Aboelmagd Noureldin
Tashfeen B. Karamat
Jacques Georgy

Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration

Springer

Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration

Aboelmagd Noureldin
Tashfeen B. Karamat · Jacques Georgy

Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration

Springer

Dr. Aboelmagd Noureldin
Department of Electrical and Computer Engineering
Royal Military College of Canada/
Queen's University
Kingston
Canada

Tashfeen B. Karamat
Department of Electrical
and Computer Engineering
Queen's University
Kingston
Canada

Jacques Georgy
Trusted Positioning Inc.
Calgary
Canada

ISBN 978-3-642-30465-1 ISBN 978-3-642-30466-8 (eBook)
DOI 10.1007/978-3-642-30466-8
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012945733

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

I dedicate this book to my mother and father for their love and sacrifices and to my wife and three sons, Abdelrahman, Yehia and Tareq for their support, encouragement and patience

Aboelmagd Noureldin

To my parents, Karamat and Safeena for their love, my brother Khaver for his kind patronage after the early demise of my parents, my wife Shazia and my sons Fahaam and Saarim for their unwavering support

Tashfeen Karamat

To my wife Sarah for her great love and valuable support, to my parents, Ford and Lucie, for all their great love and its acts, continuous encouragement and support throughout my life, and to my sister Basma for her love and encouragement

Jacques Georgy

Contents

Abbreviations

1D	One dimension
2D	Two dimension
3D	Three dimension
ADOP	Ambiguity DOP
AFM	Ambiguity function method
A-GPS	Assisted GPS
AI	Artificial intelligence
AOA	Angle of arrival
AR	Ambiguity resolution
AR	Autoregressive
ARNS	Aeronautical radio navigation services
ARW	Angle random walk
BD-1	Beidou-1
BD-2	Beidou-2
BOC	Binary offset carrier
BPSK	Binary phase shifted key
CASC	China aerospace science and technology
CAST	China academy of space technology
CDMA	Code division multiple access
CNSS	Compass navigation satellite system
CORS	Continuously operating reference station
CTP	Conventional terrestrial pole
CUPT	Coordinate update
CWAAS	Canadian wide area augmentation system
DCM	Direction cosine matrices
DD	Double difference
DGPS	Differential GPS
DLL	Delay-lock loops
DOP	Dilution of precision
DR	Dead-reckoning

DTG	Dynamically tuned gyroscopes
DVB-T	Digital video broadcasting - terrestrial
ECEF	Earth-centered earth-fixed
ECI	Earth-centered inertial
EGNOS	European geostationary navigation overlay service
EKF	Extended Kalman filter
ENU	East North Up
EOTD	Enhanced observed time difference
ESA	European space agency
FAA	Federal aviation administration
FASF	Fast ambiguity search filter
FOC	Full operational capability
FOG	Fiber optic gyroscopes
GAGAN	Geo-augmented navigation system
GBAS	Ground-based augmentation systems
GDOP	Geometric dilution of precision
GDPS	Global differential GPS
GLONASS	Global navigation satellite system
GM	Gauss-Markov
GNSS	Global navigation satellite systems
GPS	Global positioning system
GRAS	Ground-based regional augmentation system
GRS	Geographic reference system
GSM	Global system of mobile
HDOP	Horizontal dilution of precision
HOT	Higher order terms
HRG	Hemispherical resonant gyroscopes
IA	Integer ambiguity
IFOG	Interferometric fiber-optic gyroscopes
IGS	International GNSS service
IMU	Inertial measurement unit
INS	Inertial navigation system
IOC	Initial operational capability
IRNSS	Indian regional navigational satellite system
ISA	Inertial sensor assembly
ISDB	Integrated services digital broadcasting
KF	Kalman filter
LAAS	Local area augmentation system
LADGPS	Local area DGPS
LAMBDA	Least-squares ambiguity decorrelation adjustment
LIDAR	Light detection and ranging
LKF	Linearized Kalman filter
LLF	Local-level frame

LOS	Line of sight
LSAST	Least-squares ambiguity search technique
MBOC	Multiplexed binary offset carrier
MCS	Master control station
MEMS	Micro-electro-mechanical system
MSAS	Multifunction transport satellite (MTSAT) satellite augmentation system
MTSAT	Multifunction transport satellite
NED	North East Down
NGA	National geospatial-intelligence agency
NGDGPS	Nationwide differential GPS system
NTSC	National television system committee
OTDOA	Observed time difference of arrival
PDOP	Position dilution of precision
PF	Particle filter
PLL	Phase-lock loops
PPS	Precise positioning service
PRN	Pseudo random noise
QZSS	Quasi-zenith satellite system
RAAN	Right ascension of the ascending node
RF	Radio frequency
RISS	Reduced inertial sensor system
RLG	Ring laser gyroscopes
RMS	Root mean square
RNSS	Radio navigation satellite services
RS	Reference station
RSD	Receiver single difference
SA	Selective availability
SBAS	Space-based augmentation systems
SECAM	Séquentiel couleur à mémoire
SNAS	Satellite navigation augmentation system
SPAN	Synchronized position attitude navigation
SPS	Standard positioning service
SSD	Satellite single difference
STM	State transition matrix
TD	Triple difference
TDOA	Time difference of arrival
TDOP	Time dilution of precision
TEC	Total electron content
TOA	Time of arrival
TTFF	Time-to-first-fix
TV	Television
UERE	User equivalent range error

U-TDOA	Uplink time difference of arrival
UWB	Ultra-wide band
VBA	Vibrating beam accelerometers
VRW	Velocity random walk
WAAS	Wide-area augmentation system
WADGPS	Wide area DGPS
WGS	World geodetic system
WLAN	Wireless local area networks
ZUPT	Zero velocity update

Chapter 1
Introduction

The word 'navigation' comes from Latin *navigare* which means 'to sail'. The word *navigare* itself is derived from *navis*, which stands for 'ship', and *agere*, meaning 'to drive' (Esmat Sep 2007). Early navigation primarily dealt with vessels traveling in sea. However, it has now permeated into every imaginable form of transportation as well as various other applications including location-based services, search and rescue, law enforcement, road and air travel, remote asset tracking, fleet management, intelligence gathering, sports, public safety, and environmental assessment and planning (El-Rabbany 2002). Advances in microelectronics and miniaturization of integrated circuits have facilitated the production of inexpensive inertial sensors, global positioning system (GPS) receivers and powerful computers. This has placed navigation systems within easy reach of low cost applications.

Navigation is a technique for the determination of position and velocity of a moving platform with respect to a known reference, but it can also include the attitude of the platform (Groves Dec 2007). A navigation system can either be autonomous or be dependent on external sources, or in some cases a combination of the two. The fusion of the two systems is traditionally based on the technique of Kalman filtering, developed in 1960 mainly for navigating in space. The sensors for a navigation system typically include accelerometers, gyroscopes and radio receivers. There are two fundamental methods for finding a navigation solution: position fixing and dead reckoning (DR). Position fixing is based on the information of external sources with known locations, with GPS being a typical example. On the other hand, dead reckoning is autonomous and relies on knowledge of the initial location, speed and heading information. Inertial navigation is a dead reckoning system which uses accelerometers and gyroscopes to monitor translational motion and rotational motion respectively.

Position, velocity and attitude are called the navigation states. In cases where only the position is required, the term positioning is used rather than navigation. For this book, the term vehicle will be mostly used for the moving platform whose position and attitude are to be determined.

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_1,

1.1 General Classification of Positioning Techniques

According to Borenstein et al. 1997, positioning techniques (either indoor or outdoor) are divided into seven categories falling in two groups.

1.1.1 Techniques Using Relative Measurements (Known as DR)

1.1.1.1 Odometry

The odometry data is obtained by using sensors that measure the rotation of the wheel axes and the steer axes (e.g. high resolution encoders). Wheel rotation is translated into linear displacement. The advantages of this method are that it has short term accuracy, is low cost and allows high sampling rates. But it cannot take into account wheel slippage. For a specified initial position estimate, the task is to integrate incremental motion information over time. This leads to the disadvantage that any small constant error increases without bound. In particular, orientation errors can cause large position errors that increase with the distance traveled, thus impairing the long term accuracy of the solution.

1.1.1.2 Inertial Navigation

This employs inertial sensors (gyroscopes and accelerometers) which measure the rotation rates and the specific forces from which acceleration can be obtained. Inertial navigation systems are autonomous, which means they are self-contained; i.e. they don't need external references. Starting from a known position and orientation, measurements are integrated once for gyroscopes and twice for accelerometers to provide orientation and position respectively. The positioning solutions obtained tend to drift with time due to the integrations performed, which can lead to unbounded accumulation of errors. Inertial navigation alone, especially with low cost sensors, is thus unsuitable for accurate positioning over an extended period of time.

1.1.2 Techniques Using Absolute Measurements (Known as Reference-based Systems)

1.1.2.1 Electronic Compasses

Although not providing a positioning solution by its own, an electronic compass is a device which uses magnetometers to provide heading measurements relative to

the Earth's magnetic north by observing the direction of the local magnetic field. To convert the compass heading into an actual north heading, the declination angle, which is the angle between geographic and magnetic north, is needed. The declination angle is position dependent, so it is necessary to know the position of the compass in order to calculate the heading relative to geographic north. The magnetic compass has the disadvantage that the local magnetic field is distorted near power lines or metal structures such as bridges and buildings along the trajectory of the vehicle. This can result in large and unpredictable errors in the heading estimated by the compass, in turn making the application of this system to vehicular navigation questionable.

1.1.2.2 Active Beacons

This approach can be used if the moving platform is to navigate in an already known environment, and can provide accurate positioning information. However, very precise mounting of the beacons is required in order to facilitate accurate positioning. Furthermore, their installation and maintenance imposes a relatively high cost. Thus the supporting infrastructure limits the utility of active beacons to specified environments. There are several positioning algorithms that can be used with different active beacon systems, such as the trilateration-based algorithm, the triangulation-based algorithms, and the fingerprinting algorithms. Trilateration is the calculation of a vehicle's position based on distance measurements relative to a known beacon using, for example, time-of-flight information. Triangulation is the calculation of a vehicle's position and possibly its orientation based on the angles at which beacons are seen relative to the moving platform's longitudinal axis. Examples of active beacon systems are wireless communication systems for positioning, such as those which use cellular networks, television (TV) networks, ultra-wide band (UWB) networks and wireless local area networks (WLAN), the latter also known as Wi-Fi systems.

1.1.2.3 Global Navigation Satellite Systems

GNSS is mainly a technology for outdoor navigation. Currently, the most popular example is GPS, which is a constellation of satellites that transmit encoded radio frequency (RF) signals. By means of trilateration, ground receivers can calculate their position using the travel time of the satellites's signals and information about their current location, this being included in the transmitted signal. Calculation of the receiver's latitude, longitude and altitude is possible when its exact distance from a number of satellites is known, three being the theoretical minimum number and four being able to correct for clock bias in the receiver. GPS provides good absolute localization for outdoor navigation, so long as there is sufficient satellite coverage. However, it is normally not available in deep indoor environments.

1.1.2.4 Landmark Navigation

This approach can be used if the moving platform is to navigate in an environment that is well known. Landmarks are distinct objects or features such as geometric shapes that can be detected and distinguished by appropriate sensors on a vehicle. They can be either natural or artificial. Artificial landmarks are objects added to the environment specifically for positioning and navigation, whereas natural ones are already present in the environment. Each landmark must have a fixed position, and the vehicle requires a database of their characteristics and locations. Some landmarks may include additional information like bar-codes. The vehicle must be able to recognize landmarks reliably from its sensor inputs and process the data in a manner that determines its own position.

1.1.2.5 Map-Based Positioning (Or Model Matching)

This approach can be used if the moving platform is to navigate in a specific mapped environment. In this approach, the moving platform uses its sensors to perceive its local environment, and this perception is then compared to a map previously stored in its memory. If a match is found, then the vehicle can calculate its position and orientation in this specific environment. Cameras and laser range finders are examples of sensors that can be used with this type of positioning. The stored map of the environment can be an already available model, or it can be constructed from prior sensor data. This approach can be used only in a structured environment, which means indoor environments and some outdoor environments. There is a somewhat similar technique known as map matching that is normally used together with GPS to determine the position of a vehicle within a street map structure. Although there are similarities between the two techniques, the latter is used to constrain another positioning solution such as GPS, it is not a standalone positioning technique and is not the one discussed in this subcategory.

1.1.3 Combined Systems

For the first category that uses relative measurements (i.e. dead reckoning) the determination of the current vehicle position uses the knowledge of the previous position and the measurement of its latest movements. For the second category of absolute measurements (i.e. reference-based systems) the current vehicle position is calculated by measuring to known reference points but without knowledge of its previous trajectory. Usually two methods or more, involving at least, one from each group, are combined in order to obtain a reliable navigation solution.

1.2 GNSS-Based Positioning Techniques

A GNSS such as GPS is currently popular for outdoor positioning. These systems calculate the position of the receiver by ranging to several visible satellites with known locations using trilateration. The United States' GPS is the most widely used. Another one is the Russian **GLO**bal'naya **NA**vigatsionnaya **S**putnikovaya **S**istema (**GLO**bal **NA**vigation **S**atellite **S**ystem, in English), GLONASS, which is similar to GPS in concept but uses different technical standards. The first satellite of this system was launched in 1982. Like GPS, GLONASS offers a position and navigation service to civilians and an encrypted service to military users (Space-Based Positioning, Navigation and Timing 2008). As of the time of writing, there are 31 satellites in space, 24 of which are operational (FSA 2012).

Galileo is another navigation system being developed by the European Union. The first test satellite was launched in December 2005 and the second in April 2008, primarily to occupy the allotted frequencies and test some of the important technologies. Two operational satellites were launched together in October 2011, with two more are to follow in 2012. The system is expected to reach its initial operational capability (IOC) around mid-decade and its full operational capability (FOC) by the end of this decade (ESA 2012). It is primarily intended for civilian use and will provide two services, at least one of which will be freely available to any user. The final satellite constellation will include 30 satellites.

China has also developed an experimental satellite navigation system known as BeiDou-1 (BD-1) in Chinese for 'compass navigation satellite experimental system'. It was developed by the China academy of space technology (CAST), a subsidiary of the China aerospace science and technology (CASC). The system provides all-weather two-dimensional positioning, navigation and communication for both military and civilian users over most of the East Asia region. It comprises two main BeiDou-1 satellites and one backup. The launch of the third satellite in May 2003 made the system fully operational (Beidou 1 2012). The full navigation system called 'compass navigation satellite system' (CNSS) or BeiDou-2 (BD-2) is in the development stage. It will consist of a constellation of 25–35 satellites, including four in geostationary orbits, and will provide complete coverage of the Earth (Beidou-2 2012). The free service is expected to provide an accuracy of 10 m but the licensed service will be more accurate.

India and Japan are also developing satellite based navigation systems known as the Indian regional navigational satellite system (IRNSS) and the quasi-zenith satellite system (QZSS) respectively.

Since GPS is the most widely used satellite navigation system and has matured since its inception in 1973, it will be the focus of our discussion.

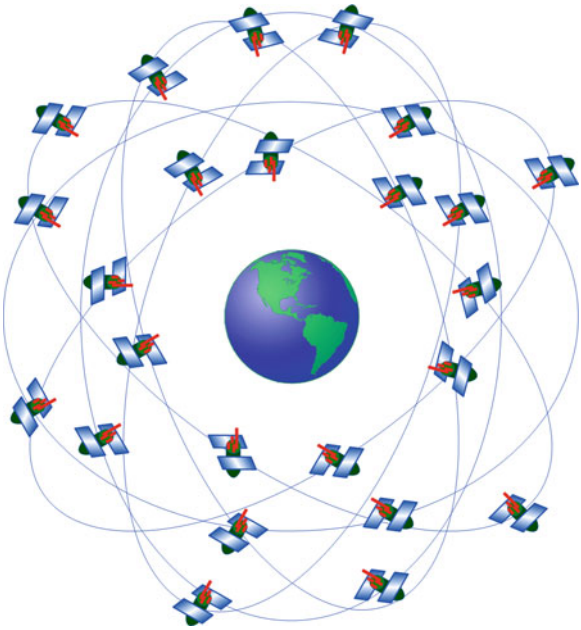


Fig. 1.1 The GPS satellite constellation

1.2.1 Global Positioning System

The US Navy introduced the world's first satellite navigation system in 1961 with the launch of the experimental Transit satellite. This system became operational in 1964. The construction of the global positioning systems (GPS) started in 1973, and it achieved IOC in 1993. It provides accurate position, navigation and timing information on a worldwide basis free of charge to anyone with a GPS receiver, although initially the military service was more accurate than the civilian one.

GPS consists of a nominal constellation of 24 satellites orbiting the Earth (as depicted in Fig. 1.1) at a mean radius of 26,560 km with capacity for some additional ones (Misra and Enge Dec 2001). The geometry of the constellation ensures the visibility of at least four satellites at any location on or near the surface, this being the minimum number to determine a solution. In principle, a GPS receiver requires the ranges to only three satellites in order to calculate its three-dimensional position using the principle of trilateration, but a fourth satellite is needed to estimate the offset of the receiver's clock from the system clock. Figure 1.2 shows the concept of GPS positioning by three satellites. The fact that more than four satellites are usually available serves to improve the accuracy of the solution. The GPS system is discussed in detail in Chap. 3.

GPS (indeed any GNSS) may suffer from outages due to signal blockage, multipath effects, interference or jamming. GPS outages happen in urban canyons, tunnels and other GPS-denied environments. Some receivers have extra features to improve performance in challenging environments. To enhance accuracy, some receivers are dual-frequency rather than single-frequency. One benefit of a dual-

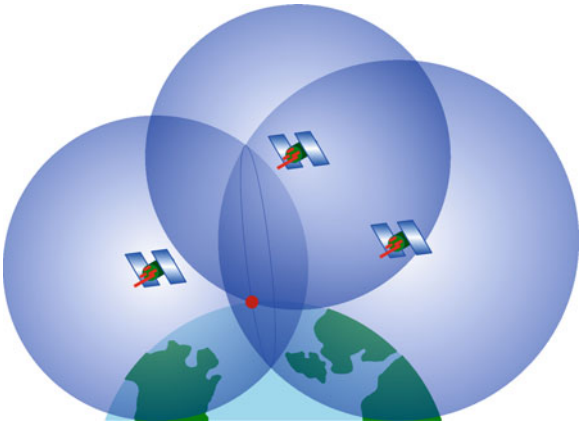


Fig. 1.2 The concept of position fixing by trilateration using signals from three satellites. The user's position is indicated by the *red/dark dot*

frequency receiver is that it can estimate the ionospheric delay, which is frequency dependent, and eliminate this error in the measured pseudo-ranges. Furthermore, some GPS receivers can employ GLONASS signals as an augmentation system. One of the extra features in receivers is multipath mitigation. Several methods of multipath estimation and mitigation have been implemented. A classification of multipath mitigation techniques can be found in Yang et al. 2004.

Another feature in GPS receivers is their sensitivity, which is defined as the minimum received signal strength necessary in order for it to provide a solution. The GPS signal is typically buried in thermal noise, and the minimum signal level at the surface of the Earth is −130 dBm. While the signal structure itself provides a means to recover the navigation data from the noise, these inherent properties fail under dense foliage or in urban canyons (Sudhir et al. 2001). GPS receivers which exhibit a sensitivity of −150 to −160 dBm are now available. Using a high-sensitivity receiver leads to faster time-to-first-fix (TTFF) and better tracking, and can provide acceptable GPS functionality to previously impossible locations such as inner city urban canyons, as well as regions populated by dense foliage and in vehicles with ultraviolet-coated or thermal windshields, although they may not be able to maintain the same levels of accuracy in these environments.

Despite the abovementioned features which enhance the performance of a GPS receiver, some environments still suffer from severely degraded performance (like in deep urban canyons encountered in downtown scenarios) or complete blockage (like tunnels and green tunnels). To provide a more accurate and uninterrupted navigation solution, GPS must be integrated with other systems which can exhibit complementary characteristics (Skog and Handel 2009). Other systems that can be integrated with GPS (some complementary and offering an uninterrupted solution whilst others merely enhance the performance) are

1. GPS Augmentation Systems
   (a) Other GNSS such as the Russian GLONASS or the European Galileo.
   (b) Space-based augmentation systems (SBAS).

(c) Ground-based augmentation systems (GBAS).

2. Local wireless-based positioning systems which use cellular base stations, Wi-Fi access points and TV towers.
3. Motion sensors such as inertial sensors, odometers and a compass.
4. Aiding sensors such as cameras and light detection and ranging (LIDAR).
5. Digital maps.

1.3 Integration of GPS with Other Systems

Accuracy, integrity and availability of GPS can be increased by integrating it with other systems, sensors and aiding devices. Some of the important augmentation techniques are summarized below.

1.3.1 GPS Augmentation Systems

1.3.1.1 Other GNSS

For GPS receivers capable of utilizing signals from other satellite systems such as GLONASS and Galileo, these extra systems can increase the number of satellite signals in view. However, any environment that highly degrades GPS will also tend to impair or block these other systems.

1.3.1.2 Space-based Augmentation System

SBAS is a system that supports wide-area or regional augmentation by the use of additional messages from geostationary satellites. These systems are commonly composed of multiple ground stations, located at sites that have been accurately surveyed. These stations take measurements of one or more of the GPS satellites' signals, and of environmental factors that could impair the signal received by the end users. After these measurements are processed, information messages in the form of a GPS-like signal are generated and sent up to one or more geostationary satellites for broadcasting to the users. The prime examples of SBAS are the wide-area augmentation system (WAAS) operated by the United States' federal aviation administration (FAA) and its European counterpart, the European geostationary navigation overlay service (EGNOS) which is operated by the European space agency (ESA). GPS augmented with WAAS can consistently achieve accuracies of 1–2 m in horizontal positioning and 2–3 m in the vertical direction (FAA 2012b). Of course, all these space-based systems will be negatively affected or blocked in highly degraded GPS environments.

1.3.1.3 Ground-Based Augmentation System

A GBAS is a system which provides augmentation by the use of terrestrial radio messages (FAA 2012a). As with satellite-based augmentation systems, ground-based augmentation systems are commonly composed of one or more accurately surveyed ground stations. The stations take measurements of the GPS signals, and then one or more radio transmitters broadcast the information directly to the end user. Examples of the GBAS are United States' local area augmentation system (LAAS) mainly used around airports to improve aircraft safety during approaches and landings, and the Differential GPS (DGPS) which provides a level of accuracy that can approach 1 m depending on whether the measurement is code-based or carrier-phase-based and whether the processing is real-time or post-processing. LAAS provides an accuracy of better than 1 m, but is mainly used in airports and out to a range of 50 km (Stanford 2012). Another example of ground stations serving as additional satellites is what are known as pseudolites (pseudo-satellites). By locating these at favorable sites, the accuracy and continuity of the navigation solution available to a GNSS receiver can be enhanced in GPS-denied environments. However, pseudolites can only solve the coverage problem locally, the receiver must be designed to handle the additional signals, and this technique requires an additional infrastructure.

1.3.2 Local Wireless-Based Positioning Systems

The location of a mobile receiver can also be determined by using the wireless communication signals transmitted from different systems. Since the focus here is on automobile navigation, the discussion is limited to those systems which can be effective in such an application, namely cellular networks and TV towers. Other positioning systems in this category that rely on WLAN, UWB or radio-frequency identification (RFID) are more suitable for indoor navigation. Positioning based on cellular networks or TV towers can be used in a standalone manner or aid GPS to enhance its performance.

1.3.2.1 Using Cellular Networks

A number of systems have used a global system of mobile/code division multiple access (GSM/CDMA) mobile cellular networks to estimate the location of outdoor mobile clients. The two major categories of location systems are network-based and handset-based. Network-based systems employ technologies which determine the position of a handset solely from measurements taken at cellular base stations, imposing no requirements on the unit beyond its normal communications function. In a handset-based system, the location measurements are made in the handset and special software and/or hardware is required. Hybrid systems can also be utilized,

where both the network and the handset are modified to perform the positioning function (Bensky Dec, 2007). Cellular phone carriers use a number of systems to provide location services based on different techniques, including

1. Cell identification (Cell-ID).
2. Time of arrival (TOA).
3. Time difference of arrival (TDOA) with modes for enhanced observed time difference (E-OTD), observed time difference of arrival (OTDOA), and uplink time difference of arrival (U-TDOA).
4. Angle of arrival (AOA).
5. Assisted-GPS (A-GPS).

The first four techniques are used for positioning mobile clients by only a few cellular phone carriers. In general they cannot achieve better positioning accuracy than GPS. They can readily achieve accuracies of 50–200 m, but hardly ever less than 10 m (Hui et al. Nov, 2007). However, positions calculated by these systems can be used to augment GPS. One limitation of such integrated systems is that positioning using cellular networks is unable to achieve a better accuracy than GPS and is usually much worse. On the other hand, the accuracy of cellular-based positioning is higher in densely covered areas compared to less densely covered regions. Another drawback of this integrated positioning solution is that service charges are imposed for accessing a cellular network. It also lacks generality and universality because it depends on the specifications in the operating environment, such as network coverage.

A-GPS uses the GPS infrastructure to determine a user's location. Signals to aid GPS acquisition are transmitted over the cellular phone network to a receiver contained within the handset. Generally, differential corrections are not sent to the user, the information transmitted is to assist the phone in rapidly acquiring the satellite signals. Assistance from the cell networks can also extend to performing some calculations remotely and sending the results to the phone (Jacobson 2007). Unlike standalone GPS, systems and devices that are A-GPS-only are tied to good signals from subscriber cellular phone networks and this substantially limits their area of operation. Of course, this technique has no positioning capabilities in GPS-denied environments. Its benefit is that the signals to aid GPS acquisition are sent by the cell phone network.

1.3.2.2 Using TV Towers

Some companies use TV signals to get a location fix in places where GPS signals are impaired or blocked, such as dense urban areas and indoors. To determine its position, the receiver measures the pseudo-ranges from TV signals broadcast by three or more different towers which contain synchronization information. Using precision timing, it is possible to work out how far a TV signal travels before it is picked up by a suitably equipped receiver. The company's location server receives the timing and stability of TV signals from monitoring units in the region and then

forwards that as aiding information. The measurements are then compared against the data collected by proprietary listening stations, and the position of the device is calculated (Kolodziej and Hjelm 2006). There are many advantages to using TV signals, such as

1. Power: Most of the time TV towers are tens of kilometers from the device to be positioned, and the synchronization codes in the TV signal are of the order of ten thousand times stronger than GPS signals.
2. Frequency: TV signals are at a much lower frequency (50–700 MHz) than the GPS, enabling them to penetrate buildings.
3. Bandwidth: In general, the broader the bandwidth of the signal the more accurately it can resolve the multipath or reflected signals that characterize urban and indoor environments.

