computed from its equivalent unit quaternion representation,

$$\mathbf{q}^{\text{body}}_{\text{inertial}} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}, \tag{3.40}$$

as

$$\mathbf{C}^{\text{body}}_{\text{inertial}} = (2q_0^2 - 1)\mathbf{I}_3 + 2\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} \times \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}^T - 2q_0 \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} \otimes \tag{3.41}$$

$$= \begin{bmatrix} (2q_0^2 - 1 + 2q_1^2) & (2q_1q_2 + 2q_0q_3) & (2q_1q_3 - 2q_0q_2) \\ (2q_1q_2 - 2q_0q_3) & (2q_0^2 - 1 + 2q_2^2) & (2q_2^2 + 2q_0q_1) \\ (2q_1q_3 + 2q_0q_2) & (2q_2^2 - 2q_0q_1) & (2q_0^2 - 1 + 2q_3^2) \end{bmatrix}. \tag{3.42}$$

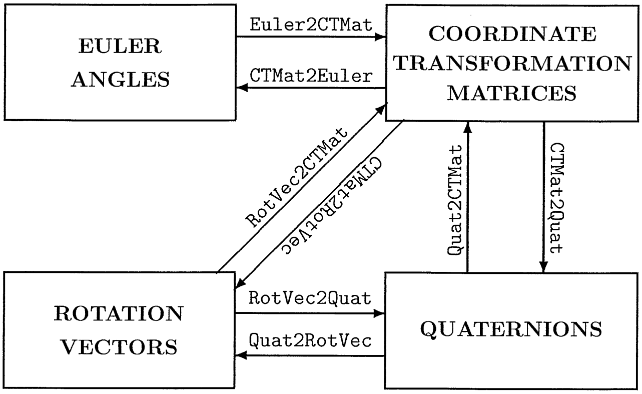

Fig. 3.20 Attitude representation formats and MATLAB® transformations.

3.7.5.6 MATLAB® Implementations The diagram in Fig. 3.20 shows four different representations used for relative attitudes and the names of the MATLAB® script m-files (i.e., with the added ending .m) on the accompanying website for transforming from one representation to another.

3.7.6 Strapdown Navigation Implementation

The basic signal processing functions for strapdown INS navigation are diagrammed in Fig. 3.21, where the common symbols used in Fig. 3.15 have the same meaning as before, and

$\mathbf{G}$ is the estimated gravitational acceleration, computed as a function of estimated position.

$\mathrm{POS}_{\mathrm{NAV}}$ is the estimated position of the host vehicle in navigation coordinates.

$\mathrm{VEL}_{\mathrm{NAV}}$ is the estimated velocity of the host vehicle in navigation coordinates.

$\mathrm{ACC}_{\mathrm{NAV}}$ is the estimated acceleration of the host vehicle in navigation coordinates, which may be used for trajectory control (i.e., vehicle guidance).

$\mathrm{ACC}_{\mathrm{SENSOR}}$ is the estimated acceleration of the host vehicle in sensor-fixed coordinates, which may be used for steering stabilization and control.

$\mathbf{C}_{\mathrm{NAV}}^{\mathrm{SENSOR}}$ is the 3×3 coordinate transformation matrix from sensor-fixed coordinates to navigation coordinates, representing the attitude of the sensors in navigation coordinates.

Ω_{SENSOR} is the estimated angular velocity of the host vehicle in sensor-fixed coordinates, which may be used for vehicle attitude stabilization and control.