However, the drawbacks of using the TV-positioning systems are

1. The technology depends on the installation of several proprietary nearby monitoring units as well as location servers, and this limits the positioning solution environment to certain areas.
2. Decoding the synchronization information from the TV signals varies with the TV standard employed, which can be NTSC, PAL, SECAM, DVB-T or ISDB amongst others.

1.3.3 Vehicle Motion Sensors

1.3.3.1 Inertial Sensors

Inertial sensors are gyroscopes and accelerometers which respectively measure the rotation rates and the specific forces from which acceleration can be obtained. An inertial measurement unit (IMU) consists of a triad of gyroscopes and a triad of accelerometers. A navigation system using an IMU and a navigation processor is called an inertial navigation system (INS). An INS has the advantage that it is self-contained; i.e. it does not need external references. Another advantage is that such sensors have high sampling rates. With knowledge of the initial navigation states of position, velocity and attitude, the system can calculate the three-dimensional position, velocity and attitude of a moving platform. Measurements are integrated twice for the accelerometers and once for the gyroscopes in order to yield position and attitude. However, these calculated navigation states drift with time due to the mathematical integrations required to obtain the position and attitude. The sensor errors (such as biases, scale factors and noise) are also integrated, and can result in unbounded errors. Consequently, inertial sensors alone are unsuitable for accurate positioning over an extended period of time. As will be discussed later, the pros and cons of an INS show that its characteristics are complementary with those of GPS, and these can be combined so that each mitigates the other's problems.

1.3.3.2 Electronic Compass

The heading is the most significant of the navigation parameters because of its influence on accumulated dead reckoning errors. Hence sensors to measure the absolute heading (e.g. an electronic compass) can be an essential component of an integrated navigation systems. On the other hand, magnetic compass usage has the disadvantage that the Earth's magnetic field is often distorted near power lines or metal structures.

1.3.3.3 Odometry

As mentioned earlier, odometry data is extracted by sensors which measure the rotation of the wheel axes and the steer axes. Curvilinear distance is obtained from the wheel rotation data. The advantages of this technique are: (1) short term accuracy, (2) low price, and (3) high sampling rates. The main disadvantages are that any small constant error increases without bound, and orientation errors cause large position errors that increase with the distance traveled. The pros and cons of odometry indicate that its characteristics are complementary with those of GPS, so it is a candidate for integration with GPS.

1.3.4 Other Aiding Sensors

1.3.4.1 Camera

Cameras and vision systems can be employed in navigation systems to perform a variety of tasks in different moving platforms. For outdoor vehicular navigation, camera images can be used to detect road edges, lanes and their transitions, and also road intersections (DeSouza and Kak 2002). Such tasks on their own cannot achieve absolute positioning for the vehicle, but can assist an existing positioning solution. The drawbacks of these systems are sensitivity to illumination changes throughout the day, the presence of clutters in imagery, the presence of shadows cast by trees and other artifacts, and the presence of water or snow on the ground. A vehicular positioning system must function properly in all conditions but these problems mean that the solutions derived from vision systems can be unreliable in certain conditions, which is unacceptable. As another use of cameras, stereo vision can assist in the detection and avoidance of obstacles but it too suffers from such limitations. Although vision can provide positioning by model matching, and is commonly used for indoor mobile robot navigation, it requires the environment to be known beforehand and the vehicle to be provided with a model or an image database, which means that this technique cannot be applied to general vehicular navigation.

1.3.4.2 Sonar

Sonar can be used for indoor localization of mobile robots using model matching techniques, but is not suitable for outdoor vehicle navigation. Another use of such sensors is for the detection and avoidance of obstacles.

1.3.4.3 LIDAR

Light detection and ranging (LIDAR) can be used in several tasks related to navigation, in particular for positioning using model matching in applications such as indoor mobile robot localization. The limitation of model matching is that it is not a general solution for vehicular navigation, it is viable only in environments for which a model has been generated. Another way of using LIDAR in outdoor positioning is to augment an integrated GPS and INS. When GPS is available the LIDAR will detect buildings, and when GPS is degraded or blocked, as in an urban canyon, the LIDAR will measure the relative positions of buildings to aid the INS. But LIDAR is most often used in mapping, in obstacle detection and avoidance, and for three-dimensional motion capture. Another application is in traffic speed law enforcement, as an alternative to radar guns for vehicle speed measurement. Airborne LIDAR can be used for aerial vehicle navigation using terrain elevation databases, and thus augment other positioning systems. One drawback of LIDAR in commercial vehicular applications is its high cost. Furthermore, working with LIDAR data can be computationally expensive for real-time applications.

1.3.5 Digital Maps

A digital map is not only used to assist the driver in relating the vehicle's position information to a physical location, it can also supply extra data to the positioning solution. Because the location and trajectory of a vehicle is usually restricted by the road network, a digital map of this network can be used to impose constraints on the positioning solution (Skog and Handel 2009) using a process referred to as map matching. Although digital maps cannot be used alone, they can enhance an existing solution.

1.4 Inertial Navigation

Of the many navigation techniques available, inertial navigation and satellite navigation are the most commonly used. As described earlier, GPS is a *reference based* system whereas an inertial navigation system (INS) is a form of dead reckoning that uses three accelerometers and three gyroscopes along with a

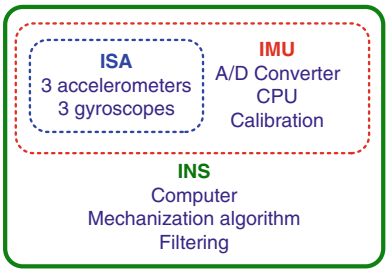


Fig. 1.3 An illustration of the various terminologies used in inertial navigation system and their interrelationship

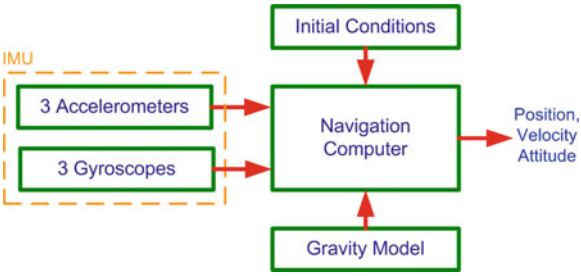


Fig. 1.4 A high level block diagram of an inertial navigation system (INS)

processor. The accelerometers and gyroscopes constitute the inertial sensor assembly (ISA) that is housed along with related electronics in a unit called the inertial measurement unit (IMU). The relationship of the ISA, IMU and INS is summarized in Fig. 1.3.

By measuring angular rates, gyroscopes provide attitude information in three dimensions. Accelerometers measure the specific forces (accelerations) that can be converted to velocity and position by a navigation computer using the process of integration. Starting with initial values of position, velocity and attitude, together with a gravity model, the computer operates on the IMU's outputs to furnish the current position, velocity and attitude of the host vehicle, as shown in Fig. 1.4. The gravity model assists in obtaining acceleration from specific forces while the initial conditions are used for integrating the acceleration and velocity. The details of inertial navigation are explored in Chap. 4.

1.5 Integrated INS/GPS Navigation

Due to their disparate principles of operation, dead reckoning and position fixing systems possess complementary characteristics and, if suitably combined, they can attain better performance than either can achieve in a standalone mode. Table 1.1 summarizes the main characteristic of the INS (dead reckoning system) and GPS (position fixing system).

Table 1.1 Summary of important characteristics of INS and GPS

Characteristics	INS	GPS
Accuracy of navigational solution	Good short term accuracy which deteriorates with time	Good long term accuracy but noisy in short term
Initial conditions	Required	Not required
Attitude information	Available	Typically not available[a]
Sensitive to gravity	Yes	No
Self-contained	Yes	No
Jamming immunity	Yes	No
Output data rate	High	Low

[a] With multiple antennae, some GPS receivers can render attitude information as well

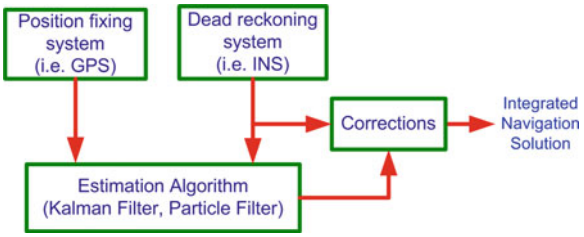


Fig. 1.5 A high level block diagram of a typical integrated navigation system

Based on their complementary properties, these systems are usually combined using some estimation technique (traditionally based on the Kalman filter) in order to obtain (Solimeno 2007)

1. Higher position and velocity accuracy.
2. Precise attitude information.
3. Higher data output rate.
4. Navigational solution during GPS signal blockages.

Figure 1.5 shows the basic scheme of a typical integrated navigation system.

1.6 Types of INS/GPS Integration

Different forms of INS/GPS integration architectures have been proposed to attain maximum advantage depending upon the type of use and the degree of simplicity versus robustness. The three main integration architectures are

1. Loosely coupled.
2. Tightly coupled.
3. Ultra-tightly or deeply coupled.

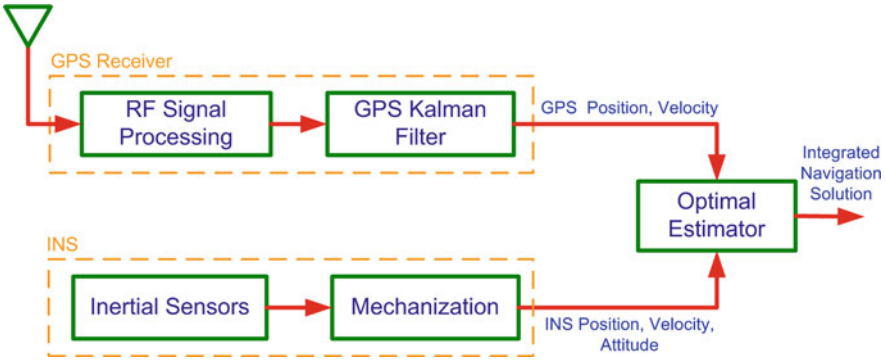


Fig. 1.6 A basic block diagram of a loosely coupled integration of INS/GPS

1.6.1 Loosely Coupled INS/GPS Integration

In this architecture, also known as a *decentralized* integration, the GPS and INS function independently and provide separate solutions for position, velocity and attitude.[1] To get the best of both solutions, this information is fused together by an optimal estimator to obtain a third and much improved solution. This arrangement is shown in Fig. 1.6.

1.6.2 Tightly Coupled INS/GPS Integration

In this architecture, also known as *centralized* integration, the GPS and INS work as basic sensors and their raw outputs of pseudo-range and pseudo-range rate from the GPS and accelerations and rotation rates from the INS are blended by a single estimator in order to achieve a synergistic solution (Yang 2007). This arrangement is shown in Fig. 1.7.

As discussed in Chap. 8, GPS/INS integration can be achieved in a variety of ways, each of which has its advantages and limitations.

1.6.3 Ultra-Tightly or Deeply Coupled Integration

Deeply coupled integration provides an increased symbiosis between the INS and GPS because the integration is at tracking loop level. The main advantage of this is that the dynamics of the host vehicle are estimated and compensated in the GPS

[1] Not all GPS receivers can provide attitude information, which requires two antennae.

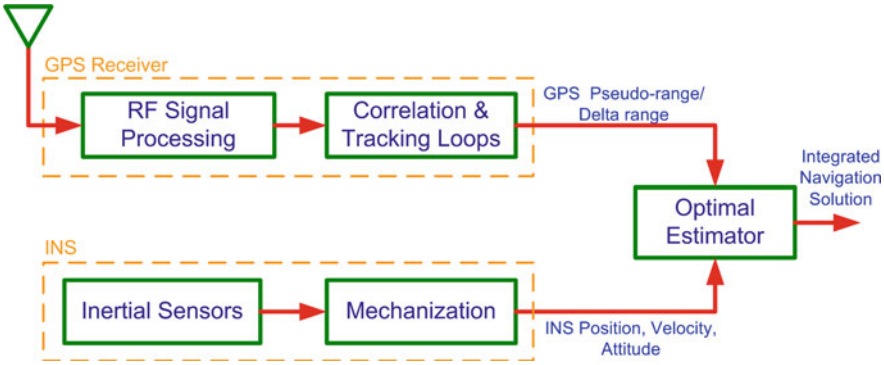


Fig. 1.7 A basic block diagram of a tightly coupled integration of INS/GPS

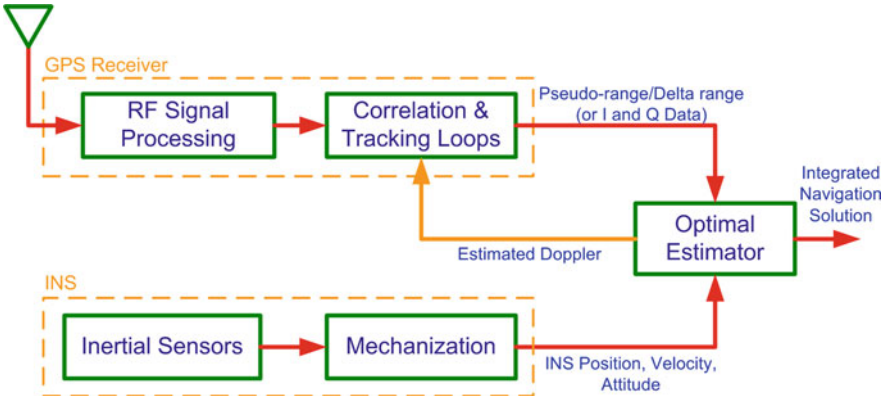


Fig. 1.8 A basic block diagram of an ultra-tightly coupled integration of INS/GPS

tracking loops by using Doppler information. Various configurations of ultra-tight integration exist, and Fig. 1.8 shows a basic one. The estimator combines either the pseudo-ranges/Doppler or I (in-phase) and Q (quadrature) measurements from the GPS with the INS navigation parameters so as to render the estimated Doppler (Alban et al. 2003). The estimated Doppler is used to remove the dynamics from the GPS signal entering the tracking loops, thereby reducing the carrier tracking loop bandwidth. Although this integration is more complex and requires access to the GPS hardware, it can improve the quality of the raw measurements and also the anti-jamming performance of the signals (Cox Jr 1978).

1.7 INS/GPS Fusion Algorithm

There are several algorithms for optimal fusion of GPS and INS data, the major ones being various forms of Kalman filter (KF), particle filter (PF) and artificial intelligence (AI). Traditionally, Kalman filtering has been the method of choice for fusing navigational information from various sources. It is an optimal recursive algorithm (Maybeck 1979) which processes all of the available measurements, regardless of their precision, to optimally estimate the current value of the state of interest and, importantly, also furnishes the uncertainty of its estimate. It uses the knowledge of a system model derived from the motion equations of the platform, including the inertial readings and a measurement model for the GPS information, a statistical description of system noises, the nature of measurement errors, and rough estimates of initial conditions. Chapter 7 discusses KF in detail.

1.8 Summary of the Chapters

Chapter 2 describes the mathematics of navigation, the coordinate frames used in navigation and the transformations that are used compute a navigation solution in suitable coordinates.

Chapter 3 deals with GPS, detailing its various segments, its signal structure, and the process of position and velocity estimation from pseudo-range and delta-range measurements. It also discusses various errors that can occur in processing GPS data. It concludes with details of how to calculate the position and velocity of a satellite from ephemeris data, including the atmospheric corrections for a single-frequency GPS receiver.

Chapter 4 deals with inertial sensors, their classification and associated errors, and their calibration, initialization and alignment procedures, then describes the INS mechanization process and its equations.

Chapter 5 gives some basics mathematics of motion and describes the details of the mechanization process in various reference frames, giving a step by step computation of navigation parameters in most common navigational frames.

Chapter 6 models INS errors, along with the linearization of INS equations for various navigational reference frames.

Chapter 7 discusses the process and measurement models for KF, the filtering algorithm, non-linear filtering techniques and how to model system disturbances.

Chapter 8 pulls all of these ideas together with a detailed account of INS/GPS integration and associated nuances and subtleties. It describes types of integration, along with their feedback schemes. It derives dynamic INS error model equations as well as inertial sensor error model equations, and discusses the implementation of linearized KF for loosely and tightly coupled INS/GPS fusion.

Chapter 9 discusses a possible trend for wheeled vehicles that cuts down the number of low cost inertial sensors by using speed readings and introduces various

constraints to eliminate the errors associated with the six sensors in a full IMU and achieve similar, if not better, accuracy.

Chapter 10 deals with the experimental results of INS/GPS fusion. The results of integrating a full IMU (with all six sensors) and the reduced system (consisting of fewer sensors) are analyzed based on real road test trajectories.

Each chapter concludes with a reference list where the interested reader can find further information on some of the topics that could not be covered in detail in this book owing to limitations of scope.

References

Alban S, Akos DM, Rock SM, Gebre-Egziabher D (2003) Performance analysis and architectures for INS-aided GPS tracking loops. In: Proceedings of ION National Technical Meeting *(NTM)*, Anaheim CA, 11–24 Jan, U.S. Institute of Navigation, Fairfax pp 611–622

Beidou-1 (2012) BeiDou 1 satellite navigation experimental system. www.sinodefence.com/satellites/beidou1.asp. Accessed 16 March 2012

Beidou-2 (2012) Compass (Beidou 2) satellite navigation system. www.sinodefence.com/space/satellite/beidou2.asp. Accessed 16 March 2012

Bensky A (2007) Wireless positioning technologies and applications-technology and applications. Artech House Publishers, London

Borenstein J, Everett H, Feng L, Wehe D (1997) Mobile robot positioning: sensors and techniques. J Rob Syst 14(4):231–249

Cox DB Jr (1978) Integration of GPS with inertial navigation systems. Navigation: J Inst Navig 25(2):206–245

DeSouza GN, Kak AC (2002) Vision for mobile robot navigation: a survey. IEEE Trans Pattern Anal Mach Intell 24(2):237–267

El-Rabbany A (2002) Introduction to GPS: the global positioning system, mobile communication series. Artech House Inc., Norwood

ESA (2012) What is Galileo? european space agency www.esa.int/esaNA/galileo.html. Accessed 12 Feb 2012

Esmat B (Sep 2007) Introduction to modern navigation systems. World Scientific Publishing, NJ

FAA (2012a) Navigation services—ground based augmentation system. www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/gnss/laas/. Accessed Mar 16, 2012

FAA (2012b) Wide area augmentation system (WAAS). www.faa.gov/air_traffic/technology/waas/. Accessed Mar 18, 2012 2012

FSA (2012) GLONASS constellation status. www.glonass-center.ru/en/GLONASS/. Accessed Mar 18, 2012

Groves PD (Dec 2007) Principles of GNSS, inertial, and multi-sensor integrated navigation systems. Artech House

Hui L, Darabi H, Banerjee P, Jing L (2007) Survey of wireless indoor positioning techniques and systems. IEEE Trans syst, Man, Cybern-Part C 37(6):1067–1080 (applications and reviews)

Jacobson L (2007) GNSS markets and applications: GNSS technology and applications series. Artech House, Norwood

Kolodziej KW, Hjelm J (2006) Local positioning systems LBS applications and services. CRC/Taylor and Francis, Boca Raton

Maybeck PS (1979) Stochastic models, estimation, and control, vol 1. mathematics in science and engineering, vol 141. Academic, New York

Misra P, Enge P (Dec 2001) Global positioning system: signals, measurements and performance. Lincoln, Mass., Ganga-Jamuna Press

Skog I, Handel P (2009) In-car positioning and navigation technologies: a survey. IEEE Trans Intell Transp Syst 10(1):4–21

Solimeno A (2007) Low-cost INS/GPS data fusion with extended kalman filter for airborne applications. M.Sc. thesis, Technical University of Lisbon, Lisbon

Space-Based Positioning, Navigation and Timing (2008). http://pnt.gov/. Accessed Dec 4 2008

Stanford (2012) Local aread augmentation system. http://waas.stanford.edu/research/laas.htm. Accessed Mar 18 2012

Sudhir NS, Vimala C, Ray JK (2001) Receiver sensitivity analysis and results. In: Proceedings of the 14th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2001), Salt Lake City, Utah, USA, Sept 11–14 2001. Institute of Navigation

Yang Y (2007) Tightly coupled MEMS INS/GPS integration with INS aided receiver tracking loops. University of Calgary, Calgary

Yang Y, Hatch RR, Sharpe RT (2004) GPS multipath mitigation in measurement domain and its applications for high accuracy navigation. In: Proceedings of the 17th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2004), Long Beach, CA, USA. Institute of Navigation, pp 1124–1130

Chapter 2
Basic Navigational Mathematics, Reference Frames and the Earth's Geometry

Navigation algorithms involve various coordinate frames and the transformation of coordinates between them. For example, inertial sensors measure motion with respect to an inertial frame which is resolved in the host platform's body frame. This information is further transformed to a navigation frame. A GPS receiver initially estimates the position and velocity of the satellite in an inertial orbital frame. Since the user wants the navigational information with respect to the Earth, the satellite's position and velocity are transformed to an appropriate Earth-fixed frame. Since measured quantities are required to be transformed between various reference frames during the solution of navigation equations, it is important to know about the reference frames and the transformation of coordinates between them. But first we will review some of the basic mathematical techniques.

2.1 Basic Navigation Mathematical Techniques

This section will review some of the basic mathematical techniques encountered in navigational computations and derivations. However, the reader is referred to (Chatfield 1997; Rogers 2007 and Farrell 2008) for advanced mathematics and derivations. This section will also introduce the various notations used later in the book.

2.1.1 Vector Notation

In this text, a vector is depicted in bold lowercase letters with a superscript that indicates the coordinate frame in which the components of the vector are given. The vector components do not appear in bold, but they retain the superscript. For example, the three-dimensional vector r for a point in an arbitrary frame k is depicted as

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_2,

$$\mathbf{r}^k = \begin{bmatrix} x^k \\ y^k \\ z^k \end{bmatrix} \tag{2.1}$$

In this notation, the superscript k represents the k-frame, and the elements (x^k, y^k, z^k) denote the coordinate components in the k-frame. For simplicity, the superscript is omitted from the elements of the vector where the frame is obvious from the context.

2.1.2 Vector Coordinate Transformation

Vector transformation from one reference frame to another is frequently needed in inertial navigation computations. This is achieved by a transformation matrix. A matrix is represented by a capital letter which is not written in bold. A vector of any coordinate frame can be represented into any other frame by making a suitable transformation. The transformation of a general k-frame vector $\mathbf{r}^k$ into frame m is given as

$$\mathbf{r}^m = R_k^m \mathbf{r}^k \tag{2.2}$$

where R_k^m represents the matrix that transforms vector $\mathbf{r}$ from the k-frame to the m-frame. For a valid transformation, the superscript of the vector that is to be transformed must match the subscript of the transformation matrix (in effect they cancel each other during the transformation).

The inverse of a transformation matrix R_k^m describes a transformation from the m-frame to the k-frame

$$\mathbf{r}^k = \left(R_k^m\right)^{-1} \mathbf{r}^m = R_m^k \mathbf{r}^m \tag{2.3}$$

If the two coordinate frames are mutually orthogonal, their transformation matrix will also be orthogonal and its inverse is equivalent to its transpose. As all the computational frames are orthogonal frames of references, the inverse and the transpose of their transformation matrices are equal. Hence for a transformation matrix R_k^m we see that

$$R_k^m = \left(R_m^k\right)^T = \left(R_m^k\right)^{-1} \tag{2.4}$$

A square matrix (like any transformation matrix) is orthogonal if all of its vectors are mutually orthogonal. This means that if

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{2.5}$$

where

$$\mathbf{r}_1 = \begin{bmatrix} r_{11} \\ r_{21} \\ r_{31} \end{bmatrix}, \mathbf{r}_2 = \begin{bmatrix} r_{12} \\ r_{22} \\ r_{33} \end{bmatrix}, \mathbf{r}_3 = \begin{bmatrix} r_{13} \\ r_{23} \\ r_{33} \end{bmatrix} \tag{2.6}$$

then for matrix R to be orthogonal the following should be true

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = 0, \mathbf{r}_1 \cdot \mathbf{r}_3 = 0, \mathbf{r}_2 \cdot \mathbf{r}_3 = 0 \tag{2.7}$$

2.1.3 Angular Velocity Vectors

The angular velocity of the rotation of one computational frame about another is represented by a three component vector $\boldsymbol{\omega}$. The angular velocity of the k-frame relative to the m-frame, as resolved in the p-frame, is represented by ω^p_{mk} as

$$\boldsymbol{\omega}^p_{mk} = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{2.8}$$

where the subscripts of $\boldsymbol{\omega}$ denote the direction of rotation (the k-frame with respect to the m-frame) and the superscripts denote the coordinate frame in which the components of the angular velocities $(\omega_x, \omega_y, \omega_z)$ are given.

The rotation between two coordinate frames can be performed in two steps and expressed as the sum of the rotations between two different coordinate frames, as shown in Eq. (2.9). The rotation of the k-frame with respect to the p-frame can be performed in two steps: firstly a rotation of the m-frame with respect to the p-frame and then a rotation of the k-frame with respect to the m-frame

$$\boldsymbol{\omega}^k_{pk} = \boldsymbol{\omega}^k_{pm} + \boldsymbol{\omega}^k_{mk} \tag{2.9}$$

For the above summation to be valid, the inner indices must be the same (to cancel each other) and the vectors to be added or subtracted must be in the same reference frame (i.e. their superscripts must be the same).

2.1.4 Skew-Symmetric Matrix

The angular rotation between two reference frames can also be expressed by a skew-symmetric matrix instead of a vector. In fact this is sometimes desired in order to change the cross product of two vectors into the simpler case of matrix multiplication. A vector and the corresponding skew-symmetric matrix forms of an angular velocity vector $\boldsymbol{\omega}^p_{mk}$ are denoted as

$$\underbrace{\boldsymbol{\omega}_{mk}^{p} = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}}_{\text{Angular velocity vector}} \Rightarrow \underbrace{\Omega_{mk}^{p} = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}}_{\text{Skew-symmetric form of angular the velocity vector}} \tag{2.10}$$

Similarly, a velocity vector $\mathbf{v}^p$ can be represented in skew-symmetric form V^p as

$$\underbrace{\mathbf{v}^{p} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}}_{\text{Velocity vector}} \Rightarrow \underbrace{\mathrm{V}^{p} = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix}}_{\text{Skew-symmetric form of the velocity vector}} \tag{2.11}$$

Note that the skew-symmetric matrix is denoted by a non-italicized capital letter of the corresponding vector.

2.1.5 Basic Operations with Skew-Symmetric Matrices

Since a vector can be expressed as a corresponding skew-symmetric matrix, the rules of matrix operations can be applied to most vector operations. If **a**, **b** and **c** are three-dimensional vectors with corresponding skew-symmetric matrices A, B and C, then following relationships hold

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T\mathbf{b} = \mathbf{b}^T\mathbf{a} \tag{2.12}$$

$$\mathbf{a} \times \mathbf{b} = \mathrm{A}\mathbf{b} = B^T\mathbf{a} = -\mathrm{B}\mathbf{a} \tag{2.13}$$

$$[\![\mathrm{A}\mathbf{b}]\!] = \mathrm{AB} - \mathrm{BA} \tag{2.14}$$

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{a}^T\mathrm{B}\mathbf{c} \tag{2.15}$$

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathrm{AB}\mathbf{c} \tag{2.16}$$

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = \mathrm{AB}\mathbf{c} - \mathrm{BA}\mathbf{c} \tag{2.17}$$

where $[\![\mathrm{A}\mathbf{b}]\!]$ in Eq. (2.14) depicts the skew-symmetric matrix of vector A**b**.

2.1.6 Angular Velocity Coordinate Transformations

Just like any other vector, the coordinates of an angular velocity vector can be transformed from one frame to another. Hence the transformation of an angular velocity vector $\boldsymbol{\omega}_{mk}$ from the k-frame to the p-frame can be expressed as

$$\boldsymbol{\omega}_{mk}^{p} = R_{k}^{p}\boldsymbol{\omega}_{mk}^{k} \tag{2.18}$$

The equivalent transformation between two skew-symmetric matrices has the special form

$$\Omega_{mk}^{p} = R_{k}^{p}\Omega_{mk}^{k}R_{p}^{k} \tag{2.19}$$

2.1.7 Least Squares Method

The method of least squares is used to solve a set of equations where there are more equations than the unknowns. The solution minimizes the sum of the squares of the residual vector. Suppose we want to estimate a vector $\mathbf{x}$ of n parameters $(x_1, x_2, \ldots, x_n)$ from vector $\mathbf{z}$ of m noisy measurements $(z_1, z_2, \ldots, z_m)$ such that $m > n$. The measurement vector is linearly related to parameter $\mathbf{x}$ with additive error vector ε such that

$$\mathbf{z} = H\mathbf{x} + \varepsilon \tag{2.20}$$

where H is a known matrix of dimension $m \times n$, called the design matrix, and it is of rank n (linearly independent row or columns).

In the method of least square, the sum of the squares of the components of the residual vector $(\mathbf{z} - H\mathbf{x})$ is minimized in estimating vector $\mathbf{x}$, and is denoted by $\hat{\mathbf{x}}$. Hence

$$\text{minimize } \|\mathbf{z} - H\hat{\mathbf{x}}\|^2 = (\mathbf{z} - H\hat{\mathbf{x}})_{1\times m}^{T}(\mathbf{z} - H\hat{\mathbf{x}})_{m\times 1} \tag{2.21}$$

This minimization is achieved by differentiating the above equation with respect to $\hat{\mathbf{x}}$ and setting it to zero. Expanding the above equation gives

$$\|\mathbf{z} - H\hat{\mathbf{x}}\|^2 = \mathbf{z}^T\mathbf{z} - \mathbf{z}^T H\hat{\mathbf{x}} - \hat{\mathbf{x}}^T H^T\mathbf{z} + \hat{\mathbf{x}}^T H^T H\hat{\mathbf{x}} \tag{2.22}$$

Using the following relationships

$$\frac{\partial(\mathbf{x}^T\mathbf{a})}{\partial\mathbf{x}} = \frac{\partial(\mathbf{a}^T\mathbf{x})}{\partial\mathbf{x}} = \mathbf{a}^T \tag{2.23}$$

and

$$\frac{\partial(\mathbf{x}^T A\mathbf{x})}{\partial\mathbf{x}} = (A\mathbf{x})^T + \mathbf{x}^T A \tag{2.24}$$

the derivative of the scalar quantity represented by Eq. (2.22) is obtained as follows

$$\begin{aligned}
\frac{\partial\left(\|\mathbf{z}-H\hat{\mathbf{x}}\|^2\right)}{\partial\hat{\mathbf{x}}} &= 0-\mathbf{z}^TH-(H^T\mathbf{z})^T+(H^TH\hat{\mathbf{x}})^T+\hat{\mathbf{x}}^TH^TH \\
\frac{\partial\left(\|\mathbf{z}-H\hat{\mathbf{x}}\|^2\right)}{\partial\hat{\mathbf{x}}} &= -\mathbf{z}^TH-\mathbf{z}^TH+\hat{\mathbf{x}}^TH^TH+\hat{\mathbf{x}}^TH^TH \\
\frac{\partial\left(\|\mathbf{z}-H\hat{\mathbf{x}}\|^2\right)}{\partial\hat{\mathbf{x}}} &= -2(\mathbf{z}^TH+\hat{\mathbf{x}}^TH^TH)
\end{aligned} \tag{2.25}$$

To obtain the maximum, the derivative is set to zero and solved for $\hat{\mathbf{x}}$

$$-2(\mathbf{z}^TH-\hat{\mathbf{x}}^TH^TH)=0 \tag{2.26}$$

$$\begin{aligned}
\hat{\mathbf{x}}^TH^TH &= \mathbf{z}^TH \\
(\hat{\mathbf{x}}^TH^TH)^T &= (\mathbf{z}^TH)^T \\
H^TH\hat{\mathbf{x}} &= H^T\mathbf{z}
\end{aligned} \tag{2.27}$$

and finally

$$\hat{\mathbf{x}}=(H^TH)^{-1}H^T\mathbf{z} \tag{2.28}$$

We can confirm that the above value of $\hat{\mathbf{x}}$ produces the minimum value of the cost function (2.22) by differentiating Eq. (2.25) once more that results in $2H^TH$ which is positive definite.

2.1.8 Linearization of Non-Linear Equations

The non-linear differential equations of navigation must be linearized in order to be usable by linear estimation methods such as Kalman filtering. The non-linear system is transformed to a linear system whose states are the deviations from the nominal value of the non-linear system. This provides the estimates of the errors in the states which are added to the estimated state.

Suppose we have a non-linear differential equation

$$\dot{\mathbf{x}}=f(\mathbf{x},t) \tag{2.29}$$

and that we know the nominal solution to this equation is $\tilde{\mathbf{x}}$ and we let $\delta\mathbf{x}$ be the error in the nominal solution, then the new estimated value can be written as

$$\mathbf{x}=\tilde{\mathbf{x}}+\delta\mathbf{x} \tag{2.30}$$

The time derivative of the above equation provides

$$\dot{\mathbf{x}}=\dot{\tilde{\mathbf{x}}}+\delta\dot{\mathbf{x}} \tag{2.31}$$

Substituting the above value of $\dot{\mathbf{x}}$ in the original Eq. (2.29) gives

$$\dot{\tilde{\mathbf{x}}} + \delta\dot{\mathbf{x}} = f(\tilde{\mathbf{x}} + \delta\mathbf{x}, t) \tag{2.32}$$

Applying the Taylor series expansion to the right-hand side about the nominal value $\tilde{\mathbf{x}}$ yields

$$f(\tilde{\mathbf{x}} + \delta\mathbf{x}, t) = f(\tilde{\mathbf{x}}, t) + \left.\frac{\partial f(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\tilde{\mathbf{x}}} \delta\mathbf{x} + HOT \tag{2.33}$$

where the *HOT* refers to the higher order terms that have not been considered. Substituting Eq. (2.32) for the left-hand side gives

$$\dot{\tilde{\mathbf{x}}} + \delta\dot{\mathbf{x}} \approx f(\tilde{\mathbf{x}}, t) + \left.\frac{\partial f(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\tilde{\mathbf{x}}} \delta\mathbf{x} \tag{2.34}$$

and since $\tilde{\mathbf{x}}$ also satisfies Eq. (2.29)

$$\dot{\tilde{\mathbf{x}}} = f(\tilde{\mathbf{x}}, t) \tag{2.35}$$

substituting this in Eq. (2.34) gives

$$\dot{\tilde{\mathbf{x}}} + \delta\dot{\mathbf{x}} \approx \dot{\tilde{\mathbf{x}}} + \left.\frac{\partial f(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\tilde{\mathbf{x}}} \delta\mathbf{x} \tag{2.36}$$

The linear differential equations whose states are the errors in the original states is give as

$$\delta\dot{\mathbf{x}} \approx \left.\frac{\partial f(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\tilde{\mathbf{x}}} \delta\mathbf{x} \tag{2.37}$$

After solving this, we get the estimated errors that are added to the estimated state in order to get the new estimate of the state.

2.2 Coordinate Frames

Coordinate frames are used to express the position of a point in relation to some reference. Some useful coordinate frames relevant to navigation and their mutual transformations are discussed next.

2.2.1 Earth-Centered Inertial Frame

An inertial frame is defined to be either stationary in space or moving at constant velocity (i.e. no acceleration). All inertial sensors produce measurements relative to an inertial frame resolved along the instrument's sensitive axis. Furthermore,

we require an inertial frame for the calculation of a satellite's position and velocity in its orbit around the Earth. The frame of choice for near-Earth environments is the Earth-centered inertial (ECI) frame. This is shown in Fig. 2.1 and defined[1] by (Grewal et al. 2007; Farrell 1998) as

a. The origin is at the center of mass of the Earth.
b. The z-axis is along axis of the Earth's rotation through the conventional terrestrial pole (CTP).
c. The x-axis is in the equatorial plane pointing towards the vernal equinox.[2]
d. The y-axis completes a right-handed system.

In Fig. 2.1, the axes of the ECI frame are depicted with superscript i as x^i, y^i, z^i, and in this book the ECI frame will be referred to as the i-frame.

2.2.2 Earth-Centered Earth-Fixed Frame

This frame is similar to the i-frame because it shares the same origin and z-axis as the i-frame, but it rotates along with the Earth (hence the name *Earth-fixed*). It is depicted in Fig. 2.1 along with the i-frame and can be defined as

a. The origin is at the center of mass of the Earth.
b. The z-axis is through the CTP.
c. The x-axis passes through the intersection of the equatorial plane and the reference meridian (i.e. the Greenwich meridian).
d. The y-axis completes the right-hand coordinate system in the equatorial plane.

In Fig. 2.1 the axes of the Earth-Centered Earth-Fixed Frame (ECEF) are shown as X^e, Y^e, Z^e and $(t - t_0)$ represents the elapsed time since reference epoch t_0. The term ω_{ie}^e represents the Earth's rotation rate with respect to the inertial frame resolved in the ECEF frame. In this book the ECEF frame will be referred to as the e-frame.

2.2.3 Local-Level Frame

A local-level frame (LLF) serves to represent a vehicle's attitude and velocity when on or near the surface of the Earth. This frame is also known as the local geodetic or navigation frame. A commonly used LLF is defined as follows

[1] Strictly speaking this definition does not satisfy the requirement defined earlier for an inertial frame because, in accordance with Kepler's second law of planetary motion, the Earth does not orbit around the sun at a fixed speed; however, for short periods of time it is satisfactory.

[2] The vernal equinox is the direction of intersection of the equatorial plane of the Earth with the ecliptic (the plane of Earth's orbit around the sun).

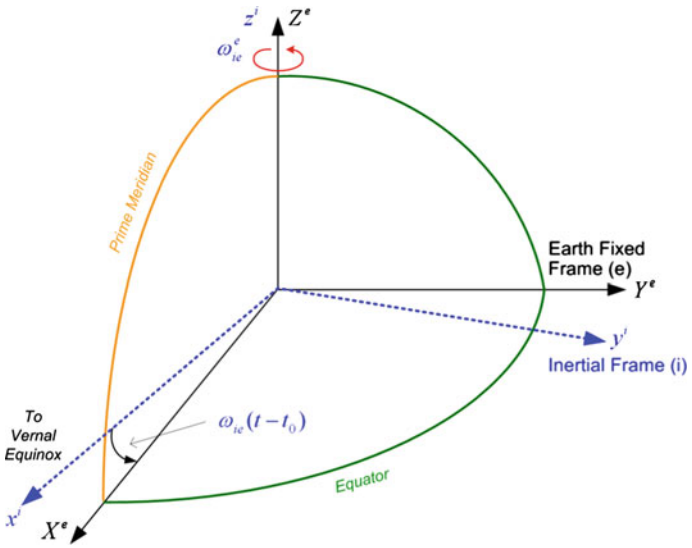


Fig. 2.1 An illustration of the ECI and ECEF coordinate frames

a. The origin coincides with the center of the sensor frame (origin of inertial sensor triad).
b. The y-axis points to true north.
c. The x-axis points to east.
d. The z-axis completes the right-handed coordinate systems by pointing up, perpendicular to reference ellipsoid.

This frame is referred to as ENU since its axes are aligned with the east, north and up directions. This frame is shown in Fig. 2.2. There is another commonly used LLF that differs from the ENU only in that the z axis completes a left-handed coordinate system and therefore points downwards, perpendicular to the reference ellipsoid. This is therefore known as the NED (north, east and down) frame. This book will use the ENU convention, and the LLF frame will be referred to as the l-frame.

2.2.4 Wander Frame

In the l-frame the y-axis always points towards true north, so higher rotation rates about the z-axis are required in order to maintain the orientation of the l-frame in the polar regions (higher latitudes) than near the equator (lower latitudes). As is apparent in Fig. 2.3b, the l-frame must rotate at higher rates to maintain its orientation when moving towards the pole, reaching its maximum when it crosses the north pole. This rate can even become infinite (a singularity condition) if the l-frame passes directly over the pole. The wander frame avoids higher rotation rates and singularity problems. Instead of always pointing northward, this rotates

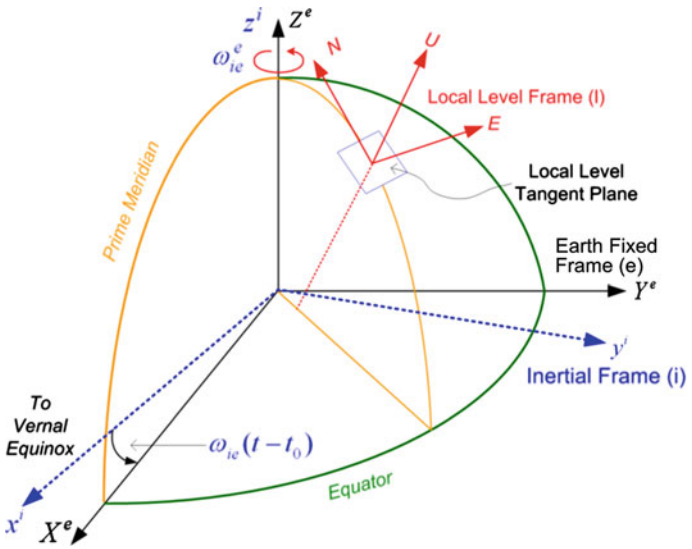


Fig. 2.2 The local-level ENU reference frame in relation to the ECI and ECEF frames

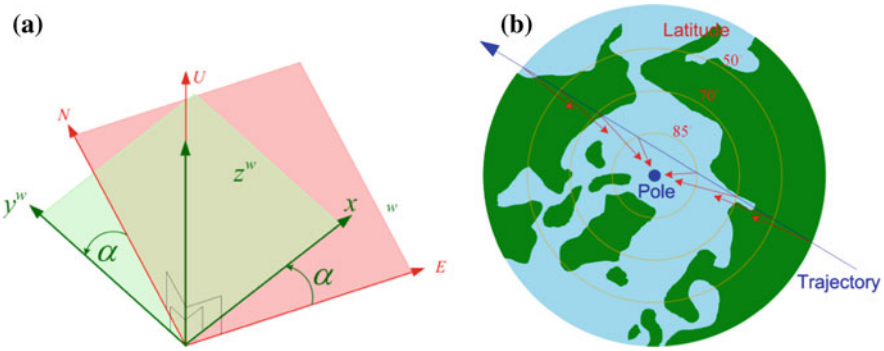


Fig. 2.3 **a** The wander frame shown with respect to the local-level frame. **b** Rotation of the y-axis of the local-level frame (shown with *red/dark arrows*) for a near polar crossing trajectory at various latitudes

about the z-axis with respect to the l-frame. The angle between the y-axis of the wander frame and north is known as the wander angle α. The rotation rate of this angle is given as

$$\dot{\alpha} = -\dot{\lambda} \sin \varphi \tag{2.38}$$

The wander frame (in relation to the l-frame) is shown in Fig. 2.3a, and is defined as

a. The origin coincides with the center of the sensor frame (origin of inertial sensor triad).

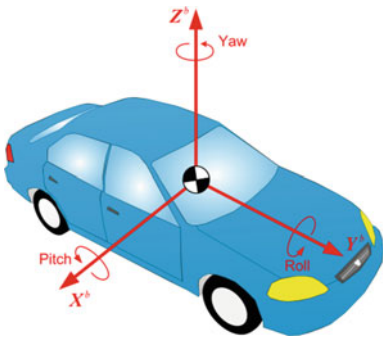


Fig. 2.4 The body frame of a moving platform

b. The z-axis is orthogonal to the reference ellipsoid pointing upward.
c. The y-axis rotates by an angle α anticlockwise from north.
d. The x-axis is orthogonal to the y and z axes and forms a right-handed coordinate frame.

In this book the wander frame is referred to as the w-frame.

2.2.5 Computational Frame

For the ensuing discussion, the computational frame is defined to be any reference frame used in the implementation of the equations of motion. It can be any of the abovementioned coordinate frames, and is referred to as the k-frame.

2.2.6 Body Frame

In most applications, the sensitive axes of the accelerometer sensors are made to coincide with the axes of the moving platform in which the sensors are mounted. These axes are usually known as the body frame.

The body frame used in this book is shown in Fig. 2.4, and is defined as

a. The origin usually coincides with the center of gravity of the vehicle (this simplifies derivation of kinematic equations).
b. The y-axis points towards the forward direction. It is also called the roll axis as the roll angle is defined around this axis using the right-hand rule.
c. The x-axis points towards the transverse direction. It is also called the pitch axis, as the pitch angle corresponds to rotations around this axis using the right-hand rule.
d. The z-axis points towards the vertical direction completing a right-handed coordinate system. It is also called the yaw axis as the yaw angle corresponds to rotations around this axis using the right-hand rule.

In this book the body frame is referred to as b-frame.

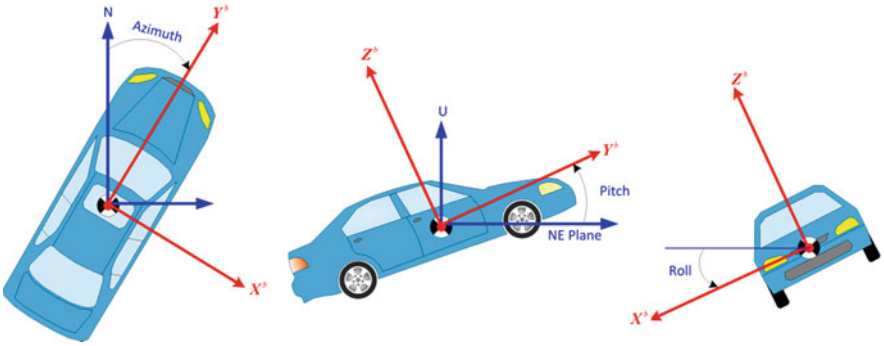


Fig. 2.5 A depiction of a vehicle's azimuth, pitch and roll angles. The body axes are shown in *red*

2.2.6.1 Vehicle Attitude Description

Apart from a vehicle's position, we are also interested in its orientation in order to describe its heading and tilt angles. This involves specifying its rotation about the vertical (z), transversal (x) and forward (y) axes of the b-frame with respect to the l-frame. In general, the rotation angles about the axes of the b-frame are called the Euler angles. For the purpose of this book, the following convention is applied to vehicle attitude angles (Fig. 2.5)

a. **Azimuth** (**or yaw**): Azimuth is the deviation of the vehicle's forward (y) axis from north, measured clockwise in the E-N plane. The yaw angle is similar, but is measured counter clockwise from north. In this book, the azimuth angle is denoted by '*A*' and the yaw angle by '*y*'. Owing to this definition, the vertical axis of the b-frame is also known as the yaw axis (Fig. 2.4).
b. **Pitch**: This is the angle that the forward (y) axis of the b-frame makes with the E-N plane (i.e. local horizontal) owing to a rotation around its transversal (x) axis. This axis is also called the pitch axis, the pitch angle is denoted by '*p*' and follows the right-hand rule (Fig. 2.5).
c. **Roll**: This is the rotation of the b-frame about its forward (y) axis, so the forward axis is also called the roll axis and the roll angle is denoted by '*r*' and follows the right-hand rule.

2.2.7 Orbital Coordinate System

This is a system of coordinates with Keplerian elements to locate a satellite in inertial space. It is defined as follows

a. The origin is located at the focus of an elliptical orbit that coincides with the center of the mass of the Earth.

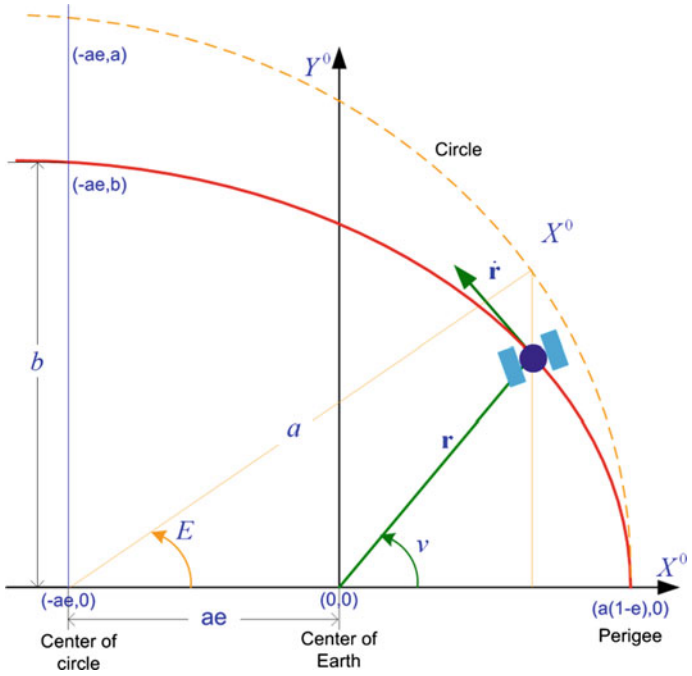


Fig. 2.6 The orbital coordinate system for a satellite

b. The y-axis points towards the descending node, parallel to the minor axis of the orbital ellipse.
c. The x-axis points to the perigee (the point in the orbit nearest the Earth's center) and along the major axis of the elliptical orbit of the satellite.
d. The z-axis is orthogonal to the orbital plane.

The orbital coordinate system is illustrated in Fig. 2.6. It is mentioned here to complete the discussion of the frames used in navigation (it will be discussed in greater detail in Chap. 3).

2.3 Coordinate Transformations

The techniques for transforming a vector from one coordinate frame into another can use direction cosines, rotation (Euler) angles or quaternions. They all involve a rotation matrix which is called either the transformation matrix or the direction cosine matrices (DCM), and is represented as R_k^l where the subscript represents the frame from which the vector originates and the superscript is the target frame. For example, a vector $\mathbf{r}^k$ in a coordinate frame k can be represented by another vector $\mathbf{r}^l$ in a coordinate frame l by applying a rotation matrix R_k^l as follows

$$\mathbf{r}^l = R_k^l \mathbf{r}^k \tag{2.39}$$

If Euler angles are used, these readily yield the elementary matrices required to construct the DCM.

2.3.1 Euler Angles and Elementary Rotational Matrices

A transformation between two coordinate frames can be accomplished by carrying out a rotation about each of the three axes. For example, a transformation from the reference frame a to the new coordinate frame b involves first making a rotation of angle γ about the z-axis, then a rotation of an angle β about the new x-axis, and finally a rotation of an angle α about the new y-axis. In these rotations, α, β and γ are the Euler angles.

To transform a vector $\mathbf{r}^a = [x^a, y^a, z^a]$ from frame a to frame d where the two frames are orientated differently in space, we align frame a with frame d using the three rotations specified above, each applying a suitable direction cosine matrix. The individual matrices can be obtained by considering each rotation, one by one.

First we consider the x-y plane of frame a in which the projection of vector r (represented by r_1) makes an angle θ_1 with the x-axis. We therefore rotate frame a around its z-axis through an angle γ to obtain the intermediate frame b, as illustrated in Fig. 2.7.