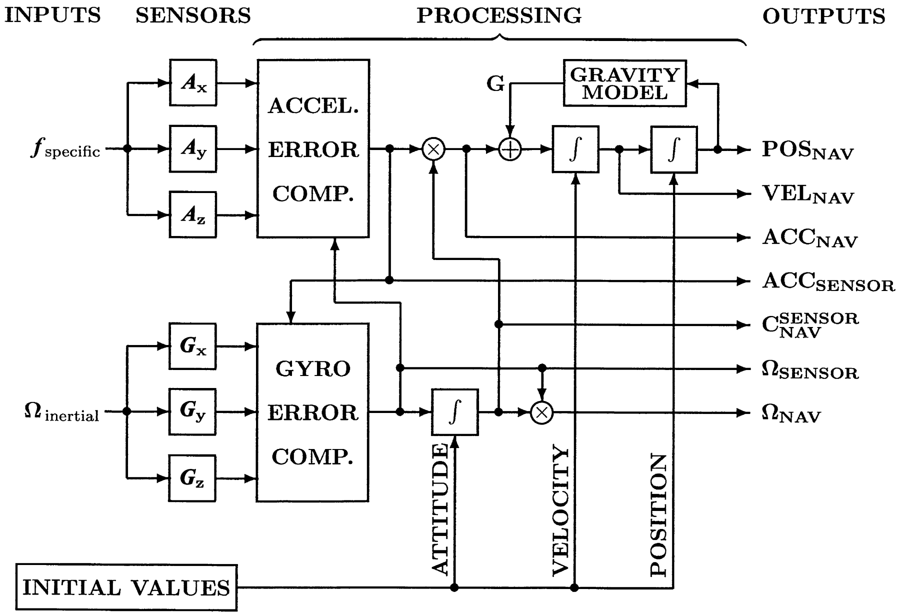

Fig. 3.21 Essential navigation signal processing for strapdown INS.

Ω_{NAV} is the estimated angular velocity of the host vehicle in navigation coordinates, which may be used in a vehicle pointing and attitude control loop.

The essential processing functions include double integration (represented by boxes containing integration symbols) of acceleration to obtain position, and computation of (unsensed) gravitational acceleration as a function of position. The sensed angular rates also need to be integrated to maintain the knowledge of sensor attitudes. The initial values of all the integrals (i.e., position, velocity, and attitude) must also be known before integration can begin.

The position vector POS_{NAV} is the essential navigation solution. The other outputs shown are not needed for all applications, but most of them (except Ω_{NAV}) are intermediate results that are available "for free" (i.e., without requiring further processing). The velocity vector VEL_{NAV}, for example, characterizes speed and heading, which are also useful for correcting the course of the host vehicle to bring it to a desired location. Most of the other outputs shown would be required for implementing control of an unmanned or autonomous host vehicle to follow a desired trajectory and/or to bring the host vehicle to a desired final position.

Navigation functions that are not shown in Fig. 3.21 include

1. How initialization of the integrals for position, velocity, and attitude is implemented. Initial position and velocity can be input from other

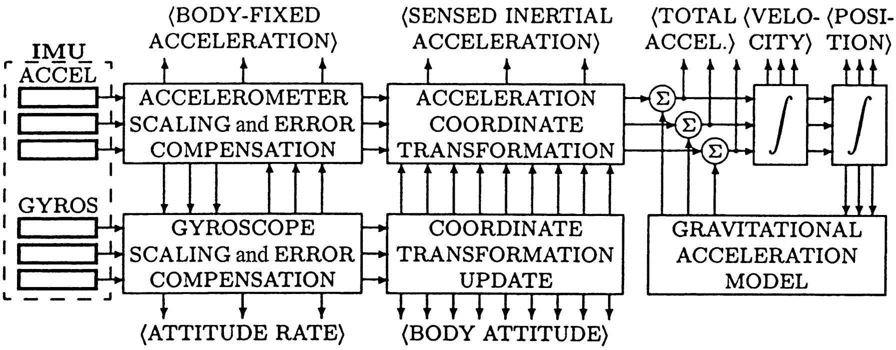

Fig. 3.22 Outputs (in angular brackets) of simple strapdown INS.

sources (e.g., GNSS), and attitude can be inferred from some form of trajectory matching (e.g., using GNSS) or by **gyrocompassing**.

2. How attitude rates are integrated to obtain attitude, described in Section 3.7.5.
3. For the case that navigation coordinates are earth-fixed, the computation of navigational coordinate rotation due to earthrate as a function of position, and its summation with sensed rates before integration.
4. For the case that navigation coordinates are locally level, the computation of the rotation rate of navigation coordinates due to vehicle horizontal velocity, and its summation with sensed rates before integration.
5. Calibration of the sensors for error compensation. If the errors are sufficiently stable, it needs to be done only once. Otherwise, it can be implemented using GNSS/INS integration techniques (Chapter 12).

Figure 3.22 is a process flow diagram for the same implementation, arranged such that the variables available for other functions is around the periphery. These are the sorts of variables that might be needed for driving GNSS phase-tracking, cockpit displays, antennas, weaponry, sensors, or other surveillance assets.

3.7.7 Navigation Computer and Software Requirements

It is not a problem today, but throughout the early history of INS development, the pace of computer development had been a limiting factor in INS development. There was no computer industry until the 1950s, no flight-qualified computers until the 1960s, and no suitable microprocessor technology until the 1970s.

The following subheadings list some of the requirements placed on navigation computers and software that tend to set them apart.