According to this figure, the new coordinates are represented by x^b, y^b, z^b and can be expressed as

$$x^b = r_1 \cos(\theta_1 - \gamma) \tag{2.40}$$

$$y^b = r_1 \sin(\theta_1 - \gamma) \tag{2.41}$$

Since the rotation was performed around the z-axis, this remains unchanged

$$z^b = z^a \tag{2.42}$$

Using the following trigonometric identities

$$\begin{aligned} \sin(A \pm B) &= \sin A \cos B \pm \cos A \sin B \\ \cos(A \pm B) &= \cos A \cos B \mp \sin A \sin B \end{aligned} \tag{2.43}$$

Equations (2.40) and (2.41) can be written as

$$x^b = r_1 \cos\theta_1 \cos\gamma + r_1 \sin\theta_1 \sin\gamma \tag{2.44}$$

$$y^b = r_1 \sin\theta_1 \cos\gamma - r_1 \cos\theta_1 \sin\gamma \tag{2.45}$$

The original coordinates of vector $\mathbf{r}_1$ in the x-y plane can be expressed in terms of angle θ_1 as

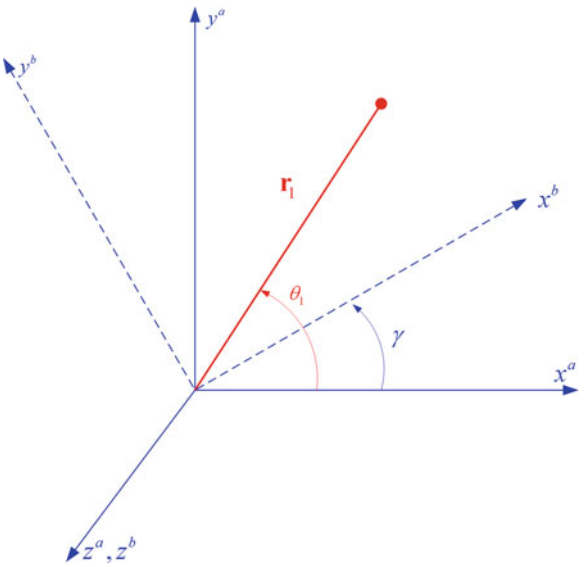


Fig. 2.7 The first rotation of frame 'a' about its z^a-axis

$$x^a = r_1 \cos\theta_1 \tag{2.46}$$

$$y^a = r_1 \sin\theta_1 \tag{2.47}$$

Substituting the above values in Eqs. (2.44) and (2.45) produces

$$x^b = x^a \cos\gamma + y^a \sin\gamma \tag{2.48}$$

$$y^b = -x^a \sin\gamma + y^a \cos\gamma \tag{2.49}$$

and we have shown that

$$z^b = z^a \tag{2.50}$$

In matrix form, the three equations above can be written as

$$\begin{bmatrix} x^b \\ y^b \\ z^b \end{bmatrix} = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^a \\ y^a \\ z^a \end{bmatrix} \tag{2.51}$$

$$\begin{bmatrix} x^b \\ y^b \\ z^b \end{bmatrix} = R_a^b \begin{bmatrix} x^a \\ y^a \\ z^a \end{bmatrix} \tag{2.52}$$

where R_a^b is the elementary DCM which transforms the coordinates x^a, y^a, z^a to x^b, y^b, z^b in a frame rotated by an angle γ around the z-axis of frame a.

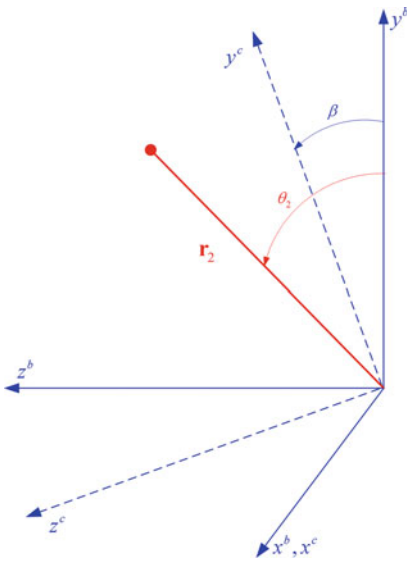


Fig. 2.8 The second rotation of rotated frame 'b' about x^b-axis

For the second rotation, we consider the y-z plane of the new coordinate frame b, and rotate it by an angle β around its x-axis to an intermediate frame c as shown in Fig. 2.8.

In a similar fashion it can be shown that the new coordinates x^c, y^c, z^c can be expressed in terms of x^b, y^b, z^b as follows

$$\begin{bmatrix} x^c \\ y^c \\ z^c \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x^b \\ y^b \\ z^b \end{bmatrix} \tag{2.53}$$

$$\begin{bmatrix} x^c \\ y^c \\ z^c \end{bmatrix} = R_b^c \begin{bmatrix} x^b \\ y^b \\ z^b \end{bmatrix} \tag{2.54}$$

where R_b^c is the elementary DCM which transforms the coordinates x^b, y^b, z^b to x^c, y^c, z^c in a frame rotated by an angle β around the x-axis of frame b.

For the third rotation, we consider the x-z plane of new coordinate frame c, and rotate it by an angle α about its y-axis to align it with coordinate frame d as shown in Fig. 2.9.

The final coordinates x^d, y^d, z^d can be expressed in terms of x^c, y^c, z^c as follows

$$\begin{bmatrix} x^d \\ y^d \\ z^d \end{bmatrix} = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix} \begin{bmatrix} x^c \\ y^c \\ z^c \end{bmatrix} \tag{2.55}$$

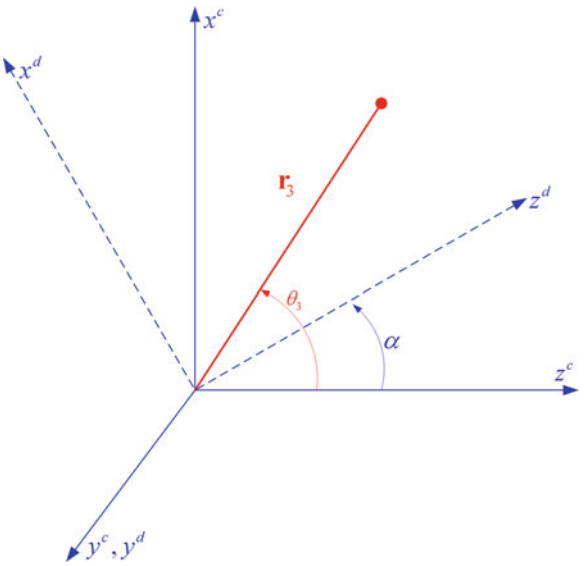


Fig. 2.9 The third rotation of rotated frame 'c' about y^c-axis

$$\begin{bmatrix} x^d \\ y^d \\ z^d \end{bmatrix} = R_c^d \begin{bmatrix} x^c \\ y^c \\ z^c \end{bmatrix} \tag{2.56}$$

where R_c^d is the elementary DCM which transforms the coordinates x^c, y^c, z^c to x^d, y^d, z^d in the final desired frame d rotated by an angle α around the y-axis of frame c.

We can combine all three rotations by multiplying the cosine matrices into a single transformation matrix as

$$R_a^d = R_c^d R_b^c R_a^b \tag{2.57}$$

The final DCM for these particular set of rotations can be given as

$$R_a^d = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.58}$$

$$R_a^d = \begin{bmatrix} \cos\alpha\cos\gamma - \sin\beta\sin\alpha\sin\gamma & \cos\alpha\sin\gamma + \cos\gamma\sin\beta\sin\alpha & -\cos\beta\sin\alpha \\ -\cos\beta\sin\gamma & \cos\beta\cos\gamma & \sin\beta \\ \cos\gamma\sin\alpha + \cos\alpha\sin\beta\sin\gamma & \sin\alpha\sin\gamma - \cos\alpha\cos\gamma\sin\beta & \cos\beta\cos\alpha \end{bmatrix} \tag{2.59}$$

The inverse transformation from frame d to a is therefore

$$\begin{aligned} R_d^a = \left(R_a^d\right)^{-1} = \left(R_a^d\right)^T &= \left(R_c^d R_b^c R_a^b\right)^T \\ &= \left(R_a^b\right)^T \left(R_b^c\right)^T \left(R_c^d\right)^T \end{aligned} \tag{2.60}$$

It should be noted that the final transformation matrix is contingent upon the order of the applied rotations, as is evident from the fact that $R_b^c R_a^b \neq R_a^b R_b^c$. The order of rotations is dictated by the specific application. We will see in Sect. 2.3.6 that a different order of rotation is required and the elementary matrices are multiplied in a different order to yield a different final transformation matrix.

For small values of α, β and γ we can use the following approximations

$$\cos\theta \approx 1\,,\ \sin\theta \approx \theta \tag{2.61}$$

Using these approximations and ignoring the product of the small angles, we can reduce the DCM to

$$\begin{aligned} R_a^d &\approx \begin{bmatrix} 1 & \gamma & -\alpha \\ -\gamma & 1 & \beta \\ \alpha & -\beta & 1 \end{bmatrix} \\ R_a^d &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & -\gamma & \alpha \\ \gamma & 0 & -\beta \\ -\alpha & \beta & 0 \end{bmatrix} \\ R_a^d &= I - \Psi \end{aligned} \tag{2.62}$$

where Ψ is the skew-symmetric matrix for the small Euler angles. For the small-angle approximation, the order of rotation is no longer important since in all cases the final result will always be the matrix of the Eq. (2.62). Similarly, it can be verified that

$$\begin{aligned} R_d^a &\approx \begin{bmatrix} 1 & \gamma & -\alpha \\ -\gamma & 1 & \beta \\ \alpha & -\beta & 1 \end{bmatrix}^T \\ R_d^a &= I - \Psi^T \end{aligned} \tag{2.63}$$

2.3.2 Transformation Between ECI and ECEF

The angular velocity vector between the i-frame and the e-frame as a result of the rotation of the Earth is

$$\boldsymbol{\omega}_{ie}^e = (0, 0, \omega_e)^T \tag{2.64}$$

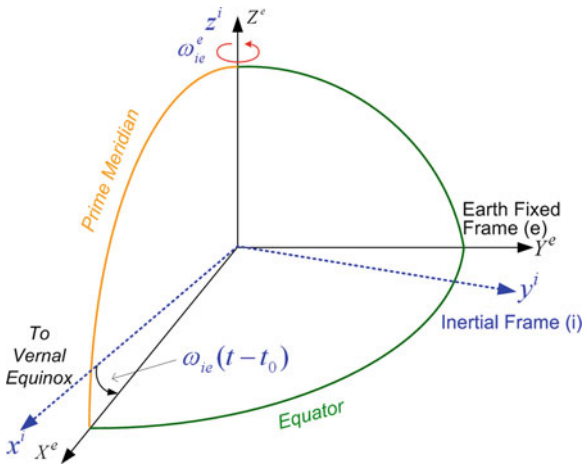


Fig. 2.10 Transformation between the e-frame and the i-frame

where ω_e denotes the magnitude of the Earth's rotation rate. The transformation from the i-frame to the e-frame is a simple rotation of the i-frame about the z-axis by an angle $\omega_e t$ where t is the time since the reference epoch (Fig. 2.10). The rotation matrix corresponds to the elementary matrix R_a^b, and when denoted R_i^e which can be expressed as

$$R_i^e = \begin{bmatrix} \cos\omega_e t & \sin\omega_e t & 0 \\ -\sin\omega_e t & \cos\omega_e t & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.65}$$

Transformation from the e-frame to the i-frame can be achieved through R_e^i, the inverse of R_i^e. Since rotation matrices are orthogonal

$$R_e^i = \left(R_i^e\right)^{-1} = \left(R_i^e\right)^T \tag{2.66}$$

2.3.3 Transformation Between LLF and ECEF

From Fig. 2.11 it can be observed that to align the l-frame with the e-frame, the l-frame must be rotated by $\varphi - 90$ degrees around its x-axis (east direction) and then by $-90 - \lambda$ degrees about its z-axis (up direction).

For the definition of elementary direction cosine matrices, the transformation from the l-frame to the e-frame is

$$R_l^e = R_a^b(-\lambda - 90)R_b^c(\varphi - 90) \tag{2.67}$$

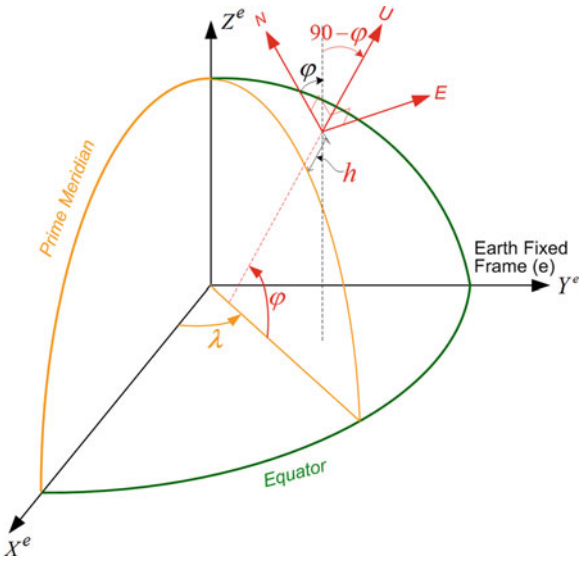


Fig. 2.11 The LLF in relation to the ECEF frame

$$R_l^e = \begin{bmatrix} \cos(-\lambda - 90) & \sin(-\lambda - 90) & 0 \\ -\sin(-\lambda - 90) & \cos(-\lambda - 90) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi - 90) & \sin(\varphi - 90) \\ 0 & -\sin(\varphi - 90) & \cos(\varphi - 90) \end{bmatrix} \tag{2.68}$$

$$R_l^e = \begin{bmatrix} -\sin\lambda & -\cos\lambda & 0 \\ \cos\lambda & -\sin\lambda & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sin\varphi & -\cos\varphi \\ 0 & \cos\varphi & \sin\varphi \end{bmatrix} \tag{2.69}$$

$$R_l^e = \begin{bmatrix} -\sin\lambda & -\sin\varphi\cos\lambda & \cos\varphi\cos\lambda \\ \cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi\sin\lambda \\ 0 & \cos\varphi & \sin\varphi \end{bmatrix} \tag{2.70}$$

The reverse transformation is

$$R_e^l = \left(R_l^e\right)^{-1} = \left(R_l^e\right)^T \tag{2.71}$$

2.3.4 Transformation Between LLF and Wander Frame

The wander frame has a rotation about the z-axis of the l-frame by a wander angle α, as depicted in Fig. 2.12. Thus the transformation matrix from the w-frame frame to the l-frame corresponds to the elementary matrix R_a^b with an angle $-\alpha$, and is expressed as

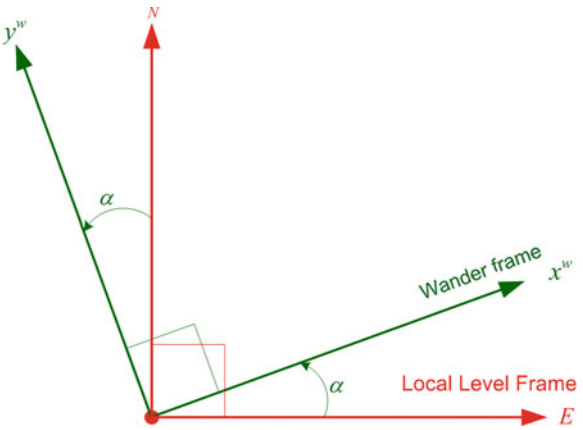


Fig. 2.12 The relationship between the l-frame and the w-frame (the third axes of these the frames are not shown because they coincide and point out of the page towards the reader)

$$R_w^l = \begin{bmatrix} \cos(-\alpha) & \sin(-\alpha) & 0 \\ -\sin(-\alpha) & \cos(-\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.72}$$

$$R_w^l = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.73}$$

and

$$R_l^w = \left(R_w^l\right)^{-1} = \left(R_w^l\right)^T \tag{2.74}$$

2.3.5 Transformation Between ECEF and Wander Frame

This transformation is obtained by first going from the w-frame to the l-frame and then from the l-frame to the e-frame

$$R_w^e = R_l^e R_w^l \tag{2.75}$$

$$R_w^e = \begin{bmatrix} -\sin\lambda & -\sin\varphi\cos\lambda & \cos\varphi\cos\lambda \\ \cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi\sin\lambda \\ 0 & \cos\varphi & \sin\varphi \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.76}$$

$$R_w^e = \begin{bmatrix} -\sin\lambda\cos\alpha - \cos\lambda\sin\varphi\sin\alpha & \sin\lambda\sin\alpha - \cos\lambda\sin\varphi\cos\alpha & \cos\lambda\cos\varphi \\ \cos\lambda\cos\alpha - \sin\lambda\sin\varphi\sin\alpha & -\cos\lambda\sin\alpha - \sin\lambda\sin\varphi\cos\alpha & \sin\lambda\cos\varphi \\ \cos\varphi\sin\alpha & \cos\varphi\cos\alpha & \sin\varphi \end{bmatrix} \tag{2.77}$$

The inverse is

$$R_e^w = \left(R_w^e\right)^{-1} = \left(R_w^e\right)^T \tag{2.78}$$

2.3.6 Transformation Between Body Frame and LLF

One of the important direction cosine matrices is R_b^l, which transforms a vector from the b-frame to the l-frame, a requirement during the mechanization process. This is expressed in terms of yaw, pitch and roll Euler angles. According to the definitions of these specific angles and the elementary direction cosine matrices, R_b^l can be expressed as

$$\begin{aligned} R_b^l &= \left(R_l^b\right)^{-1} = \left(R_l^b\right)^T = \left(R_c^d R_b^c R_a^b\right)^T \\ &= \left(R_a^b\right)^T \left(R_b^c\right)^T \left(R_c^d\right)^T \end{aligned} \tag{2.79}$$

Substituting the elementary matrices into this equation gives

$$R_b^l = \begin{bmatrix} \cos y & \sin y & 0 \\ -\sin y & \cos y & 0 \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos p & \sin p \\ 0 & -\sin p & \cos p \end{bmatrix}^T \begin{bmatrix} \cos r & 0 & -\sin r \\ 0 & 1 & 0 \\ \sin r & 0 & \cos r \end{bmatrix}^T \tag{2.80}$$

$$R_b^l = \begin{bmatrix} \cos y & -\sin y & 0 \\ \sin y & \cos y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos p & -\sin p \\ 0 & \sin p & \cos p \end{bmatrix} \begin{bmatrix} \cos r & 0 & \sin r \\ 0 & 1 & 0 \\ -\sin r & 0 & \cos r \end{bmatrix} \tag{2.81}$$

$$R_b^l = \begin{bmatrix} \cos y\cos r - \sin y\sin p\sin r & -\sin y\cos p & \cos y\sin r + \sin y\sin p\cos r \\ \sin y\cos r + \cos y\sin p\sin r & \cos y\cos p & \sin y\sin r - \cos y\sin p\cos r \\ -\cos p\sin r & \sin p & \cos p\cos r \end{bmatrix} \tag{2.82}$$

where 'p', 'r' and 'y' are the pitch, roll and yaw angles. With a known R_b^l, these angles can be calculated as

$$p = \tan^{-1}\left\{\frac{R_b^l(3,2)}{\sqrt{\left[R_b^l(1,2)\right]^2+\left[R_b^l(2,2)\right]^2}}\right\} \tag{2.83}$$

$$y = -\tan^{-1}\left[\frac{R_b^l(1,2)}{R_b^l(2,2)}\right] \tag{2.84}$$

$$r = -\tan^{-1}\left[\frac{R_b^l(3,1)}{R_b^l(3,3)}\right] \tag{2.85}$$

A transformation from the l-frame to the b-frame can be achieved by the inverse rotation matrix, R_b^l, as follows

$$R_l^b = \left(R_b^l\right)^{-1} = \left(R_b^l\right)^T \tag{2.86}$$

2.3.7 Transformation From Body Frame to ECEF and ECI Frame

Two other important transformations are from the b-frame to the e-frame and the i-frames. Their rotation matrices can be computed from those already defined as follows.

For the body frame to the e-frame

$$R_b^e = R_l^e R_b^l \tag{2.87}$$

For the body frame to the i-frame

$$R_b^i = R_e^i R_b^e \tag{2.88}$$

Their inverses are

$$R_e^b = \left(R_b^e\right)^{-1} = \left(R_b^e\right)^T \tag{2.89}$$

$$R_i^b = \left(R_b^i\right)^{-1} = \left(R_b^i\right)^T \tag{2.90}$$

2.3.8 Time Derivative of the Transformation Matrix

If a coordinate reference frame k rotates with angular velocity $\boldsymbol{\omega}$ relative to another frame m, the transformation matrix between the two is composed of a set of time variable functions. The time rate of change of the transformation matrix $\dot{R}_k^m$ can be described using a set of differential equations. The frame in which the time differentiation occurs is usually identified by the superscript of the variable.

At time t, the two frames m and k are related by a DCM $R_k^m(t)$. After a short time δt, frame k rotates to a new orientation and the new DCM at time $t+\delta t$ is $R_k^m(t+\delta t)$. The time derivative of the DCM R_k^m is therefore

$$\begin{aligned}\dot{R}_k^m &= \lim_{\delta t\to 0}\frac{\delta R_k^m}{\delta t}\\ \dot{R}_k^m &= \lim_{\delta t\to 0}\frac{R_k^m(t+\delta t)-R_k^m(t)}{\delta t}\end{aligned} \tag{2.91}$$

The transformation at time $t+\delta t$ is the outcome of the transformation up to time t followed by the small change of the m frame that occurs during the brief interval δt. Hence $R_k^m(t+\delta t)$ can be written as the product of two matrices

$$R_k^m(t+\delta t) = \delta R^m R_k^m(t) \tag{2.92}$$

From Eq. (2.62), the small angle transformation can be given as

$$\delta R^m = I - \Psi^m \tag{2.93}$$

Substituting Eq. (2.93) into (2.92) gives

$$R_k^m(t+\delta t) = (I-\Psi^m)R_k^m(t) \tag{2.94}$$

and substituting this back into Eq. (2.91) produces

$$\begin{aligned}\dot{R}_k^m &= \lim_{\delta t\to 0}\frac{(I-\Psi^m)R_k^m(t)-R_k^m(t)}{\delta t}\\ \dot{R}_k^m &= \lim_{\delta t\to 0}\frac{(I-\Psi^m-I)R_k^m(t)}{\delta t}\\ \dot{R}_k^m &= \lim_{\delta t\to 0}\frac{-\Psi^m R_k^m(t)}{\delta t}\\ \dot{R}_k^m &= -\left(\lim_{\delta t\to 0}\frac{\Psi^m}{\delta t}\right)R_k^m(t)\end{aligned} \tag{2.95}$$

When $\delta t\to 0$, $\Psi^m/\delta t$ is the skew-symmetric form of the angular velocity vector of the m frame with respect to the k frame during the time increment δt. Due to the limiting process, the angular velocity can also be referenced to the k frame

$$\lim_{\delta t\to 0}\frac{\Psi^m}{\delta t} = \Omega_{km}^m \tag{2.96}$$

Substituting Eq. (2.96) into (2.95) gives

$$\dot{R}_k^m = -\Omega_{km}^m R_k^m \tag{2.97}$$

From the relation $\Omega_{km}^m = -\Omega_{mk}^m$, this becomes

$$\dot{R}_k^m = \Omega_{mk}^m R_k^m \tag{2.98}$$

From Eq. (2.19)

$$\Omega^m_{mk} = R^m_k \Omega^k_{mk} R^k_m \tag{2.99}$$

Substituting this into (2.98) gives

$$\dot{R}^m_k = R^m_k \Omega^k_{mk} R^k_m R^m_k \tag{2.100}$$

Finally we get the important equation for the rate of change of the DCM as

$$\dot{R}^m_k = R^m_k \Omega^k_{mk} \tag{2.101}$$

This implies that the time derivative of the rotation matrix is related to the angular velocity vector $\boldsymbol{\omega}$ of the relative rotation between the two coordinate frames. If we have the initial transformation matrix between the body and inertial frames R^i_b, then we can update the change of the rotation matrix using gyroscope output Ω^b_{ib}.

2.3.9 Time Derivative of the Position Vector in the Inertial Frame

For a position vector $\mathbf{r}^b$, the transformation of its coordinates from the b-frame to the inertial frame is

$$\mathbf{r}^i = R^i_b \mathbf{r}^b \tag{2.102}$$

Differentiating both sides with respect to time leads to

$$\dot{\mathbf{r}}^i = \dot{R}^i_b \mathbf{r}^b + R^i_b \dot{\mathbf{r}}^b \tag{2.103}$$

Substituting the value of $\dot{R}^i_b$ from Eq. (2.101) into (2.103) gives

$$\dot{\mathbf{r}}^i = \left(R^i_b \Omega^b_{ib}\right)\mathbf{r}^b + R^i_b \dot{\mathbf{r}}^b \tag{2.104}$$

A rearrangement of the terms gives

$$\dot{\mathbf{r}}^i = R^i_b \left(\dot{\mathbf{r}}^b + \Omega^b_{ib} \mathbf{r}^b\right) \tag{2.105}$$

which describes the transformation of the velocity vector from the b-frame to the inertial frame. This is often called the Coriolis equation.

2.3.10 Time Derivative of the Velocity Vector in the Inertial Frame

The time derivative of the velocity vector is obtained by differentiating Eq. (2.105) as follows

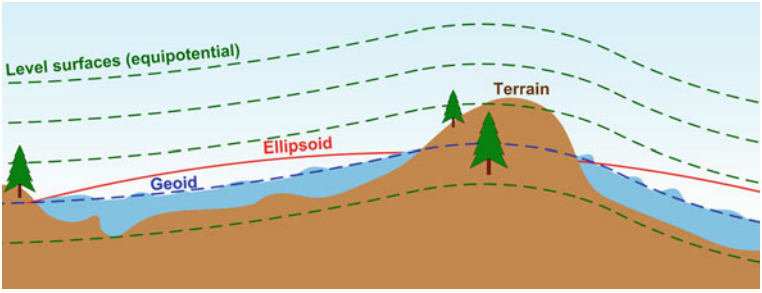


Fig. 2.13 A depiction of various surfaces of the Earth

$$\begin{aligned}\ddot{\mathbf{r}}^i &= \dot{R}_b^i \dot{\mathbf{r}}^b + R_b^i \ddot{\mathbf{r}}^b + \dot{R}_b^i \Omega_{ib}^b \mathbf{r}^b + \left(\dot{\Omega}_{ib}^b \mathbf{r}^b + \Omega_{ib}^b \dot{\mathbf{r}}^b\right) R_b^i \\ \ddot{\mathbf{r}}^i &= \dot{R}_b^i \dot{\mathbf{r}}^b + R_b^i \ddot{\mathbf{r}}^b + \dot{R}_b^i \Omega_{ib}^b \mathbf{r}^b + R_b^i \dot{\Omega}_{ib}^b \mathbf{r}^b + R_b^i \Omega_{ib}^b \dot{\mathbf{r}}^b\end{aligned} \tag{2.106}$$

Substituting the value of $\dot{R}_b^i$ from Eq. (2.101) yields

$$\begin{aligned}\ddot{\mathbf{r}}^i &= R_b^i \Omega_{ib}^b \dot{\mathbf{r}}^b + R_b^i \ddot{\mathbf{r}}^b + R_b^i \Omega_{ib}^b \Omega_{ib}^b \mathbf{r}^b + R_b^i \dot{\Omega}_{ib}^b \mathbf{r}^b + R_b^i \Omega_{ib}^b \dot{\mathbf{r}}^b \\ \ddot{\mathbf{r}}^i &= R_b^i \left(\Omega_{ib}^b \dot{\mathbf{r}}^b + \ddot{\mathbf{r}}^b + \Omega_{ib}^b \Omega_{ib}^b \mathbf{r}^b + \dot{\Omega}_{ib}^b \mathbf{r}^b + \Omega_{ib}^b \dot{\mathbf{r}}^b\right) \\ \ddot{\mathbf{r}}^i &= R_b^i \left(2\Omega_{ib}^b \dot{\mathbf{r}}^b + \ddot{\mathbf{r}}^b + \Omega_{ib}^b \Omega_{ib}^b \mathbf{r}^b + \dot{\Omega}_{ib}^b \mathbf{r}^b\right)\end{aligned} \tag{2.107}$$

and rearrangement gives

$$\ddot{\mathbf{r}}^i = R_b^i \left(\ddot{\mathbf{r}}^b + 2\Omega_{ib}^b \dot{\mathbf{r}}^b + \dot{\Omega}_{ib}^b \mathbf{r}^b + \Omega_{ib}^b \Omega_{ib}^b \mathbf{r}^b\right) \tag{2.108}$$

where

$\ddot{\mathbf{r}}^b$	is the acceleration of the moving object in the b-frame
Ω_{ib}^b	is the angular velocity of the moving object measured by a gyroscope
$2\Omega_{ib}^b \dot{\mathbf{r}}^b$	is the Coriolis acceleration
$\dot{\Omega}_{ib}^b \mathbf{r}^b$	is the tangential acceleration
$\Omega_{ib}^b \Omega_{ib}^b \mathbf{r}^b$	is the centripetal acceleration

2.4 The Geometry of the Earth

Although the Earth is neither a sphere nor a perfect ellipsoid, it is approximated by an ellipsoid for computational convenience. The ellipsoid and various surfaces that are useful for understanding the geometry of the Earth's shape are depicted in Fig. 2.13.

2.4.1 Important Definitions

At this point, it will be useful to describe some of the important definitions which will assist in understanding the ensuing analysis. For further details, the reader is referred to (Titterton and Weston 2005; Vanícek and Krakiwsky 1986).

- **Physical Surface—Terrain**": This is defined as the interface between the solid and fluid masses of the Earth and its atmosphere. It is the actual surface that we walk or float on.
- **Geometric Figure—Geoid**": This is the equipotential surface (surface of constant gravity) best fitting the average sea level in the least squares sense (ignoring tides and other dynamical effects in the oceans). It can be thought of as the idealized mean sea level extended over the land portion of the globe. The geoid is a smooth surface but its shape is irregular and it does not provide the simple analytic expression needed for navigational computations.
- **Reference Ellipsoid—Ellipsoid**": This mathematically defined surface approximates the geoid by an ellipsoid that is made by rotating an ellipse about its minor axis, which is coincident with the mean rotational axis of the Earth. The center of the ellipsoid is coincident with the Earth's center of mass.

The ellipsoid is the most analytically convenient surface to work with for navigational purposes. Its shape is defined by two geometric parameters called the semimajor axis and the semiminor axis. These are typically represented by a and b respectively, as in Fig. 2.14. The geoid height N is the distance along the ellipsoidal normal from the surface of the ellipsoid to the geoid. The orthometric height H is the distance from the geoid to the point of interest. The geodetic height h (also known as altitude) is the sum of the geoid and orthometric heights ($h = H + N$). Various parameter sets have been defined to model the ellipsoid. This book uses the world geodetic system (WGS)-84 whose defining parameters (Torge 1980; Vanícek and Krakiwsky 1986) are

Semimajor axis (equatorial radius) $a = 6,378,137.0\,\text{m}$
Reciprocal flattening $\frac{1}{f} = 298.257223563$
Earth's rotation rate $\omega_e = 7.292115 \times 10^{-5}\,\text{rad/s}$
Gravitational constant $GM = 3.986004418 \times 10^{14}\,\text{m}^3/\text{s}^2$

Other derived parameters of interest are

Flatness $f = \frac{a-b}{a} = 0.00335281$
Semiminor axis $b = a(1-f) = 6356752.3142\,\text{m}$
Eccentricity $e = \sqrt{\frac{a^2-b^2}{a^2}} = \sqrt{f(2-f)} = 0.08181919$

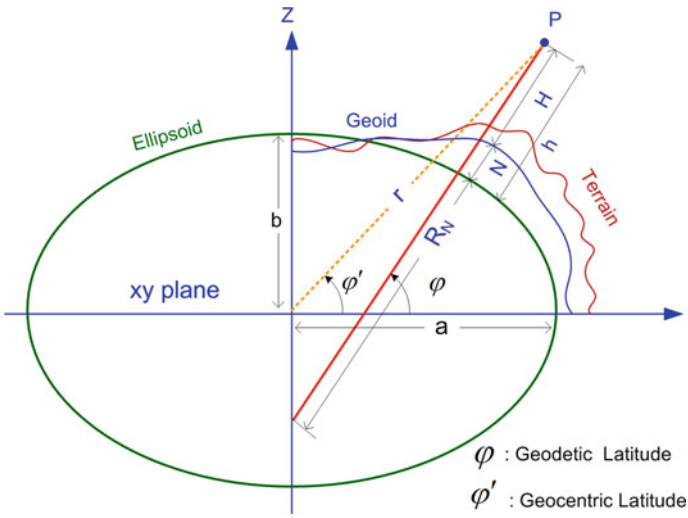


Fig. 2.14 The relationship between various Earth surfaces (highly exaggerated) and a depiction of the ellipsoidal parameters

2.4.2 Normal and Meridian Radii

In navigation two radii of curvature are of particular interest, the normal radius and the meridian radius. These govern the rates at which the latitude and longitude change as a navigating platform moves on or near the surface of the Earth.

The normal radius R_N is defined for the east-west direction, and is also known as the great normal or the radius of curvature of the prime vertical

$$R_N = \frac{a}{\left(1 - e^2 \sin^2 \varphi\right)^{\frac{1}{2}}} \tag{2.109}$$

The meridian radius of curvature is defined for the north-south direction and is the radius of the ellipse

$$R_M = \frac{a(1 - e^2)}{\left(1 - e^2 \sin^2 \varphi\right)^{\frac{3}{2}}} \tag{2.110}$$

A derivation of these radii can be found in Appendix A, and for further insight the reader is referred to (Grewal et al. 2007; Rogers 2007).

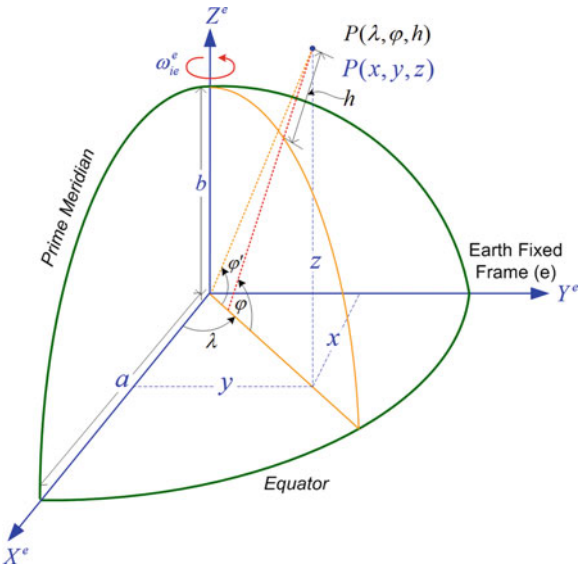


Fig. 2.15 Two types of ECEF coordinates and their interrelationship

2.5 Types of Coordinates in the ECEF Frame

It is important to distinguish between the two common coordinate systems of the e-frame, known as the 'rectangular' and 'geodetic' systems.

2.5.1 Rectangular Coordinates in the ECEF Frame

Rectangular coordinates are like traditional Cartesian coordinates, and represent the position of a point with its x, y and z vector components aligned parallel to the corresponding e-frame axes (Fig. 2.15).

2.5.2 Geodetic Coordinates in the ECEF Frame

Geodetic (also referred to as ellipsoidal or curvilinear) coordinates are defined in a way that is more intuitive for positioning applications on or near the Earth. These coordinates are defined (Farrell 2008) as

a. Latitude (φ) is the angle in the meridian plane from the equatorial plane to the ellipsoidal normal at the point of interest.
b. Longitude (λ) is the angle in the equatorial plane from the prime meridian to the projection of the point of interest onto the equatorial plane.

c. Altitude h is the distance along the ellipsoidal normal between the surface of the ellipsoid and the point of interest.

The two types of e-frame coordinates and their interrelationship are illustrated in Fig. 2.15.

2.5.3 Conversion From Geodetic to Rectangular Coordinates in the ECEF Frame

In navigation, it is often necessary to convert from geodetic e-frame coordinates to rectangular e-frame coordinates. The following relationship (see Appendix B for a derivation) accomplishes this

$$\begin{bmatrix} x^e \\ y^e \\ z^e \end{bmatrix} = \begin{bmatrix} (R_N + h)\cos\varphi\cos\lambda \\ (R_N + h)\cos\varphi\sin\lambda \\ \{R_N(1-e^2) + h\}\sin\varphi \end{bmatrix} \tag{2.111}$$

where

(x^e, y^e, z^e)	are the e-frame rectangular coordinates
R_N	is the normal radius
h	is the ellipsoidal height
λ	is the longitude
φ	is the latitude
e	is the eccentricity.

2.5.4 Conversion From Rectangular to Geodetic Coordinates in the ECEF Frame

Converting rectangular to geodetic coordinates is not straightforward, because the analytical solution results in a fourth-order equation. There are approximate closed form solutions but an iterative scheme is usually employed.

2.5.4.1 Closed-Form Algorithm

This section will describe a closed form algorithm to calculate e-frame geodetic coordinates directly from e-frame rectangular coordinates though series expansion (Hofmann-Wellenhof et al. 2008). An alternate method is detailed in Appendix C.

Longitude is

$$\lambda = 2 \arctan \frac{y^e}{x^e + \sqrt{(x^e)^2 + (y^e)^2}} \tag{2.112}$$

latitude is

$$\varphi = \arctan \frac{z^e + (e')^2 b \sin^3 \theta}{p - e^2 a \cos^3 \theta} \tag{2.113}$$

where

$$\theta = \arctan \frac{z^e a}{pb}$$

$$e' = \sqrt{\frac{a^2 - b^2}{b^2}}$$

$$p = \sqrt{(x^e)^2 + (y^e)^2}$$

and height is

$$h = \frac{p}{\cos \varphi} - N \tag{2.114}$$

where

$$N = \frac{a^2}{\sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}}$$

2.5.4.2 Iterative Algorithm

The iterative algorithm is derived in Appendix D and implemented by taking the following steps

a. Initialize the altitude as

$$h_0 = 0 \tag{2.115}$$

b. Choose an arbitrary value of latitude either from a previous measurement (if one is available) or by using the approximation

$$\varphi_0 = \tan^{-1}\left[\frac{z^e}{P^e(1 - e^2)}\right] \tag{2.116}$$

c. The geodetic longitude is calculated as

$$\lambda = \tan^{-1}\left(\frac{y^e}{x^e}\right) \tag{2.117}$$

d. Starting from $i = 1$, iterate as follows

$$R_{N_i} = \frac{a}{\left(1 - e^2 \sin^2 \varphi_{i-1}\right)^{1/2}} \tag{2.118}$$

$$h_i = \frac{\sqrt{(x^e)^2 + (y^e)^2}}{\cos \varphi_{i-1}} - R_{N_i} \tag{2.119}$$

$$\varphi_i = \tan^{-1}\left\{\frac{z^e}{\sqrt{(x^e)^2 + (y^e)^2}} \cdot \frac{(R_{N_i} + h_i)}{R_{N_i}(1 - e^2) + h_i}\right\} \tag{2.120}$$

e. Compare φ_i, φ_{i-1} and h_i, h_{i-1}; if convergence has been achieved then stop, otherwise repeat step d using the new values.

2.6 Earth Gravity

The gravity field vector is different from the gravitational field vector. Due to the Earth's rotation, the gravity field is used more frequently and is defined as

$$\mathbf{g} = \bar{\mathbf{g}} - \boldsymbol{\Omega}_{ie}\boldsymbol{\Omega}_{ie}\mathbf{r} \tag{2.121}$$

where $\bar{\mathbf{g}}$ is the gravitational vector, $\boldsymbol{\Omega}_{ie}$ is the skew-symmetric representation of the Earth's rotation vector $\boldsymbol{\omega}_{ie}$ with respect to the i-frame, and $\mathbf{r}$ is the geocentric position vector. The second term in the above equation denotes the centripetal acceleration due to the rotation of the Earth around its axis. Usually, the gravity vector is given in the l-frame. Because the normal gravity vector on the ellipsoid coincides with the ellipsoidal normal, the east and the north components of the normal gravity vector are zero and only third component is non-zero

$$\mathbf{g}^l = [\,0 \quad 0 \quad -g\,]^T \tag{2.122}$$

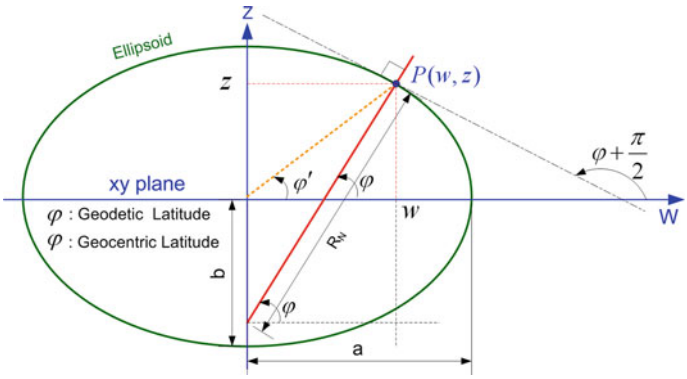


Fig. 2.16 A meridian cross section of the reference ellipsoid containing the projection of the point of interest P

The magnitude of the normal gravity vector over the surface of the ellipsoid can be computed as a function of latitude and height by a closed form expression known as the Somigliana formula (Schwarz and Wei Jan 1999), which is

$$\gamma = a_1\left(1 + a_2 \sin^2 \varphi + a_3 \sin^4 \varphi\right) + \left(a_4 + a_5 \sin^2 \varphi\right)h + a_6 h^2 \tag{2.123}$$

where h is the height above the Earth's surface and the coefficients a_1 through a_6 for the 1980 geographic reference system (GRS) are defined as

$$\begin{aligned} a_1 &= 9.7803267714\ \text{m/s}^2; & a_4 &= -0.0000030876910891/\text{s}^2; \\ a_2 &= 0.0052790414; & a_5 &= 0.0000000043977311/\text{s}^2; \\ a_3 &= 0.0000232718; & a_6 &= 0.0000000000007211/\text{ms}^2 \end{aligned}$$

Appendix A

Derivation of Meridian Radius and Normal Radius

For Earth ellipsoids, every meridian is an ellipse with equatorial radius a (called the semimajor axis) and polar radius b (called the semiminor axis). Figure 2.16 shows a meridian cross section of one such ellipse (Rogers 2007).

This ellipse can be described by the equation

$$\frac{w^2}{a^2} + \frac{z^2}{b^2} = 1 \tag{2.124}$$

and the slope of the tangent to point P can be derived by differentiation

$$\frac{2wdw}{a^2} + \frac{2zdz}{b^2} = 0 \tag{2.125}$$

$$\frac{dz}{dw} = -\frac{b^2 w}{a^2 z} \tag{2.126}$$

An inspection of the Fig. 2.16 shows that the derivative of the curve at point P, which is equal to the slope of the tangent to the curve at that point, is

$$\frac{dz}{dw} = \tan\left(\frac{\pi}{2} + \varphi\right) \tag{2.127}$$

$$\frac{dz}{dw} = \frac{\sin\left(\frac{\pi}{2} + \varphi\right)}{\cos\left(\frac{\pi}{2} + \varphi\right)} = \frac{\cos\varphi}{-\sin\varphi} = -\frac{1}{\tan\varphi} \tag{2.128}$$

Therefore

$$\frac{1}{\tan\varphi} = \frac{b^2 w}{a^2 z} \tag{2.129}$$

From the definition of eccentricity, we have

$$\begin{aligned} e^2 &= 1 - \frac{b^2}{a^2} \\ \frac{b^2}{a^2} &= 1 - e^2 \end{aligned} \tag{2.130}$$

and Eq. (2.129) becomes

$$z = w\left(1 - e^2\right)\tan\varphi \tag{2.131}$$

The ellipse described by Eq. (2.124) gives

$$w^2 = a^2\left(1 - \frac{z^2}{b^2}\right) \tag{2.132}$$

Substituting the value of z from Eq. (2.131) yields

$$\begin{aligned} w^2 &= \left(a^2 - \frac{a^2 w^2 (1 - e^2)^2 \tan^2\varphi}{b^2}\right) \\ w^2 &+ \frac{a^2 w^2 (1 - e^2)^2 \tan^2\varphi}{b^2} = a^2 \\ w^2 &\left(\frac{b^2 + a^2 (1 - e^2)^2 \tan^2\varphi}{b^2}\right) = a^2 \end{aligned}$$

$$
\begin{aligned}
w^2 &= \frac{a^2 b^2}{b^2 + a^2(1-e^2)^2 \tan^2 \varphi} \\
w^2 &= \frac{a^2[a^2(1-e^2)]}{a^2(1-e^2) + a^2(1-e^2)^2 \tan^2 \varphi} \\
w^2 &= \frac{a^2}{1 + (1-e^2)\tan^2 \varphi} \\
w^2 &= \frac{a^2}{1 + (1-e^2)\frac{\sin^2 \varphi}{\cos^2 \varphi}} \\
w^2 &= \frac{a^2 \cos^2 \varphi}{\cos^2 \varphi + (1-e^2)\sin^2 \varphi} \\
w^2 &= \frac{a^2 \cos^2 \varphi}{1 - e^2 \sin^2 \varphi}
\end{aligned}
\tag{2.133}
$$

$$
w = \frac{a \cos \varphi}{\left(1 - e^2 \sin^2 \varphi\right)^{\frac{1}{2}}} \tag{2.134}
$$

Substituting this expression for w in Eq. (2.131) produces

$$
z = \frac{a(1-e^2)\sin \varphi}{\left(1 - e^2 \sin^2 \varphi\right)^{\frac{1}{2}}} \tag{2.135}
$$

which will be used later to derive the meridian radius.

It can easily be proved from Fig. 2.16 that

$$
w = R_N \cos \varphi \tag{2.136}
$$

From this and Eq. (2.134) we have the expression for the normal radius, R_N, also known as the radius of curvature in the prime vertical

$$
R_N = \frac{a}{\left(1 - e^2 \sin^2 \varphi\right)^{\frac{1}{2}}} \tag{2.137}
$$

The radius of curvature of an arc of constant longitude is

$$
R_M = \frac{\left[1 + \left(\frac{dz}{dw}\right)^2\right]^{\frac{3}{2}}}{\pm \frac{d^2 z}{dw^2}} \tag{2.138}
$$

The second derivative of Eq. (2.126) provides

$$\frac{dz}{dw} = -\frac{b^2 w}{a^2 z} = -\frac{b^2 w}{a^2}\left(b^2 - \frac{b^2 w^2}{a^2}\right)^{-1/2}$$

$$\frac{d^2 z}{dw^2} = -\frac{b^2}{a^2}\left(b^2 - \frac{b^2 w^2}{a^2}\right)^{-1/2} - \frac{b^2 w}{a^2}\left(-\frac{1}{2}\right)\left(b^2 - \frac{b^2 w^2}{a^2}\right)^{-3/2}\left(-\frac{b^2}{a^2}\right)2w$$

$$\frac{d^2 z}{dw^2} = -\frac{b^2}{a^2 z} - \frac{b^4 w^2}{a^4 z^3}$$

$$\frac{d^2 z}{dw^2} = \frac{-b^2 a^2 z^2 - b^4 w^2}{a^4 z^3} = \frac{-b^2 a^2\left(b^2 - \frac{b^2 w^2}{a^2}\right) - b^4 w^2}{a^4 z^3}$$

$$\frac{d^2 z}{dw^2} = \frac{-b^4 a^2 + b^4 w^2 - b^4 w^2}{a^4 z^3}$$

which simplifies to

$$\frac{d^2 z}{dw^2} = -\frac{b^4}{a^2 z^3} \tag{2.139}$$

Substituting Eqs. (2.139) and (2.126) into (2.138) yields

$$R_M = \frac{\left[1 + \frac{b^4 w^2}{a^4 z^2}\right]^{\frac{3}{2}}}{\frac{b^4}{a^2 z^3}} \tag{2.140}$$

and since $\frac{b^2}{a^2} = 1 - e^2$

$$R_M = \frac{\left[1 + \frac{\left(1-e^2\right)^2 w^2}{z^2}\right]^{\frac{3}{2}}}{\frac{b^2(1-e^2)}{z^3}}$$

$$R_M = \frac{\left[\frac{z^2 + \left(1-e^2\right)^2 w^2}{z^2}\right]^{\frac{3}{2}}}{\frac{b^2(1-e^2)}{z^3}}$$

$$R_M = \frac{\left[z^2 + (1-e^2)^2 w^2\right]^{\frac{3}{2}}}{b^2(1-e^2)} \tag{2.141}$$

Substituting the value of w from Eq. (2.134) and z from (2.135) gives

$$R_M = \frac{\left[\frac{a^2\left(1-e^2\right)^2\sin^2\varphi}{1-e^2\sin^2\varphi} + (1-e^2)^2\frac{a^2\cos^2\varphi}{1-e^2\sin^2\varphi}\right]^{\frac{3}{2}}}{b^2(1-e^2)}$$

$$R_M = \frac{\left[\frac{a^2\left(1-e^2\right)^2\sin^2\varphi+\left(1-e^2\right)^2a^2\cos^2\varphi}{1-e^2\sin^2\varphi}\right]^{\frac{3}{2}}}{b^2(1-e^2)}$$

$$R_M = \frac{\left[a^2(1-e^2)^2\left(\sin^2\varphi+\cos^2\varphi\right)\right]^{\frac{3}{2}}}{b^2(1-e^2)\left(1-e^2\sin^2\varphi\right)^{\frac{3}{2}}}$$

$$R_M = \frac{a^3(1-e^2)^3}{b^2(1-e^2)\left(1-e^2\sin^2\varphi\right)^{\frac{3}{2}}}$$

$$R_M = \frac{a(1-e^2)^3}{(1-e^2)(1-e^2)\left(1-e^2\sin^2\varphi\right)^{\frac{3}{2}}}$$

leading to the meridian radius of curvature

$$R_M = \frac{a(1-e^2)}{\left(1-e^2\sin^2\varphi\right)^{\frac{3}{2}}} \tag{2.142}$$

Appendix B

Derivation of the Conversion Equations From Geodetic to Rectangular Coordinates in the ECEF Frame

Figure 2.17 shows the relationship between geodetic and rectangular coordinates in the ECEF frame.

It is evident that

$$\mathbf{r}_P = \mathbf{r}_Q + h\mathbf{n} \tag{2.143}$$

where

$$\mathbf{r}_Q = \begin{bmatrix} R_N\cos\varphi\cos\lambda \\ R_N\cos\varphi\sin\lambda \\ R_N\sin\varphi - R_Ne^2\sin\varphi \end{bmatrix} = R_N\begin{bmatrix} \cos\varphi\cos\lambda \\ \cos\varphi\sin\lambda \\ \sin\varphi - R_Ne^2\sin\varphi \end{bmatrix} \tag{2.144}$$

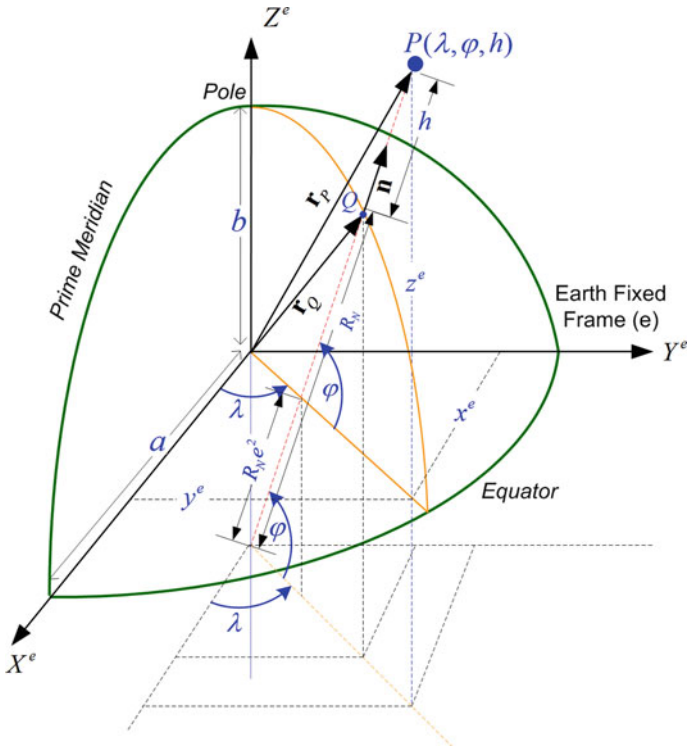


Fig. 2.17 The relationship between geodetic and rectangular coordinates in the ECEF frame

Also, the unit vector along the ellipsoidal normal is

$$\mathbf{n} = \begin{bmatrix} \cos\varphi\cos\lambda \\ \cos\varphi\sin\lambda \\ \sin\varphi \end{bmatrix} \tag{2.145}$$

Substituting Eqs. (2.144) and (2.145) into (2.143) gives

$$\mathbf{r}_P = R_N \begin{bmatrix} \cos\varphi\cos\lambda \\ \cos\varphi\sin\lambda \\ \sin\varphi - R_N e^2 \sin\varphi \end{bmatrix} + h \begin{bmatrix} \cos\varphi\cos\lambda \\ \cos\varphi\sin\lambda \\ \sin\varphi \end{bmatrix} \tag{2.146}$$

$$\mathbf{r}_P = \begin{bmatrix} (R_N + h)\cos\varphi\cos\lambda \\ (R_N + h)\cos\varphi\sin\lambda \\ \{R_N(1 - e^2) + h\}\sin\varphi \end{bmatrix} \tag{2.147}$$

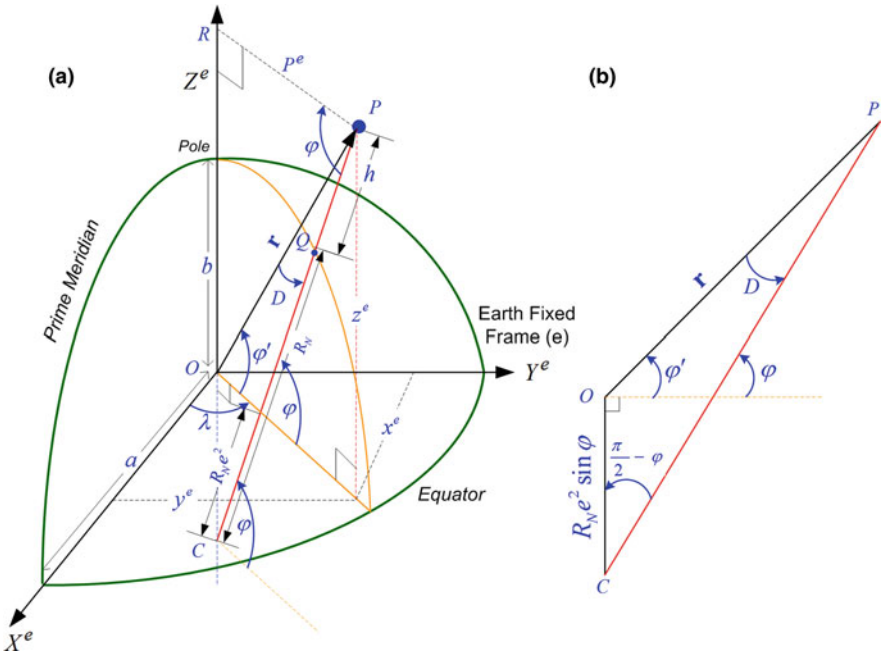


Fig. 2.18 Reference diagram for conversion of rectangular coordinates to geodetic coordinates in the ECEF frame through a closed-form method

$$\begin{bmatrix} x^e \\ y^e \\ z^e \end{bmatrix} = \begin{bmatrix} (R_N + h)\cos\varphi\cos\lambda \\ (R_N + h)\cos\varphi\sin\lambda \\ \{R_N(1-e^2) + h\}\sin\varphi \end{bmatrix} \tag{2.148}$$

Appendix C

Derivation of Closed Form Equations From Rectangular to Geodetic Coordinates in the ECEF Frame

Here we derive a closed form algorithm which uses a series expansion (Schwarz and Wei Jan 1999) to compute e-frame geodetic coordinates directly from e-frame rectangular coordinates.