3.7.7.1 Physical and Operational Requirements These include

1. size, weight, form factor, available input power;
2. environmental conditions such as shock/vibration, temperature, electromagnetic interference (EMI);
3. memory (how much and how fast), throughput (operations/s), word length/precision;
4. time required between power-on and full operation, and minimum time between turn-off and turn-on (e.g., some vehicles shut down all power during fueling);
5. reliability, shelf life and storage requirements;
6. operating life; some applications (e.g., missiles) have operating lifetimes of minutes or seconds, others (e.g., military and commercial aircraft) may operate nearly continuously for decades; and
7. additional application-specific requirements, such as radiation hardening or the ability to function during high dynamic loading.

Most of these and their associated system interfaces have largely been standardized today, especially for military applications.

3.7.7.2 Operating Systems Inertial navigation is a *real-time* process. The tasks of sampling the sensor outputs and of integrating attitude rates, accelerations, and velocities must be scheduled at precise time intervals, and the results must be available after limited delay times. The top-level operating system that prioritizes and schedules these and other tasks must be a *real-time operating system* (RTOS). It may also be required to communicate with other computers in various ways.

3.7.7.3 Interface Requirements These include not only the operational interfaces to sensors and displays but may also include communications interfaces and specialized computer interfaces to support navigation software development and verification.

3.7.7.4 Software Development Because INS failures could put host vehicle crews and passengers at risk, it is very important during system development to demonstrate high reliability of the software. Ada is often used as the programming language because it has many built-in features to assure compliance. INS software is usually developed offline on a general-purpose computer interfaced to the navigation computer. Software development environments for INS typically include code editors, cross compilers, navigation computer emulators, hardware simulators, hardware-in-the-loop interfaces and specialized source-code-online interfaces to the navigation computer for monitoring, debugging and verifying the navigation software on the navigation computer.

Software developed for manned missions must be acceptably reliable, which requires metrics for demonstrating reliability and testing for verification.

3.8 INS PERFORMANCE STANDARDS

3.8.1 Free Inertial Operation

Operation of an INS without external aiding of any sort is called **free inertial** or **pure inertial**. Because free inertial navigation in the near-earth gravitational environment is unstable in the vertical direction (due to the falloff of gravity with increasing altitude), aiding by other sensors (e.g., barometric altimeters for aircraft or surface vehicles, radar altimeters for aircraft over water, or hydrostatic pressure for submersibles) is required to avoid vertical error instability. For that reason, performance of free INSs is usually specified for horizontal position errors only.

3.8.2 INS Performance Metrics

Oversimplified error model: INS position is initialized by knowing where you are starting from at initial time $\boldsymbol{t_0}$. The position error may start out very small, but it tends to increase with time due to the influence of sensor errors. Double integration of accelerometer output errors is a major source of this growth over time. Experience has shown that the variance and standard deviation of horizontal position error,

$$\boldsymbol{\sigma}^2_{\text{position}}(t) \propto (t - t_0)^2 \tag{3.43}$$

$$\boldsymbol{\sigma}_{\text{position}}(t) \approx C \times |t - t_0|, \tag{3.44}$$

with unknown positive constant C. This constant C would then characterize performance of an INS in terms of how fast its RMS position error grows.

A problem with this model is that actual horizontal INS position errors are two-dimensional, and we would need a **2 × 2 covariance matrix** in place of C. That would not be very useful in practice. As an alternative, we replace C with something more intuitive and practical.

CEP is the radius of a horizontal circle centered at the estimated position, and of sufficient radius such that it is equally probable that the true horizontal position is inside or outside the circle. CEP is the acronym for either **circular error probable** or for **circle of equal probability**, depending on which is easier to remember.

CEP rate is the time rate of change of CEP. Traditional units of CEP rate are **nautical miles per hour** or **kilometers per hour**. The nautical mile was originally intended to designate a surface distance equivalent to 1 arc minute of latitude change at sea level, but that depends on latitude. The Système International (SI)-derived nautical mile is 1.852 km.

3.8.3 Performance Standards

In the 1970s, before GPS became a reality, the U.S. Air Force had established the following standard levels of performance for INS:

High-accuracy systems have free inertial CEP rates in the order of 0.1 nmi/h (≈**185** m/h) or better. This is the order of magnitude in accuracy required for intercontinental ballistic missiles (ICBMs) and missile-carrying submarines, for example.

Medium-accuracy systems have free inertial CEP rates in the order of 1 nmi/h (≈**1.85** km/h). This is the level of accuracy deemed sufficient for most military and commercial aircraft [2].