From Fig. 2.18a it can is evident that, for given rectangular coordinates, the calculation of geocentric coordinates (λ, φ', r) is simply

$$\mathbf{r} = \sqrt{(x^e)^2 + (y^e)^2 + (z^e)^2} \tag{2.149}$$

$$\lambda = \tan^{-1}\left(\frac{y^e}{x^e}\right) \tag{2.150}$$

$$\varphi' = \sin^{-1}\left(\frac{z^e}{r}\right) \tag{2.151}$$

From the triangle POC (elaborated in Fig. 2.18b) it is apparent that the difference between the geocentric latitude φ' and the geodetic latitude φ is

$$\begin{aligned} &D + \varphi' + \frac{\pi}{2} + \left(\frac{\pi}{2} - \varphi\right) = \pi \\ &D + \varphi' + \pi - \varphi = \pi \\ &D = \varphi - \varphi' \end{aligned} \tag{2.152}$$

where D is the angle between the ellipsoidal normal at Q and the normal to the sphere at P.

Applying the law of sines to the triangle in Fig. 2.18b provides

$$\begin{aligned} &\frac{\sin D}{R_N e^2 \sin\varphi} = \frac{\sin\left(\frac{\pi}{2} - \varphi\right)}{r} \\ &\frac{\sin D}{R_N e^2 \sin\varphi} = \frac{\cos\varphi}{r} \\ &\sin D = \frac{R_N e^2 \sin\varphi \cos\varphi}{r} \\ &D = \sin^{-1}\left(\frac{R_N e^2 \sin\varphi \cos\varphi}{r}\right) \end{aligned} \tag{2.153}$$

$$D = \sin^{-1}\left(\frac{R_N e^2 \frac{1}{2}\sin 2\varphi}{r}\right) \tag{2.154}$$

Substituting the definition of the normal radius R_N given in Eq. (2.137)

$$R_N = \frac{a}{\left(1 - e^2 \sin^2\varphi\right)^{1/2}} \tag{2.155}$$

into Eq. (2.154) gives

$$D = \sin^{-1}\left(\frac{\frac{a}{\left(1-e^2\sin^2\varphi\right)^{1/2}} e^2 \frac{1}{2}\sin 2\varphi}{r}\right) \tag{2.156}$$

$$D = \sin^{-1}\left(\frac{k \sin 2\varphi}{\left(1 - e^2 \sin^2\varphi\right)^{1/2}}\right) \tag{2.157}$$

where

$$k = \frac{e^2 a}{2} \tag{2.158}$$

To achieve a first approximation, if $\varphi = \varphi'$ then D can be computed by using the geocentric latitude

$$D_c = \sin^{-1}\left(\frac{k \sin 2\varphi'}{\left(1 - e^2 \sin^2 \varphi'\right)^{1/2}}\right) \tag{2.159}$$

Expanding Eq. (2.157) for this approximation, D_c, gives

$$D = D(\varphi = \varphi') + D'(\varphi)|_{\varphi=\varphi'}(\varphi - \varphi') + D''(\varphi)|_{\varphi=\varphi'}\frac{(\varphi - \varphi')^2}{2!} + D'''(\varphi)|_{\varphi=\varphi'}\frac{(\varphi - \varphi')^3}{3!} + \ldots \tag{2.160}$$

$$D = D_c + D'(\varphi')D + \frac{1}{2!}D''(\varphi')D^2 + \frac{1}{3!}D'''(\varphi')D^3 + \ldots \tag{2.161}$$

For a very small value of D (less than 0.005), it can be assumed that $R_N \simeq r$, and hence

$$D = \frac{e^2}{2}\sin 2\varphi \tag{2.162}$$

So for this situation, let $k = \frac{e^2}{2}$ so that

$$D = k \sin 2\varphi \tag{2.163}$$

The series used above can therefore be truncated after the fourth-order term and be considered as a polynomial equation. Solving this polynomial provides

$$D = \frac{D_c}{1 - 2k \cos 2\varphi' + 2k^2 \sin^2 \varphi'} \tag{2.164}$$

The geodetic latitude can be given as

$$\varphi = \varphi' + D = \sin^{-1}\frac{z^e}{r} + D \tag{2.165}$$

The geodetic longitude is the same as shown earlier

$$\lambda = \tan^{-1}\left(\frac{y^e}{x^e}\right) \tag{2.166}$$

The ellipsoidal height is calculated from the triangle PRC of Fig. 2.18a

$$P^e = (R_N + h)\cos\varphi \tag{2.167}$$

$$h = \frac{P^e}{\cos\varphi} - R_N \tag{2.168}$$

$$h = \frac{\sqrt{(x^e)^2+(y^e)^2}}{\cos\varphi} - R_N \tag{2.169}$$

Equations (2.165), (2.166) and (2.169) are therefore closed-form expressions to convert from rectangular coordinates to geodetic coordinates in the e-frame.

Appendix D

Derivation of the Iterative Equations From Rectangular to Geodetic Coordinates in the ECEF Frame

From Eq. (2.148), which relates geodetic and rectangular coordinates in the e-frame, we see that

$$x^e = (R_N + h)\cos\varphi\cos\lambda \tag{2.170}$$

$$y^e = (R_N + h)\cos\varphi\sin\lambda \tag{2.171}$$

$$z^e = \left\{R_N\left(1 - e^2\right) + h\right\}\sin\varphi \tag{2.172}$$

Given the rectangular coordinates and these equations, the geodetic longitude is

$$\lambda = \tan^{-1}\left(\frac{y^e}{x^e}\right) \tag{2.173}$$

Equation (2.169) specifies the relationship for the altitude as

$$h = \frac{\sqrt{(x^e)^2+(y^e)^2}}{\cos\varphi} - R_N \tag{2.174}$$

Also, from Eqs. (2.170) and (2.171) it can be shown that

$$(x^e)^2+(y^e)^2 = (R_N + h)^2\cos\varphi^2\left[\cos^2\lambda + \sin^2\lambda\right] \tag{2.175}$$

$$\sqrt{(x^e)^2+(y^e)^2} = (R_N + h)\cos\varphi \tag{2.176}$$

And finally, dividing Eq. (2.172) by Eq. (2.176) yields

$$\frac{z^e}{\sqrt{(x^e)^2+(y^e)^2}} = \frac{\left[R_N(1 - e^2) + h\right]}{(R_N + h)}\tan\varphi$$

$$\varphi = \tan^{-1}\left\{\frac{z^e(R_N + h)}{[R_N(1-e^2)+h]\sqrt{(x^e)^2+(y^e)^2}}\right\} \tag{2.177}$$

References

Chatfield AB (1997) Fundamentals of high accuracy inertial navigation. American Institute of Aeronautics and Astronautics, Reston

Farrell JA (1998) The global positioning system and inertial navigation. McGraw-Hill, New York

Farrell JA (2008) Aided navigation : GPS with high rate sensors. McGraw-Hill, New York

Grewal MS, Weill LR, Andrews AP (2007) Global positioning systems, inertial navigation, and integration, 2nd edn. Wiley, New York

Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS-global navigation satellite systems : GPS, GLONASS, Galileo, and more. Springer, New York

Rogers RM (2007) Applied mathematics in integrated navigation systems, 3rd edn. American Institute of Aeronautics and Astronautics, Reston

Schwarz KP, Wei M (Jan 1999) ENGO 623: Partial lecture on INS/GPS integration for geodetic applications University of Calgary, Department of Geomatics Engineering

Titterton D, Weston J (2005) Strapdown inertial navigation technology IEE radar, sonar, navigation and avionics series, 2nd edn. AIAA

Torge W (1980) Geodesy: an introduction. De Gruyter, Berlin

Vanícek P, Krakiwsky EJ (1986) Geodesy, the concepts. North Holland; Sole distributors for the U.S.A. and Canada. Elsevier Science Pub. Co., Amsterdam

Chapter 3
Global Positioning System

The global positioning system (GPS) was developed by the US Department of Defense in the early 1970s to serve military navigational requirements. The first satellite was launched in 1978 and the system was declared operational in 1995. It is based on a network of at least 24 satellites (with room for six further satellites) orbiting the Earth in nearly circular orbits with a mean radius of about 26,560 km. Each satellite transmits a radio signal that contains both a pseudo-random noise (PRN) code and a navigation message (Kaplan and Hegarty 2006). The PRN code is used by the receiver to find the transit time, as a preliminary to calculating the range (called pseudo-range ρ) from the satellite to the receiver by multiplying this by the speed of light. It also calculates the satellite's position from the information in the navigation message. With the information from at least three satellites the receiver can use the process of trilateration to calculate its own position in terms of latitude, longitude and altitude, as shown in Fig. 3.1. The signal from a fourth satellite is needed to cancel the receiver's clock bias (b). A master control station in Colorado Springs monitors the health of the system using information from 12 monitoring stations around the globe, and ensures its accuracy by transmitting data and control signals to the satellites through four ground antennae.

3.1 GPS Observables

The three main GPS observables are pseudo-ranges measurements, carrier phase measurements and Doppler measurements.

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_3,

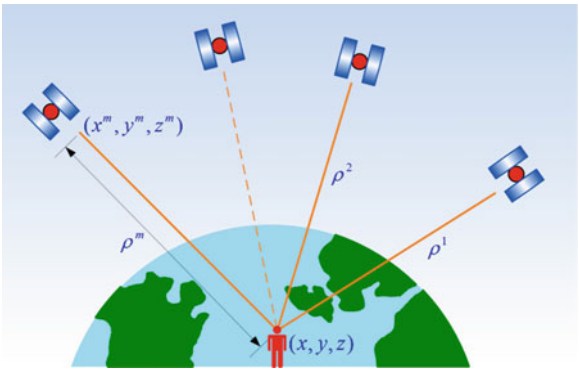


Fig. 3.1 The concept of trilateration which enables a GPS receiver to calculate its position (x,y,z) based on the four range measurements. The coordinates (x^m,y^m,z^m) indicate the position of GPS receiver and ρ^m depicts the pseudo-range between mth satellite and the receiver

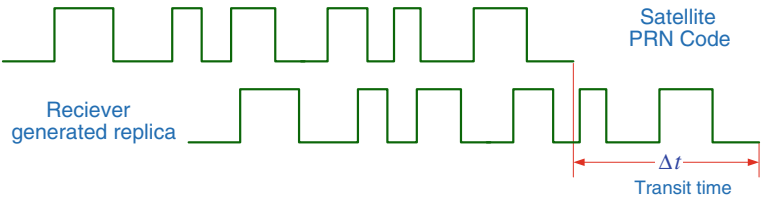


Fig. 3.2 The concept of a pseudo-range calculation based on the PRN code transmitted by a GPS satellite and its receiver-generated replica

3.1.1 Pseudo-Ranges Measurements

The pseudo-range is the measure of the distance between a satellite and a receiver. The transmission time can be calculated from the navigation message, while the reception time is measured from the correlation of the received signal (PRN code) and its receiver-generated replica (Fig. 3.2). The pseudo-range (ρ) is calculated by taking the difference of these times (Δt) and multiplying it with the speed of light (c) as follows

$$\rho = c \times \Delta t \tag{3.1}$$

The pseudo-range measurement for the mth satellite can be written as

$$\rho^m = r^m + c\delta t_r - c\delta t_s + I^m + T^m + \varepsilon_\rho^m \tag{3.2}$$

where

- ρ^m is the measured pseudo-range between the mth satellite and the receiver (meters)
- r^m is the true range between the receiver antenna at time t_r and the satellite antenna at time t_t (meters)
- δt_r is the receiver's clock offset (sec)
- δt_s is the satellite's clock offset (sec)

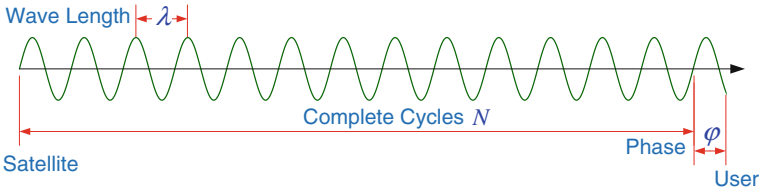


Fig. 3.3 The principle of phase measurements based on complete and partial cycles of the carrier wave

I^m is the ionospheric delay (sec)
T^m is the tropospheric delay (sec)
ε_ρ^m is the is the error in the range due to various sources, including receiver noise, multipath, satellite clock modeling and orbit prediction (meters)

3.1.2 Carrier Phase Measurements

The ranges to the satellites can also be measured through the phases of the carrier signals. The GPS receiver can accurately measure the fractional phase of a cycle but the total number of complete cycles from the satellite to the receiver is initially unknown. This is called integer ambiguity (IA). After the IA is resolved, the range can be calculated by multiplying the carrier wavelength (λ) by the sum of the complete cycles (N) and the fractional one (φ). This range is much more accurate than that calculated from the PRN code. This concept is illustrated in Fig. 3.3, and mathematically expressed as follows

$$\rho = (N + \phi) \times \lambda \tag{3.3}$$

By accounting for clock biases, atmospheric delays, measurement errors and doing some rearrangement, this can be written as

$$\phi = \frac{1}{\lambda}\left(r + c\delta t_r - c\delta t_s - I_\phi + T_\phi\right) + N + \varepsilon_\phi \tag{3.4}$$

where

I_ϕ is the ionospheric delay (meters)
T_ϕ is the tropospheric delay (meters)
δt_r is the receiver's clock offset (sec)
δt_s is the satellite's clock offset (sec)
N is the integer ambiguity (cycles)
ε_ϕ is the error in the measurement (cycles)

The carrier phase measurement can be converted to units of meters by multiplying Eq. (3.4) with the wavelength λ, thus

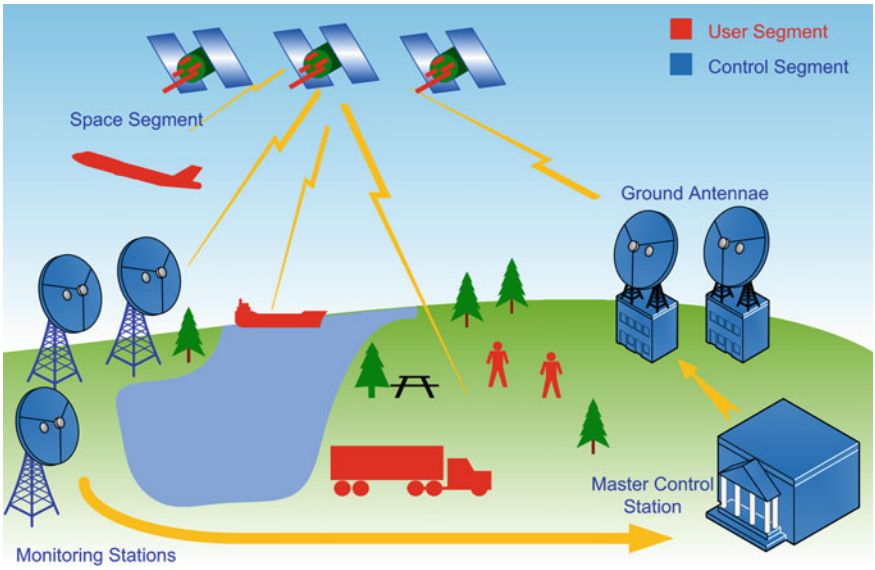


Fig. 3.4 An illustration of the control, user and space segments of GPS

$$\begin{aligned} \lambda\phi &= r + c\delta t_r - c\delta t_s - I_\phi + T_\phi + \lambda N + \varepsilon_{\lambda\phi} \\ \Phi &= r + c\delta t_r - c\delta t_s - I_\phi + T_\phi + \lambda N + \varepsilon_\Phi \end{aligned} \tag{3.5}$$

Since the ionospheric and tropospheric delays are the same for the pseudo-range and the carrier phase (except for the minus sign of the former) we can elide their subscripts

$$\Phi = r + c\delta t - c\delta t_s - I + T + \lambda N + \varepsilon_\Phi \tag{3.6}$$

Note that Eq. (3.6) is very similar to the pseudo-range Eq. (3.2) save for the IA term and the sign of the ionospheric delay.

3.1.3 Doppler Measurements

The Doppler effect is a frequency shift of the electromagnetic signal caused by the relative motion of the emitter and receiver. Based on this phenomenon, some GPS receivers also measure the Doppler frequency of the received signal to calculate the velocity of the user. The Doppler shift (f_d), line-of-sight range rate ($\dot{r}$), and wavelength (λ) of the transmitted signal are related by

$$f_d = -\frac{\dot{r}}{\lambda} \tag{3.7}$$

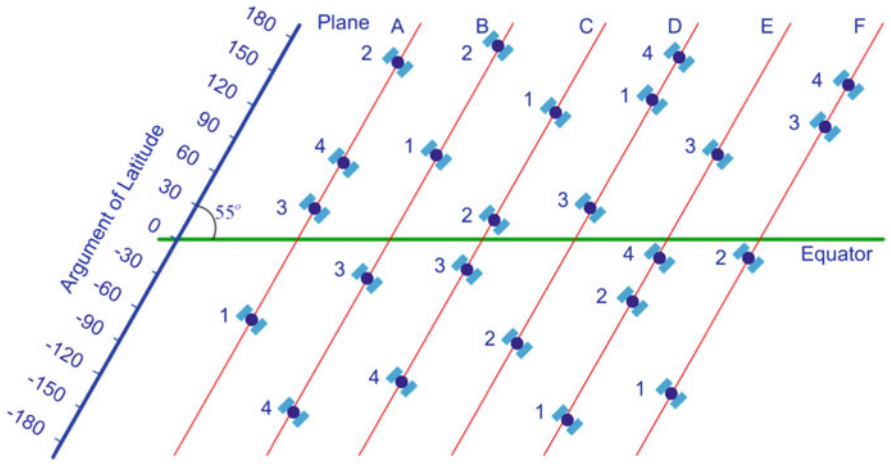


Fig. 3.5 A planar projection of the GPS constellation giving the plane and slot numbers

3.2 GPS Structure

GPS is composed (El-Rabbany 2002) of a space segment, a control segment and a user segment, as shown in Fig. 3.4.

3.2.1 Space Segment

The space segment consists of a constellation of at least 24 satellites orbiting the Earth in nearly circular orbits at an altitude of about 20,000 km. They occupy six orbits inclined at 55° to the equator, each with four primary satellites which are unevenly distributed. The orbital period is about 12 h. Figure 3.5 illustrates the planner projection of GPS satellite constellation. It can accommodate more than 30 satellites to ensure that at any given time at least four satellites are visible to a user with a clear view of the sky. The satellites broadcast radio signals contain coded information and navigation data to enable a receiver to calculate pseudo-ranges and Doppler data to estimate the position and velocity of the user.

3.2.2 Control Segment

The control segment is responsible for the overall control and maintenance of the system, and its responsibilities (Misra and Enge 2001) include

Fig. 3.6 Distribution and location of various elements of the GPS control segment

- Monitoring and maintaining the orbit of each satellite in the constellation by manoeuvring and relocation (if needed).
- Ensuring the performance of the satellites.
- Maintenance of the GPS system time.
- Prediction of the ephemerides and clock parameters of each satellite, and periodic uploading of this information to keep the navigation message up to date.

It consists of the master control station (MCS), the monitoring stations and the ground antennae. Satellite signals are tracked by six US Air Force and 11 national geospatial-intelligence agency (NGA) monitoring stations around the globe. These unmanned stations are controlled by the MCS and they observe satellite orbits and clock integrity. This information is sent to the MCS at Schriever Air Force Base near Colorado Springs, which estimates the ephemeris and clock parameters. The MCS can also relocate a satellite if needed. Ground antennae are situated at four monitoring stations that have radio links with the satellites through the S-band to receive data about their systems, as well as links to the MCS to enable it to upload commands and the latest data for the navigation message to be broadcast by the satellites. Figure 3.6 shows the components and their approximate locations.

3.2.3 User Segment

This segment consists of receivers which receive the radio frequency signals from GPS satellites and estimate their position, velocity and time. Users of GPS can be classified into civilian and military. The user segment has grown many-fold since its inception in the 1980s due to rapid advancement in the technology and the ever decreasing cost of electronics. Early receivers cost at least US$100,000 but now a reasonably accurate receiver can be purchased for under $100. The user segment has permeated many facets of daily life, including location based services, search and rescue, law enforcement, road and maritime travel, remote asset tracking, fleet management, intelligence gathering, sports, public safety, and environmental assessment and planning. Detailed accounts can be found in (Seeber 2003) and in (Parkinson et al. 1996).

3.3 GPS Signals

GPS initially transmitted ranging signals on two frequencies. Later, more signals were added under a modernisation program. This section will give an overview of the traditional signals and then discuss the modernization program.

3.3.1 Traditional GPS Signals

GPS satellites transmit signals on two radio frequencies called L1 and L2, which are centred at 1575.41 and 1227.60 MHz respectively. Each frequency is modulated by ranging codes called pseudo-random noise (PRN) sequence or code for precise range measurements (Langley Jun 1990). A coarse acquisition (C/A) code is associated with the standard positioning service (SPS) set aside for civilian use. The precise code (P-code) is associated with the precise positioning service (PPS). This is further encrypted to the P(Y)-code for security reasons and is for authorized military users. L1 is modulated by both C/A and P-codes, whereas L2 is modulated by P-code only. With the removal of selective availability (SA[1]) in May 2000 both the codes have equal accuracy, which is around 5 to 30 m for single GPS receivers. The receiver must determine the position of the satellite in order to use the navigation message to convert the range measurements into the position and velocity of the user. The navigation message is superimposed on both the L1 and L2 carriers, along with the PRN codes. The message is binary coded data consisting of information on the satellite's health, ephemeris (its position and velocity), clock bias parameters, and an almanac that is a less precise version of the ephemeris data

[1] Selective availability was an intentional degradation in the GPS signal by the US Department of Defense to limit the positional accuracy to 100 m for civilian users.

Table 3.1 New GPS signals and their characteristics

Signal	Satellite launch	Features
L2C (developmental) 2nd civil signal 1227.60 MHz	Began launching in 2005 with GPS Block IIR (M) Available on 24 GPS satellites around 2016	Radio navigation satellite services (RNSS) radio band Modern signal design (CNAV), including multiple message types and forward error correction Binary phase shifted key (BPSK) modulation Includes dedicated channel for codeless tracking
L5 (developmental) 3nd civil signal 1176.45 MHz	Began launching in 2010 with GPS Block IIF Available on 24 GPS satellites around 2019	Highly protected aeronautical radio navigation services (ARNS) radio band Higher transmitted power than L1C/A or L2C Greater bandwidth for improved jam resistance Modern signal design (CNAV), including multiple message types and forward error correction Binary offset carrier (BOC) modulation Includes dedicated channel for codeless tracking
L1C 3nd civil signal 1575.42 MHz	Begins launching in 2014 with GPS Block IIIA Available on 24 GPS satellites around 2021	Aeronautical radio navigation services (ARNS) radio band Designed for international GNSS interoperability Modern signal design (CNAV-2), including forward error correction Multiplexed binary offset carrier (MBOC) modulation

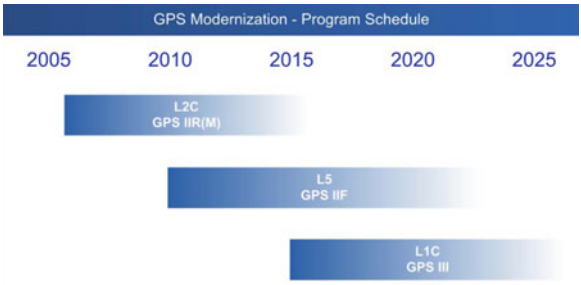


Fig. 3.7 The schedule of the GPS modernization program (Government 2012)

but for all the satellites of the constellation. The ephemeris and clock parameters are repeated every 30 s. The chipping rates of the P-code and C/A code are 10.23 and 1.023 M/sec respectively. The navigation message is broadcasted at the much lower rate of 50 bps and therefore takes 12.5 min to transmit all of the information.

3.3.2 GPS Modernization

The US government is constantly improving the performance of GPS services to maintain leadership in satellite navigation against competition from other satellite navigation systems (Government 2012). Under the GPS modernization program the space and control segments are upgraded to bring new features to improve the system's performance. A major part of this program is to add new signals for both civilian and military users. These new signals are being introduced progressively as new generations of satellites are launched to supersede earlier ones. Currently, three new signals called L2C, L5 and L1C are being introduced for civilian use. The details of the modernization program and the new GPS signals can be found on the website http://www.gps.gov. A summary of the new signals along with their defining features is given in Table 3.1, and the timeline of the GPS modernization program is shown in Fig. 3.7.

3.4 GPS Error Sources

As stated in Chap. 1, a GPS receiver uses the trilateration principle to compute its position by making range measurement to at least four satellites. These ranging measurements are plagued by errors arising from a variety of sources. To obtain an accurate position, the effect of these errors must be mitigated. The next sections give a brief description of the errors that are typically encountered by a receiver in calculating its position.

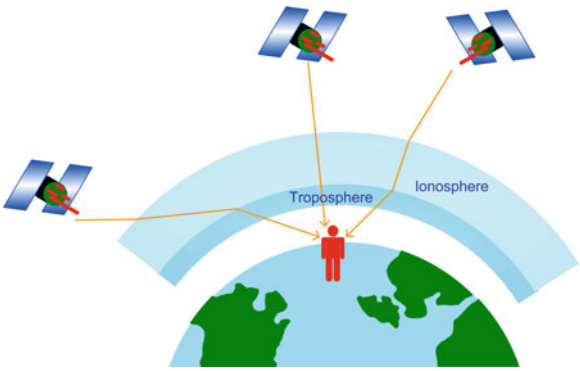


Fig. 3.8 An illustration of the ionospheric and tropospheric effects on GPS signals for various elevation angles

3.4.1 Satellite Clock Error

Although fairly accurate, over time satellite clocks drift away from GPS system time. Based on the observed satellite clock data provided by monitoring stations, the control segment estimates the correction parameters for the satellite clocks and uploads them to the satellites, which broadcast these parameters in the navigation message to enable a receiver to correct for satellite clock error in a measured range.

3.4.2 Receiver Clock Error

Receiver clocks are meant to be inexpensive for affordability. Consequently they are much less accurate than the satellite clocks and contain a bias. This clock bias error affects all measurements in same manner. Therefore if four pseudo-range measurements are available, the clock bias can be estimated along with the three components required to determine the position of the user. This is usually done by a Kalman filter, as detailed in Chap. 7.

3.4.3 Ionosphere Delay

The ionosphere is the layer of the atmosphere which contains ionized gases (free electrons and ions) and occupies the region of space approximately from 60 to 1,000 km above the Earth's surface (Parkinson et al. 1996). The ionization level of this layer changes with solar activity, affecting the refractive indices of the various layers of the ionosphere and, as a result, changing the transit time of a GPS signal

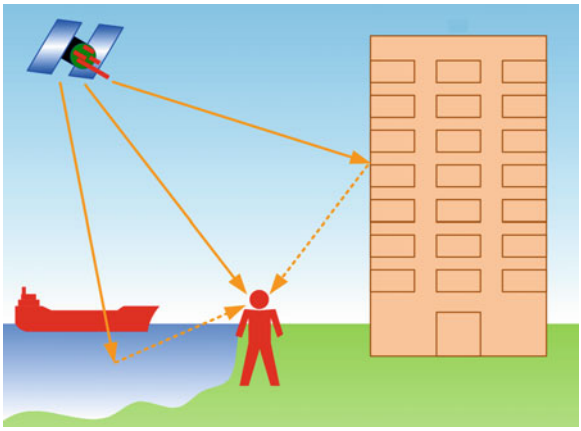


Fig. 3.9 A signal from a GPS satellite can be reflected off water and buildings, affecting both the signal direction and the transition time, degrading the accuracy of the calculated position

(Farrell 2008). Satellite elevation also adds to the variability of this error, because signals from low elevation satellites pass a greater slant-range distance through the ionosphere than those at higher elevations. This fact is graphically illustrated in Fig. 3.8. The ionospheric delay in pseudo-range and carrier phase is equal but opposite in sign [details can be found in (Hofmann-Wellenhof et al. 2008)] and is expressed as

$$I_\phi = -I_\rho = \frac{40.3TEC}{f^2} \tag{3.8}$$

where f is the carrier frequency and TEC stands for total electron count. TEC is defined as the number of electrons in a 1 m^2 cross sectional tube along the path of transmission through the ionosphere, and it varies both temporally and spatially. It is the main quantity to be determined and depends on sunspot activities, seasonal and diurnal variations and line of sight. It can be measured, estimated, modeled or eliminated.

It is evident that measurements at two frequencies

$$I_1 = \frac{40.3TEC}{f_1^2}; I_2 = \frac{40.3TEC}{f_2^2} \tag{3.9}$$

and hence

$$I_1 = \frac{f_2^2}{f_1^2} I_2 \Rightarrow I_1 f_1^2 = f_2^2 I_2 \tag{3.10}$$

Dual frequency GPS receivers equipped with both L1 and L2 are able to calculate ionospheric delay much more accurately. A single frequency receiver relies on the Klobuchar model, whose parameters are broadcast by the satellites. Section 3.7.2 explains how to calculate the ionospheric delay for a single frequency using this model, which can reduce the ionospheric delay error by 50 %.

3.4.4 Tropospheric Delay

The troposphere is the lower part of the atmosphere extending from 8 to 40 km above the Earth's surface and it is mainly composed of dry gases (N_2 and O_2) and water vapor. Unlike the ionosphere, the troposphere, being electrically neutral, is non-dispersive for GPS frequencies but since it is refractive it causes a decrease in speed relative to free space. Therefore apparent ranges appear longer by 2.5–25 m depending upon the elevation angle of the satellite (Parkinson et al. 1996). Tropospheric errors are consistent between L1 and L2 carriers. Tropospheric delay has a dry and a wet component. The wet component is responsible for 10 % of the tropospheric delay and is difficult to model because the water vapor content varies on a local scale. The dry component is better modeled and accounts for 90 % of the tropospheric delay. Several models have been developed for tropospheric delay, including the Saastamoinen model, Hopfield model and Chao model (Parkinson et al. 1996).

3.4.5 Multipath Errors

Multipath is a major error source in urban environments where the GPS signal is able to reach the receiver by several different paths, as shown in Fig. 3.9. These paths include direct line of sight and reflected signals from other objects around the receiving antenna. The signal that arrives indirectly is delayed and has a lower signal to noise ratio. Multipath distorts the original signal because of interference with the reflected signals at the receiving antenna. This can cause a position error in excess of 10 m. For a detailed treatment refer to (Parkinson et al. 1996; Kaplan and Hegarty 2006; Misra and Enge 2001). The multipath error is two orders of magnitude lower for carrier phase measurements than for pseudo-range measurements.

3.4.6 Satellite Orbital Errors

Satellite orbital errors are the difference between the actual position of a satellite in space and the position of the satellite calculated by the receiver using ephemeris data. Depending on the previous motion of the satellite and knowledge of Earth's gravity, the orbital errors are predicted by the control segment and uploaded to the satellites for broadcast to the users as ephemeris data. Since the ephemeris model is a curve fit to the measured orbit, it will include time varying residual errors relative to the actual orbit. Typically, this error is between 2 and 5 m.

3.4.7 Receiver Noise

This is a random measurement noise intrinsic to the electronics of a GPS receiver. It is caused by the cumulative effects of antenna circuitry, cables, thermal noise, RF signal interference, signal quantization and sampling. Since it is a function of the signal to noise ratio, it varies with the elevation angle of a satellite. It gives rise to an incorrect measurement of the transit time of the GPS signal. As in the case of the multipath effect, receiver noise is two orders of magnitude lower for carrier phase measurements than for pseudo-range measurements.

3.4.8 User Equivalent Range Error

The combined effect of the residual errors (after mitigation through appropriate models and the data in the navigation message) on pseudo-range measurements is called user equivalent range error (UERE). This is made under the assumption that error sources can be allocated to individual satellite pseudo-ranges, and viewed as affecting the pseudo-range values equivalently (Misra and Enge 2001). Assuming that these errors are uncorrelated, a reasonable assumption, the composite UERE for a satellite is approximated as a zero-mean Gaussian random variable that has a variance equal to the root-sum-square of the individual component variances

$$\sigma_{UERE} = \sqrt{\sigma_{eph}^2 + \sigma_{clk}^2 + \sigma_{ion}^2 + \sigma_{tro}^2 + \sigma_{mlt}^2 + \sigma_{rcv}^2} \tag{3.11}$$

where

σ_{eph} is the range error due to ephemeris data
σ_{clk} is the range error due to the satellite's clock
σ_{ion} is the range error due to the ionosphere
σ_{tro} is the range error due to the troposphere
σ_{mlt} is the range error due to multipath
σ_{rcv} is the range error due to the receiver measurement

The typical UERE for the above mentioned errors is about 5.3 m. For a detailed breakdown of the various range errors see (Parkinson et al. 1996).

3.5 GPS Augmentation

Currently, the typical accuracy of a standalone GPS (i.e. SPS) is around 10 m horizontally and 15 m vertically. Although this is adequate for many common navigational applications, a greater accuracy is required for critical safety-of-life applications, and to achieve this GPS is augmented with various other systems. The requirements for critical safety-of-life applications are as follows:

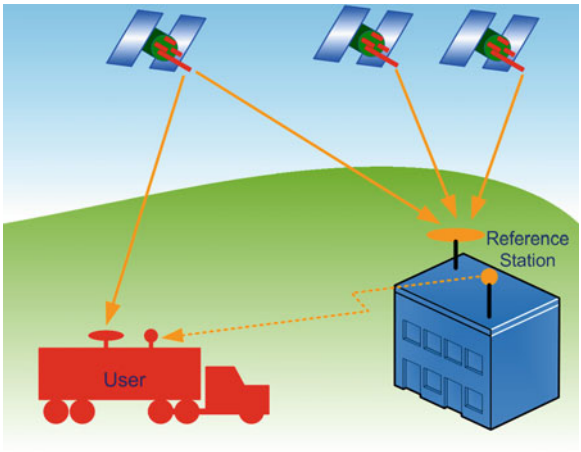


Fig. 3.10 The concept of local area differential GPS (LADGPS)

- **Integrity**: The ability to provide timely warnings to users that the system should not be used for navigation or other purposes.
- **Accuracy**: The difference between a GPS-measured position at any given time and the true position.
- **Continuity**: The ability to complete an operation without triggering an alarm.
- **Availability**: The ability of a system to be used for navigation whenever and wherever the user desires.

Various augmentation systems are employed to achieve these requirements, and they are mainly based on the following techniques

- **Addition of extra sensors**: In this method additional sensors are used to complement the GPS information and improve the overall navigational information; these are typically an altimeter, a compass, accelerometers, gyroscopes and an odometer.
- **Use of differential GPS corrections**: Some of the GPS errors are similar for users who are relatively near to one another, and they do not change quickly. These errors are said to spatially and temporally correlated. They can be estimated by a receiver whose location is already known (it being called a base or reference station). These errors are conveyed to nearby GPS users for use by a compensation algorithm. This technique is known as differential GPS (DGPS).
- **Use of pseudo-satellites**: Pseudolites (as they are called) are transmitters on the ground that provide additional GPS-like signals in areas where the reception of satellite signals is either below the minimum level necessary or is completely blocked. In cases where GPS satellite geometry is poor, pseudolites can artificially improve the geometry by supplying signals at better angles.
- **Assisted GPS**: GPS is increasingly being incorporated into cell phones, which require extra information to give location information in a timely manner. This

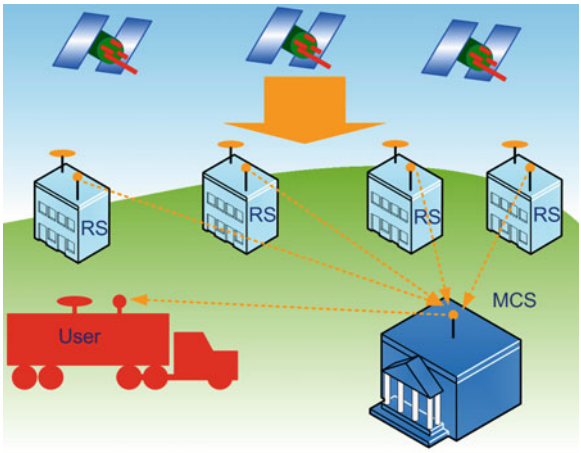


Fig. 3.11 The concept of wide area differential GPS (WADGPS)

information is provided through the cellular networks in the form of assisted GPS (A-GPS).

Further details of DGPS and A-GPS will be provided in subsequent sections of this chapter and the addition of extra sensors will be discussed in Chaps. 8 and 9.

3.5.1 Differential GPS

DGPS is broadly categorized into two techniques based on the area in which they can mitigate GPS errors, the local area DGPS (LADGPS) and the wide area DGPS (WADGPS).

3.5.2 Local Area DGPS

In LADGPS, a single reference station (RS) serves the users in a relatively small area. It calculates the pseudo-ranges with code-phase measurements and, based on accurate knowledge of its position, determines the bias in the measurements for all the visible satellites. These biases are calculated as the difference in pseudo-range measurements made by the RS and the geometric range between the RS and the satellite. These bias errors (differential corrections) are transmitted to the users in the area on a radio link for real-time applications. These errors include ephemeris prediction errors, uncorrelated satellite perturbations, and errors introduced by the atmosphere (Prasad and Ruggieri 2005). These corrections are more accurate for users who are closer to the RS than for those who are farther away. The concept of LADGPS is shown in Fig. 3.10.

Table 3.2 A summary of world GPS augmentation systems (Prasad and Ruggieri 2005; GPS Augmentations 2009; McPherson 2001; Crosby et al. 2000)

Augmentation system	Association	Type
NGDGPS: Nationwide differential GPS system	Federal railroad administration, federal highway administration, and US coast guard	SBAS
WAAS: Wide area augmentation system	US Federal aviation administration	SBAS
CWAAS: Canadian wide area augmentation system	Canada	SBAS
EGNOS: Geostationary navigation overlay system	European commission, European space agency, and Eurocontrol	SBAS
GAGAN: Geo-augmented navigation system	Indian Space Research Organization and the Airports Authority of India	SBAS
QZSS: Quasi-zenith satellite system	Japan (under development)	SBAS
MSAS: Multifunction transport satellite (MTSAT) satellite augmentation system	Civil aviation Bureau, Ministry of Land, Infrastructure and Transport, Government of Japan	SBAS
SNAS: Satellite navigation augmentation system	China	SBAS
CORS: Continuously operating reference station	National Oceanic & Atmospheric Administration(for precision positioning and atmospheric modeling)	GBAS
GDPS: Global differential GPS	Jet Propulsion Laboratory, NASA	GBAS
IGS: International GNSS service	Joint effort of 200 organization in 80 countries (for GNSS)	GBAS
LAAS: Local area augmentation System	Federal Aviation Administration, USA	GBAS
GRAS : Ground-based regional augmentation system	Australia	GBAS
OmniSTAR:	USA	SBAS

3.5.3 Wide Area DGPS

As the GPS receiver moves farther away from the associated RS, the correlation between errors reduces and they become spatially correlated, with the result that the errors estimated at the RS can become different from those experienced by the user. This situation can be ameliorated by expanding the coverage by adding more reference stations along the perimeter of the area covered by the single RS. The receiver weights the corrections based on its proximity to each individual RS. This method of differential correction, called wide area DGPS (WADGPS), is shown in Fig. 3.11.

WADGPS has a master (control) station (MS) and some number of reference stations. Each RS transfers its measurement data to the MS, which estimates the GPS error components based on the data received and *a priori* knowledge of the position of that particular RS. These errors are then sent to all the users in the area either by radio or by satellite links. The receiver uses a weighted average of these corrections to correct its estimated position.

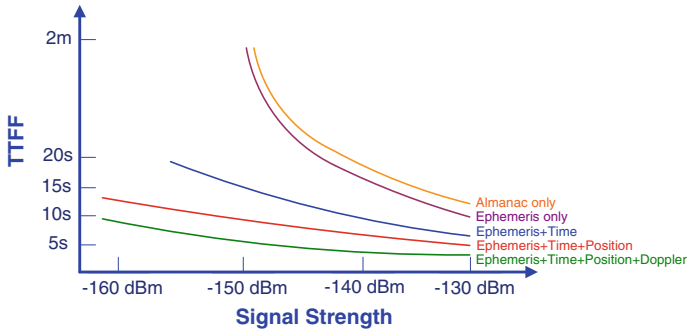


Fig. 3.12 Time to first fix (TIFF) for various types of aiding data versus the strength of the signal [adapted with permission from (Essentials of Satellite Navigation 2007)]

The DGPS can be either space-based or ground-based, differing in the way the information is transmitted to the user

- **Space-based augmentation systems** (**SBAS**): In this case the differential GPS corrections are broadcast via geostationary satellites using the same frequency band as the core GPS satellites.
- **Ground-based augmentation systems** (**GBAS**): The differential GPS corrections are provided to the users by transmitters on the ground. The utilization of pseudolites is sometimes considered to fall in this category.

There are many GPS augmentations operating with the intention of improving the integrity, accuracy, continuity and availability in various regions of the world. The detail of these is beyond the scope of this book, but the reader is referred to (Prasad and Ruggieri 2005; Kaplan and Hegarty 2006). A summary of such systems is provided in Table 3.2.

3.5.4 Assisted GPS

Cellular networks are expanding rapidly, and many mobile phones incorporate a GPS. However, in order to limit the power consumption the GPS cannot always be on. Two or more hours of inactivity requires obtaining fresh satellite orbital data before a positional fix can be made. This can take several minutes because, in the absence of orbital data, the GPS receiver must carry out a complete search to find the available satellites and then download the data prior to calculating its position. The time required to achieve this is called time-to-first-fix (TTFF). Since a longer TTFF is unacceptable for emergency services, A-GPS was proposed as a means of solving this problem.

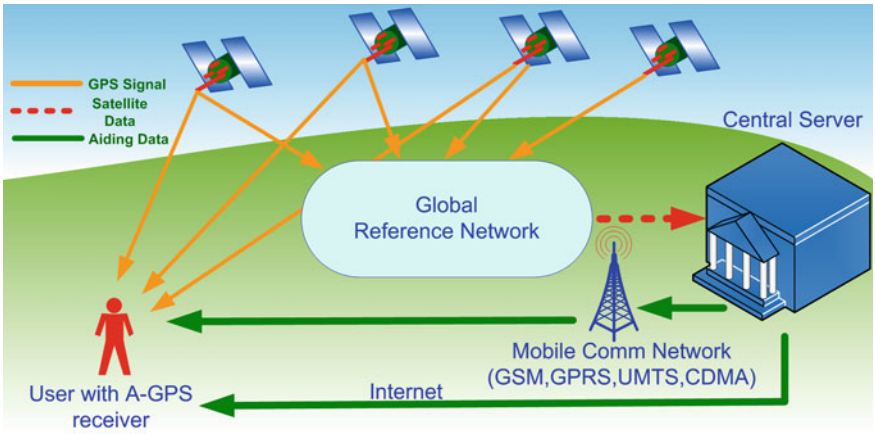


Fig. 3.13 The concept of an assisted GPS network

3.5.4.1 A-GPS Aiding Data

In A-GPS, position can be determined quickly by providing aiding data to GPS though either mobile communication channels or the Internet (Prasad and Ruggieri 2005; GPS Augmentations 2009; McPherson 2001; Crosby et al. 2000). This data usually includes

- The almanac (for the satellite constellation).
- The ephemeris data (accurate orbital information).
- GPS time information.
- Information about the Doppler frequency and the frequency offset of the GPS receiver.

Generally, the greater the availability and the accuracy of the aiding data, the faster is the startup. This concept is depicted in Fig. 3.12, which also shows that TTFF decreases with signal strength.

3.5.4.2 An A-GPS Network

A typical A-GPS network comprises following segments

1. Networks of reference GPS receivers.
2. A central server.
3. Communications media (e.g. radio or Internet)
4. A receiver capable of processing A-GPS data.

The global reference network provides the relevant satellite information to the server that calculates the aiding data which will be provided to GPS devices upon

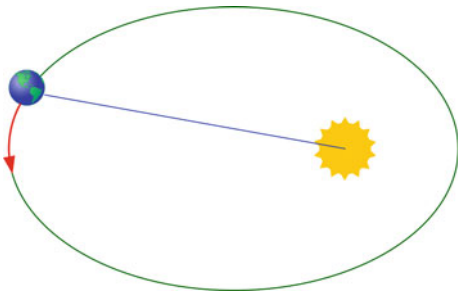


Fig. 3.14 Graphical depiction of Kepler's first law

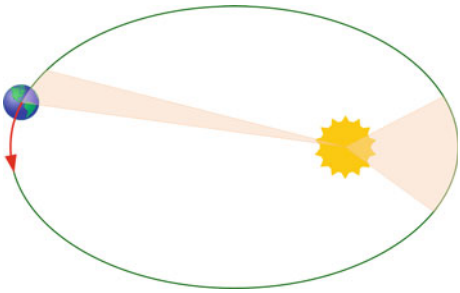


Fig. 3.15 Graphical depiction of Kepler's second law

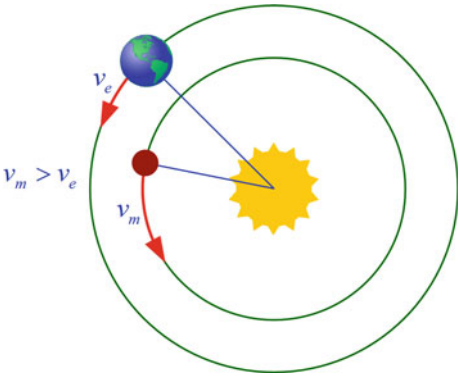


Fig. 3.16 Graphical depiction of Kepler's third law

request by way of the communications media. Using this aiding data, the receiver can quickly calculate its position. Figure 3.13 illustrates a typical A-GPS network. For a more detailed explanation the reader should consult (Essentials of Satellite Navigation 2007).

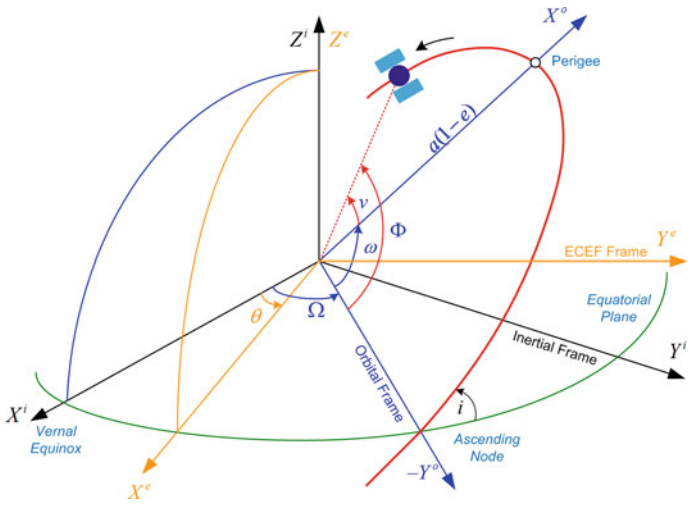


Fig. 3.17 Elements of a Keplerian orbit relative to the ECEF, ECI and orbital frames

3.6 GPS Satellite Orbits

The following subsections will describe the parameters used to define the orbit of a GPS satellite. Before going into mathematical details, it is important to know the basics of Kepler's three laws of planetary motion.

3.6.1 Kepler's Laws

Based on extensive astronomical observations by Tycho Brahe, Johannes Kepler formulated three famous laws of planetary motion.

1. A planet travels along an elliptical orbit with the sun situated at one of the foci of the ellipse (Fig. 3.14).
2. A line between the sun and a planet sweeps out equal areas in equal times (Fig. 3.15).
3. The square of the period of revolution of a planet is proportional to the cube of its mean distances from the sun (Fig. 3.16).

The motion of a satellite could therefore be characterized by a fixed elliptical orbit in space with the Earth being at one of the foci. This orbit can be specified by the six elements of the satellite's position and velocity vectors at a specific epoch.

An alternative representation uses the six Keplerian elements that are described in the next section (Diggelen 2009).