Low-accuracy systems have free inertial CEP rates in the order of 10 nmi/h (≈**18.5** km/h) or worse. This range covered the requirements for many short-range standoff weapons such as guided artillery or tactical rockets.

Table 3.2 lists accelerometer and gyroscope performance ranges compatible with the standards.

However, after GPS became available, GPS/INS integration could make a low-accuracy INS behave more like a high-accuracy INS.

3.9 TESTING AND EVALUATION

The final stage of the development cycle is testing and performance evaluation. For stand-alone inertial systems, this usually proceeds from the laboratory to a succession of host vehicles, depending on the application.

3.9.1 Laboratory Testing

Laboratory testing is used to evaluate sensors before and after their installation in the ISA, and then to evaluate the system implementation during

TABLE 3.2. INS and Inertial Sensor Performance Ranges

	Performance Ranges			
System or Sensor	High	Medium	Low	Units
INS	$\leq 10^{-1}$	≈ 1	≥ 10	nmi/h[a]
Gyroscopes	$\leq 10^{-3}$	$\approx 10^{-2}$	$\geq 10^{-1}$	deg/h
Accelerometers	$\leq 10^{-7}$	$\approx 10^{-6}$	$\geq 10^{-5}$	g (9.8 m/s^2)

[a]Nautical miles per hour CEP rate.

operation. The navigation solution from a stationary system should remain stationary, and any deviation is due to navigation errors. Testing with the system stationary can also be used to verify that position errors due to intentional initial velocity errors follow a path predicted by Schuler oscillations (≈**84.4**-min period, described in Chapter 11) and the Coriolis effect. If not, there is an implementation error. Other laboratory testing may include controlled tilts and rotations to verify the attitude estimation implementations, and detect any uncompensated sensitivities to rotation and acceleration.

Additional laboratory testing may be required for specific applications. Systems designed to operate aboard Navy ships, for example, may be required to meet their performance requirements under dynamic disturbances at least as bad as those to be expected aboard ships under the worst sea conditions. This may include what is known as a "Scoresby test," used at the U.S. Naval Observatory in the early twentieth century for testing gyrocompasses. Test conditions may include roll angles of **±80°** and pitch angles of **±15°**, at varying periods in the order of a second.

Drop tests (for survival testing) and shake-table or centrifuge tests (for assessing acceleration capabilities) can also be done in the laboratory.

3.9.2 Field Testing

After laboratory testing, systems are commonly evaluated next in highway testing.

Systems designed for tactical aircraft must be designed to meet their performance specifications under the expected peak dynamic loading, which is generally determined by the pilot's limitations.

Systems designed for rockets must be tested under conditions expected during launch, sometimes as a "piggyback" payload during the launch of a rocket for another purpose. Accelerations can reach around 3 g for manned launch vehicles, and much higher for unmanned launch vehicles.

In all cases, GNSS has become an important part of field instrumentation. The Central Inertial Guidance Test Facility (CIGTF) at Holoman AFB has elaborate range instrumentation for this purpose. This facility is used by NASA and Department of Defense (DoD) agencies for INS and GNSS/INS testing in a range of host vehicles.

3.10 SUMMARY

1. Inertial navigation has a rich technology base—more than can be covered in a single book, and certainly not in one chapter. Unfortunately, some of best technology is either classified or proprietary. Otherwise, there is some good open-source literature:
   (a) Titterton and Weston [9] is a good source for additional information on strapdown hardware and software.

   (b) Paul Savage's two-volume tome [8] on strapdown system implementations is also rather thorough.
   (c) Chapter 5 of Ref. 3 and the references therein include some recent developments.
   (d) Journals of the Institute of Electrical and Electronics Engineers (IEEE), Institution of Electrical Engineers (IEE), Institute of Navigation, and other professional engineering societies generally have the latest developments on inertial sensors and systems.
   (e) In addition, the World Wide Web includes many surveys and reports on inertial sensors and systems.
2. Inertial navigation accuracy is mostly limited by inertial sensor accuracy.
3. The accuracy requirements for inertial sensors cannot always be met within manufacturing tolerances. Some form of calibration is usually required for compensating the residual errors.
4. INS accuracy degrades over time, and the most accurate systems generally have the shortest mission times. For example, ICBMs only need their inertial systems for a few minutes.
5. Performance of inertial systems is commonly specified in terms of CEP rate.
6. Accelerometers cannot measure gravitational acceleration.
7. Both inertial and satellite navigation require accurate models of the earth's gravitational field.
8. Both navigation modes also require an accurate model of the shape of the earth.
9. The first successful navigation systems were gimbaled, in part because the computer technology required for strapdown implementations was decades away. That has not been a problem for about four decades.
10. Gimbaled systems tend to be more accurate and more expensive than strapdown systems.
11. The more reliable attitude implementations for strapdown systems use quaternions to represent attitude.
12. Systems traditionally go through a testing and evaluation process to verify performance.
13. Before testing and evaluation of an INS, its expected performance is commonly evaluated using the analytical models of Chapter 11.

PROBLEMS

Refer to Appendix B for coordinate system definitions and satellite orbit equations.

3.1 Which, if any, of the following coordinate systems is not rotating?

(a) NED

(b) ENU

(c) ECEF

(d) ECI

(e) Moon-centered, moon-fixed.

3.2 What is the minimum number of two-axis gyroscopes (i.e., gyroscopes with two, independent, orthogonal input axes) required for inertial navigation?

(a) 1

(b) 2

(c) 3

(d) Not determined.

3.3 What is the minimum number of gimbal axes required for gimbaled inertial navigators in fully maneuverable host vehicles? Explain your answer.

(a) 1

(b) 2

(c) 3

(d) 4

3.4 Define specific force.

3.5 An ISA operating at a fixed location on the surface of the earth would measure

(a) no acceleration

(b) 1 g acceleration downward

(c) 1 g acceleration upward.

3.6 Explain why an INS is not a good altimeter.

3.7 The inertial rotation rate of the earth is

(a) 1 revolution per day

(b) 15 deg/h

(c) 15 arc seconds/s

(d) ≈15.0411 arc seconds/s.

3.8 Define the CEP and the CEP rate for an INS.

3.9 The CEP rate for a medium-accuracy INS is in the order of

(a) 2 m/s

(b) 200 m/h

(c) 2000 m/h

(d) 20 km/h

3.10 In the one-dimensional line land world of Section 3.7.1, an INS requires no gyroscopes. How many gyroscopes would be required for two-dimensional navigation in flat land?

3.11 Derive the equivalent formulas in terms of Y (yaw angle), P (pitch angle), and R (roll angle) for unit vectors 1_R, 1_P, 1_Y in NED coordinates and 1_N, 1_E, 1_D in RPY coordinates.

3.12 Explain why accelerometers cannot sense gravitational accelerations.

3.13 Show that the matrix $\mathbf{C}_{\text{inertial}}^{\text{body}}$ defined in Eq. 3.42 is orthogonal by showing that $\mathbf{C}_{\text{inertial}}^{\text{body}} \times \mathbf{C}_{\text{inertial}}^{\text{body T}} = \mathbf{I}$, the identity matrix. (Hint: Use $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$.)

3.14 Calculate the numbers of computer multiplies and adds required for

(a) gyroscope scale factor/misalignment/bias compensation (Eq. 3.4 with $N = 3$),

(b) accelerometer scale factor/misalignment/bias compensation (Eq. 3.4 with $N = 3$), and

(c) transformation of accelerations to navigation coordinates (Fig. 3.22) using quaternion rotations (see Appendix B section on quaternion algebra).

If the INS performs these 100 times per second, how many operations per second will be required?

REFERENCES

[1] J. E. Bortz, "A New Mathematical Formulation for Strapdown Inertial Navigation," *IEEE Transactions on Aerospace and Electronic Systems* **AES-6**, 61–66 (1971).

[2] Chairman of Joint Chiefs of Staff, U.S. Department of Defense, *2003 CJCS Master Positioning, Navigation and Timing Plan*, Report CJCSI 6130.01C, March 2003.

[3] P. D. Groves, *Principles of GNSS, Inertialm and Multisensor Integrated Navigation Systems*. Artech House, London, 2008.

[4] *IEEE Standard for Inertial Sensor Terminology*, IEEE Standard 528-2001, Institute of Electrical and Electronics Engineers, New York, 2001.

[5] *IEEE Standard for Inertial System Terminology*, IEEE Standard 1559, Institute of Electrical and Electronics Engineers, New York, 2007.

[6] D. Mackenzie, *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance*. MIT Press, Cambridge, MA, 1990.

[7] F. K. Mueller, "A History of Inertial Navigation," *Journal of the British Interplanetary Society* **38**, 180–192 (1985).

[8] P. G. Savage, *Introduction to Strapdown Inertial Navigation Systems*, Vols. 1 & 2. Strapdown Associates, Maple Plain, MN, 1996.

[9] D. H. Titterton and J. L. Weston, *Strapdown Inertial Navigation Technology*. IEE, London, 2004.