3.6.2 Keplerian Orbital Elements

The Keplerian orbital elements depicted in Fig. 3.17 are defined as follows

a. *Semimajor axis a* : This runs from the center through the focus to the edge of the ellipse, and so is one half of the longest diameter

b. *Eccentricity e* : A measure of *ovalness* of the ellipse

} *Specifies the size and shape of an elliptical*

c. *Inclination i* : The angle of the orbital plane relative to the Earth's equatorial plane

d. *Right ascension of the ascending node* (*RAAN*) Ω : The angle in the equatorial plane between the ascending node of the satellite's orbit and a reference direction in space called the vernal equinox

} *Specifies the orientation of the orbital plane*

e. *Argument of perigee* ω : The angle in the plane of the orbit between the ascending node and the perigee (the point of the orbit that is closest to the center of the Earth)

} *Specifies the orientation of the ellipse in orbital plane*

f. *True anomaly v* : The angle between the perigee and the satellite at any particular moment. The sum of the true anomaly and the argument of perigee is equal to another parameter called the *argument of latitude*, Φ, which is given by the expression

} *Specifies the position of the satellite in the orbit*

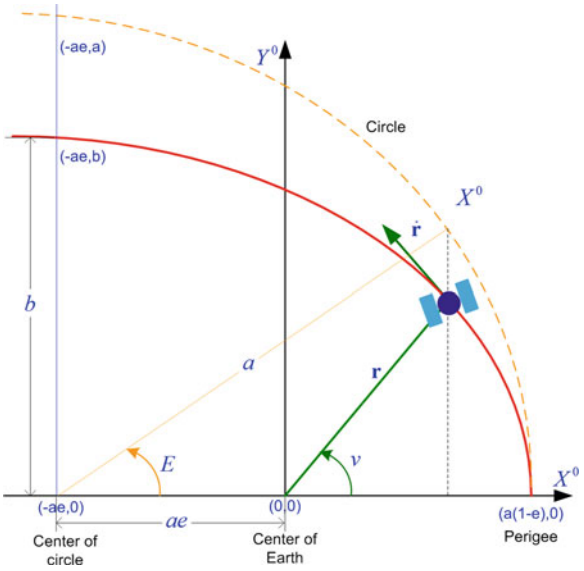


Fig. 3.18 The eccentric anomaly and the true anomaly in the orbital coordinate system

$$\Phi = \omega + v \tag{3.12}$$

The true anomaly specifies the location of a satellite (denoted by r in Fig. 3.18) in its orbit. It does not vary linearly with time for noncircular orbits. To get a parameter that varies linearly with time, two further definitions are required which transform the true anomaly to the mean anomaly that does vary linearly with time (Bate et al. 1971). These two parameters are called the *eccentric anomaly* and the *mean anomaly* (see Fig. 3.18) and are defined as follows

a. ***Eccentric anomaly***,E: The angle subtended at the center of the orbit between the perigee and the projection of the satellite onto a circle of radius a. The true anomaly is converted to the eccentric anomaly using the following relationship (Kaplan and Hegarty 2006)

$$E = 2\tan^{-1}\left[\sqrt{\frac{1-e}{1+e}}\tan\left(\frac{1}{2}\right)v\right] \tag{3.13}$$

b. ***Mean anomaly***,M: The angle between the perigee and an imaginary satellite that travels in a circular orbit that has the same focus and period as the actual satellite but does so with a constant speed (called the *mean motion*). When the eccentric anomaly has been calculated from Eq. (3.13) the mean anomaly can be calculated by Kepler's equations as

$$M = E - e\sin E \tag{3.14}$$

Table 3.3 Ephemeris parameters of GPS navigation message (Battin 1987)

t_{oe}	Ephemeris reference time (sec) is nominally at the center of the time interval over which ephemeris is useful
$\sqrt{a}$	Square root of the semimajor axis ($\sqrt{m}$)
e	Eccentricity (dimensionless)
i_0	Inclination angle at the reference time (semicircles)
Ω_0	Longitude of the ascending node of the orbital plane at the beginning of the GPS week (semicircles)
ω	Argument of perigee (semicircles)
M_0	Mean anomaly at the reference time (semicircles)
Δn	Correction to the computed mean motion (semicircles/sec)
$\dot{i}(IDOT)$	Rate of change of the inclination angle (semicircles/sec)
$\dot{\Omega}$	Rate of change of the RAAN with time (semicircle/sec)
C_{uc}, C_{us}	Amplitudes of the cosine and sine harmonic correction terms for the computed argument of latitude (radians)
C_{rc}, C_{rs}	Amplitudes of the cosine and sine harmonic correction terms for the computed orbit radius (meters)
C_{ic}, C_{is}	Amplitudes of the cosine and sine harmonic correction terms for the computed inclination angles (radians)

3.6.3 GPS Orbital Parameters

Keplerian elements describe the position of a satellite for the ideal conditions of a spherical Earth of uniform composition (and therefore a spherically symmetrical gravitational field known as the *central field*), taking into account the gravitational force of the Earth only. However, the Earth is not a uniform sphere and as shown in Fig. 3.19 there are perturbing forces acting on the satellite apart from gravity. These forces are

a. A non-central gravitational force field.
b. Gravitational fields of the sun, moon and the planets.
c. Solar radiation pressure and atmospheric drag.

GPS accounts for these perturbations with an expanded set of 12 parameters that look similar to Keplerian parameters and are specified relative to a reference epoch. These parameters, referred to as the *broadcast ephemerides*, are estimated using a least squares curve fit to the orbit based on 4 h of data. The broadcast ephemerides are uploaded to the satellites typically once per day and are good for a fortnight. Satellites broadcast this ephemeris data as part of the GPS navigation message every 30 s. The first six parameters listed in Table 3.3 describe a smooth elliptical orbit along with the mean motion, n, with the satellite's motion as a function of time since t_{oe}. The other parameters describe the deviation of the satellite's actual motion from the smooth ellipse, and account for the perturbation effects.

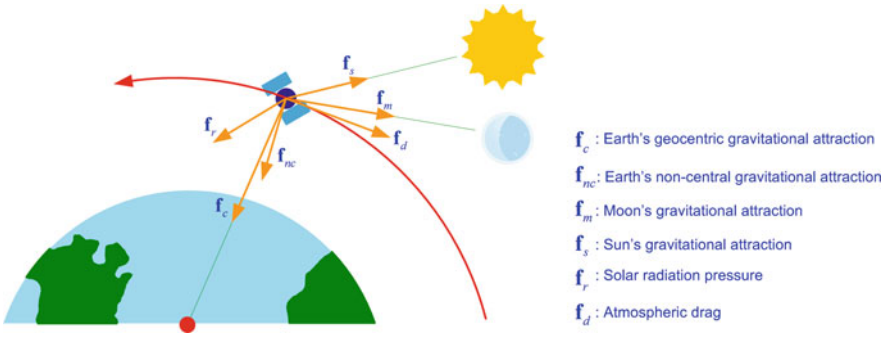


Fig. 3.19 Forces acting on a satellite

3.7 Ephemeris Data Processing

The pseudo-range measured by a GPS receiver must be corrected for a variety of errors to become accurate and useful. We also need to know the position of the satellite in order to use the corrected pseudo-ranges from a number of satellites to estimate the position of the receiver. This requires estimating the effects of these errors and calculating the position of the satellites in the e-frame. The position of a satellite is determined based on the orbital parameters broadcast by the satellite in the ephemeris data of the navigation message. These parameters are predicted by the master control station on the basis of measurements made by the monitoring stations 24 to 48 h earlier. The following sections will discuss the calculations required to find the effects of the various errors and the position and velocity of a satellite at the time of message transmission. The reader is referred to (IS-GPS-200F 2011) for more details.

3.7.1 Calculation of Satellite Clock Corrections

In GPS it is essential to have a very accurate time, as 1 microsecond of error could result in 300 m of range error. So it is very important to calculate it precisely, taking into account every possible source of error. A GPS receiver, upon receiving a signal from a satellite, time stamps the pseudo-range measurements. Hence the transit time of the signal as

$$t_t = \frac{\rho}{c} \tag{3.15}$$

where

t_t is the satellite to receiver transit time (sec)
ρ is the pseudo-range observation (meters)
c is the speed of light (meters/sec)

Then the nominal time at which the satellite sent the signal can be given as

$$t_{sv} = t_m - t_t \tag{3.16}$$

where

t_{sv} is the raw transmission time of the satellite's signal (sec)
t_m is the time at which the measurement was made by the receiver (sec)

The satellite clock offset can be calculated from the polynomial coefficients found in the GPS navigation message as

$$\Delta t_{sv} = a_{f0} + a_{f1}(t - t_{oc}) + a_{f2}(t - t_{oc})^2 \tag{3.17}$$

where

Δt_{sv} is the satellite clock correction (sec)
a_{f0} is the satellite clock offset (sec)
a_{f1} is the fractional frequency offset (sec/sec)
a_{f2} is the fractional frequency drift (sec/sec^2)
t_{oc} is the clock data reference time (sec).

} Contained in navigation message

In Eq. (3.17) t_{sv}can be substituted for t as it is not available yet (IS-GPS-200F 2011). Also, the quantity $(t - t_{oc})$ should be corrected for the end of week crossover as follows

$$\begin{aligned}
&\textit{if } [(t - t_{oc}) > 302,400] \\
&\qquad (t - t_{oc}) = (t - t_{oc}) - 604,800 \\
&\textit{elseif } [(t - t_{oc}) < -302,400] \\
&\qquad (t - t_{oc}) = (t - t_{oc}) + 604,800 \\
&\textit{endif}
\end{aligned}$$

The corrected time at which the signal was transmitted can be computed using the abovementioned clock correction as follows

$$t = t_{sv} - \Delta t_{sv} - \Delta t_r \tag{3.18}$$

where Δt_r, the relativistic correction, can be calculated as

$$\Delta t_r = -4.442807633 \times 10^{-10} e\sqrt{a}\sin E_k \text{ (sec)} \tag{3.19}$$

where

e is the eccentricity of the satellite's orbit
a is the semimajor axis
E_k is the eccentric anomaly.

Defining the satellite clock offset as

$$\delta t_s = \Delta t_{sv} + \Delta t_r \tag{3.20}$$

the final correction to pseudo-range can be applied as

$$\rho_{corrected} = \rho_{measured} + c\delta t_s \tag{3.21}$$

3.7.2 Atmospheric Corrections

3.7.2.1 Ionospheric Modeling

For single-frequency users, an empirical model is used for ionospheric effects and the parameters of this model are broadcast by the satellites. This model was developed by Mr. Klobuchar, and named after him (Farrell 1998). The Klobuchar model typically reduces the RMS errors due to the ionospheric propagation effect by at least 50 % (Klobuchar May 1987). The broadcast parameters are based on the best fit of the diurnal maximum values of the monthly average TEC.

Using variables defined in Table 3.4, a step by step procedure to calculate the ionospheric correction using the Klobuchar model (2011) is now given

a. Compute the Earth-centred latitude (semicircles)

$$\Psi = \frac{0.0137}{El + 0.11} - 0.022 \tag{3.22}$$

where El is the satellite's elevation angle.

b. Calculate the sub-ionospheric latitude φ_i (semicircles)

$$\varphi_i = \varphi_u + \Psi \cos A \tag{3.23}$$

$$\varphi_i = \begin{cases} 0.416 & if |\varphi_i| > +0.416 \\ -0.416 & if |\varphi_i| < -0.416 \end{cases} \tag{3.24}$$

where φ_u and $A$are the user's latitude and azimuth angle respectively.

c. Calculate the sub-ionospheric longitude λ_i (semicircles)

$$\lambda_i = \lambda_u + \frac{\Psi \sin A}{\cos \varphi_i} \tag{3.25}$$

Table 3.4 Variable definitions for ionosphere correction (IS-GPS-200F 2011)

Satellite transmitted terms	
α_n	Coefficients of a cubic equation representing the amplitude of the vertical delay
$\sqrt{2}$	Coefficients of a cubic equation representing the period of the model
Receiver generated terms	
El	Elevation angle between the user and the satellite (semicircles)
A	Azimuth angle between the user and the satellite (semicircles)
φ_u	The user's geodetic latitude (semicircles)
λ_u	The user's geodetic longitude (semicircles)
t_{GPS}	GPS time (sec)
Computed terms	
x	Phase (radians)
F	Obliquity factor (dimensionless)
t	Local time (sec)
φ_m	Geomagnetic latitude of the Earth projection of the ionospheric intersection point (semicircles)
λ_i	Geodetic longitude of the Earth projection of the ionospheric intersection point (semicircles)
φ_i	Geodetic latitude of the Earth projection of the ionospheric intersection point (semicircles)
Ψ	Earth's central angle between the user's position and the Earth projection of the ionospheric intersection point (semicircles)

where λ_u is the user's longitude.

d. Calculate the geomagnetic latitude (semicircles)

$$\varphi_m = \varphi_i + 0.064\cos(\lambda_i - 1.617) \tag{3.26}$$

e. Calculate the local time at the sub-ionospheric point (sec)

$$t = 4.32 \times 10^4 \lambda_i + t_{GPS} \tag{3.27}$$

and because t should be between 0 and 86,400 the following check must be performed

$$\begin{aligned} &if\,(t \geq 86,400) \\ &\qquad t = t - 86,400 \\ &elseif\ (t < 0) \\ &\qquad t = t + 86,400 \\ &endif \end{aligned}$$

f. Calculate the period of the model using the four coefficients β (provided in the navigation message) for the cubic equation (sec)

$$PER = \begin{cases} \sum_{n=0}^{3} \beta_n \varphi_m^n & if\ PER \geq 72,000 \\ 72,000 & if\ PER < 72,000 \end{cases} \tag{3.28}$$

g. Compute the argument x as

$$x = \frac{2\pi(t - 50,400)}{PER} \tag{3.29}$$

h. Calculate the obliquity factor to account for the increased delay at lower elevations

$$F = 1.0 + 16.0(0.53 - El)^3 \tag{3.30}$$

i. Calculate the amplitude of the vertical delay using the four coefficients α (provided in navigation message) for the cubic equation (sec)

$$AMP = \begin{cases} \sum_{n=0}^{3} \alpha_n \varphi_m^n & if\ \text{AMP} \geq 0 \\ 0 & if\ \text{AMP} < 0 \end{cases} \tag{3.31}$$

j. Finally, compute the ionospheric correction term I (sec)

$$I = \begin{cases} F\left\{5 \times 10^{-9} + AMP\left(1 - \frac{x^2}{2} + \frac{x^4}{24}\right)\right\} & if\ |x| < 1.57 \\ 5F \times 10^{-9} & f\ |x| \geq 1.57 \end{cases} \tag{3.32}$$

This correction can be applied to the pseudo-range as

$$\rho_{corrected} = \rho_{measured} - cI \tag{3.33}$$

3.7.2.2 Tropospheric Modeling

Estimation of tropospheric delay[2] requires knowledge of temperature, pressure and water content in the atmosphere around the location of the receiver. Since real-time meteorological measurements are not always feasible for low cost navigation applications, average meteorological conditions are assumed for estimation of the tropospheric delay based on the standard atmosphere for the day of the year and the user's general location. There are various models to determine the tropospheric delays which differ in assumptions about temperature and water vapour variations with altitude.

The delay from zenith direction (zenith delayt_{tro}^z) can be expressed in terms of corresponding dry delays $t_{tro,d}^z$ and wet delays $t_{tro,w}^z$

$$t_{tro}^z = t_{tro,d}^z + t_{tro,w}^z \tag{3.34}$$

Then an obliquity factor is included to scale the zenith delay as a function of elevation angle El, giving the tropospheric delay T as

$$T = t_{tro,d}^z . OF_d(El) + t_{tro,w}^z . OF_w(El) \tag{3.35}$$

[2] Tropospheric correction does not depend upon the ephemeris supplied data, but it is discussed in this subsection for the sake of continuity and completeness.

There are various models in use for the wet delay. For our purpose we employ the Hopfield model, for which dry and wet delays are expressed (IS-GPS-200F 2011) as

$$t^z_{trop,d} = 77.6 \times 10^{-6} \frac{P_0 \times h_d}{T_0 \times 5} \tag{3.36}$$

$$t^z_{trop,w} = 0.373 \frac{e_0 h_w}{T_0^2 \times 5} \tag{3.37}$$

where P_0 is the total pressure (mbar), T_0 is the temperature (Kelvin) and e_0 is the partial pressure (mbar) due to water vapour at the location of the GPS antenna. However, for navigation applications these values are picked as constants from the model of the standard atmosphere. In the above equation h_d and h_w are the mean values of the height constants and are 43 and 12 km respectively.

Because these delays are estimated for the zenith direction, an obliquity factor (also known as a mapping function) is needed to estimate the slant delay from the zenith delay. Of the mapping functions available, the simplest is $1/\sin(El)$ for higher angles. However, general mapping functions (including all angles) for dry and wet components (Hopfield 1969) are

$$OF_d(El) = \frac{1}{\sin El + \frac{0.00143}{\tan El + 0.0445}} \tag{3.38}$$

$$OF_w(El) = \frac{1}{\sin El + \frac{0.00035}{\tan El + 0.017}} \tag{3.39}$$

Thus Eq. (3.35) becomes

$$\begin{aligned} T = {} & 77.6 \times 10^{-6} \frac{P_0 \times h_d}{T_0 \times 5} \left(\frac{1}{\sin El + \frac{0.00143}{\tan El + 0.0445}} \right) \\ & + 0.373 \frac{e_0 h_w}{T_0^2 \times 5} \left(\frac{1}{\sin El + \frac{0.00035}{\tan El + 0.017}} \right) \end{aligned} \tag{3.40}$$

This correction can be applied to pseudo-range as

$$\rho_{corrected} = \rho_{measured} - cT \tag{3.41}$$

By applying all of the aforementioned corrections, the pseudo-range becomes

$$\rho_{corrected} = \rho_{measured} + c\delta t_s - cI - cT \tag{3.42}$$

3.7.3 Calculation of Satellite Position

Using parameters defined in Table 3.3, the following procedure (Chao 1971) will compute the position of a satellite in ECEF coordinates

a. Calculate the semimajor axis of the elliptical orbit of the satellite

$$a = \left(\sqrt{a}\right)^2 \tag{3.43}$$

where $\sqrt{a}$ is the square root of the semimajor axis specified in the ephemeris data.

b. Calculate the mean motion of the satellite

$$n_0 = \sqrt{\frac{\mu}{a^3}} \tag{3.44}$$

where μ is the Earth's gravitational constant $= 3.986005 \times 10^{14} m^3/s^2$

c. Find time t_k (the time since the reference epoch t_{oe}, as specified in the ephemeris)

$$t_k = t - t_{oe} \tag{3.45}$$

where t is the GPS system time at the time of transmission. Time t_k should be corrected for the end of week crossover as

$$\begin{aligned} &if(t_k > 302,400) \\ &\qquad t_k = t_k - 604,800 \\ &elseif\ (t_k < -302,400) \\ &\qquad t_k = t_k + 604,800 \\ &endif \end{aligned}$$

d. Adjust the mean motion by the correction Δn specified in the ephemeris

$$n = n_0 + \Delta n \tag{3.46}$$

e. Compute the mean anomaly M_k at time t_k

$$M_k = M_0 + nt_k \tag{3.47}$$

where M_o is the mean anomaly at the reference time.

f. Calculate the eccentric anomaly E_k by solving Kepler's law as given in Eq. (3.14)

$$E_k = M_k + e \sin E_k \tag{3.48}$$

where e is the eccentricity of the satellite orbit. Normally, the above equation is solved iteratively by setting an initial $E_k = M_k$. In addition, the answer should be corrected to lie between 0 and 2π.

g. Calculate the true anomaly v_k

$$v_k = \tan^{-1}\left\{\frac{\sqrt{1-e^2}\sin E_k/(1-e\cos E_k)}{(\cos E_k - e)/(1-e\cos E_k)}\right\} \tag{3.49}$$

h. Compute the argument of latitude Φ_k

$$\Phi_k = v_k + \omega \tag{3.50}$$

i. Calculate the three harmonic perturbations. The argument of latitude correction is

$$\delta u_k = C_{us}\sin 2\Phi_k + C_{uc}\cos 2\Phi_k \tag{3.51}$$

The radius correction is

$$\delta r_k = C_{rs}\sin 2\Phi_k + C_{rc}\cos 2\Phi_k \tag{3.52}$$

The inclination correction is

$$\delta i_k = C_{is}\sin 2\Phi_k + C_{ic}\cos 2\Phi_k \tag{3.53}$$

j. Compute the corrected argument of latitude

$$u_k = \Phi_k + \delta u_k \tag{3.54}$$

k. Compute the corrected radius

$$r_k = a(1 - e\cos E_k) + \delta r_k \tag{3.55}$$

l. Calculate the corrected inclination

$$i_k = i_0 + \delta i_k + (IDOT)t_k \tag{3.56}$$

m. Calculate the satellite's position in its orbital plane

$$x_k^{'} = r_k\cos u_k \tag{3.57}$$

$$y_k^{'} = r_k\sin u_k \tag{3.58}$$

n. Compute the corrected longitude of the ascending node Ω_k

$$\Omega_k = \Omega_0 + \left(\dot{\Omega} - \dot{\Omega}_e\right)t_k - \dot{\Omega}_e t_{oe} \tag{3.59}$$

where $\dot{\Omega}_e$ is the Earth's rotation rate. It is the same as ω_e, but $\dot{\Omega}_e$ is used to be consistent with the main reference (McLellan 1992; IS-GPS-200F 2011).

o. Finally, compute the satellite's position in the e-frame

$$x_k = x_k^{'}\cos\Omega_k - y_k^{'}\cos i_k\sin\Omega_k \tag{3.60}$$

$$y_k = x_k^{'}\sin\Omega_k + y_k^{'}\cos i_k\cos\Omega_k \tag{3.61}$$

$$z_k = y_k^{'} \sin i_k \tag{3.62}$$

There is one more refinement to be accounted for. Ephemeris parameters in the navigation message give $\tilde{\mathbf{P}}^m$, which is the satellite's position at the time of signal transmission expressed in ECEF. However, at the time that the signal is received by the receiver's antenna the e-frame would have rotated about the Earth's z-axis by $\omega_e(t_r - t_t)$ as a result of the globe's rotation during the transit time $(t_r - t_t)$. It is therefore necessary to express the position of the satellite in the e-frame at time t_r instead of time t_t (2011). This can be accomplished by multiplying $\tilde{\mathbf{P}}^m$, with an appropriate rotation matrix as follows

$$\mathbf{P}^m = \begin{bmatrix} \cos\omega_e(t_r - t_t) & \sin\omega_e(t_r - t_t) & 0 \\ -\sin\omega_e(t_r - t_t) & \cos\omega_e(t_r - t_t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tilde{\mathbf{P}}^m \tag{3.63}$$

$$\mathbf{P}^m = R_{et}^{er} \tilde{\mathbf{P}}^m \tag{3.64}$$

where R_{et}^{er} represents the transformation matrix from the e-frame at the time the GPS signal was transmitted to the e-frame at the time of its reception. However, since ω_e and the transit time $t_r - t_t$ are very small, we can use the approximations $\cos\omega_e(t_r - t_t) \approx 1$ and $\sin\omega_e(t_r - t_t) \approx \omega_e(t_r - t_t)$

3.7.4 Calculation of Satellite Velocity

The velocity of a satellite can be computed (Misra and Enge 2001) by taking the derivative of the position Eq. (3.60) through (3.62). This can be done by the following step-by-step procedure

a. Calculate the rate of change of the eccentric anomaly

$$\dot{E}_k = \frac{n}{1 - e\cos E_k} \tag{3.65}$$

b. Calculate the rate of change of the argument of latitude

$$\dot{\Phi}_k = \frac{\sqrt{1 - e^2}}{1 - e\cos E_k} \dot{E}_k \tag{3.66}$$

c. Compute the rate of change of the corrected argument of latitude

$$\dot{u}_k = (1 + 2C_{us}\cos 2\Phi_k - 2C_{uc}\sin 2\Phi_k)\dot{\Phi}_k \tag{3.67}$$

d. Find the rate of change of the corrected radius

$$\dot{r}_k = 2(C_{rs}\cos 2\Phi_k - C_{rc}\sin 2\Phi_k)\dot{\Phi}_k + Ae\sin(E_k)\dot{E}_k \tag{3.68}$$

e. Calculate the rate of change of the satellite's position in its orbital plane

$$\dot{x}_k^{'} = \dot{r}_k\cos u_k - r_k\sin(u_k)\dot{u}_k \tag{3.69}$$

$$\dot{y}_k^{'} = \dot{r}_k\sin u_k + r_k\cos(u_k)\dot{u}_k \tag{3.70}$$

f. Compute the rate of change of the corrected inclination

$$\frac{di_k}{dt} = 2(C_{is}\cos 2\Phi_k - C_{ic}\sin 2\Phi_k)\dot{\Phi}_k + IDOT \tag{3.71}$$

g. Find the rate of change of the corrected longitude of the ascending node

$$\dot{\Omega}_k = \dot{\Omega} - \dot{\Omega}_e \tag{3.72}$$

h. Differentiate Eqs. (3.60) through (3.62) to obtain the velocity of the satellite in the ECEF frame

$$\dot{x}_k = \dot{x}_k^{'}\cos\Omega_k - \dot{y}_k^{'}\cos i_k\sin\Omega_k + y_k^{'}\sin i_k\sin(\Omega_k)\frac{di_k}{dt} - y_k\dot{\Omega}_k \tag{3.73}$$

$$\dot{y}_k = \dot{x}_k^{'}\sin\Omega_k + \dot{y}_k^{'}\cos i_k\cos\Omega_k - y_k^{'}\sin i_k\cos(\Omega_k)\frac{di_k}{dt} + x_k\dot{\Omega}_k \tag{3.74}$$

$$\dot{z}_k = y_k^{'}\sin i_k + y_k\cos(i_k)\frac{di_k}{dt} \tag{3.75}$$

3.8 Receiver Position and Velocity Estimation

As mentioned earlier, there are three main observables related to GPS which can be used to estimate the navigational data, but the pseudo-ranges and Doppler measurements are treated next in the context of position and velocity estimation.

3.8.1 Pseudo-Range Measurements

Pseudo-ranges are obtained by measuring the time it takes for the GPS signal to propagate from the satellite to the receiver and then multiplying by the speed of light. Since the satellite's and receiver's clocks are not synchronised (mainly due to the clock in the receiver being inexpensive), these pseudo-ranges are biased by an amount equal to the receiver's clock offset. This offset is a fourth unknown in addition to the positional components of latitude, longitude and height. Therefore measurements from at least four satellites are required to solve these equations for four unknowns. Figure 3.20 depicts this concept of position estimation.

The pseudo-range measurement for the mth satellite can be written as

$$\rho^m = r^m + c\delta t_r - c\delta t_s + cI^m + cT^m + \varepsilon_\rho^m \tag{3.76}$$

where

- ρ^m is the measured pseudo-range between the mth satellite and the receiver (meters)
- r^m is the true range between the receiver's antenna at time t_r and the satellite's antenna at time t_t (meters)
- δt_r is the receiver's clock offset (sec)
- δt_s is the satellite's clock offset (sec)
- I^m is the ionospheric delay (sec)
- T^m is the tropospheric delay (sec)
- ε_ρ^m is the error in the range due to various sources such as receiver noise, multipath, satellite clock modeling, and orbit discrepancies (meters)

3.8.2 Position Estimation

After compensating for satellite clock bias, ionospheric errors and tropospheric errors (McLellan 1992; Farrell 2008) we can write the corrected pseudo-range as

$$\rho_c^m = r^m + c\delta t_r + \tilde{\varepsilon}_\rho^m \tag{3.77}$$

where, $\tilde{\varepsilon}_\rho^m$ represents the total effect of residual errors.

The geometric range from the mth satellite to the receiver is

$$r^m = \sqrt{(x - x^m)^2 + (y - y^m)^2 + (z - z^m)^2} = \|\mathbf{x} - \mathbf{x}^m\| \tag{3.78}$$

where

$\mathbf{x} = [x, y, z]^T$ is the receiver position in ECEF frame

$\mathbf{x}^m = [x^m, y^m, z^m]^T$ is the position of the m^{th} satellite in ECEF frame.

Equation (3.77) can be written in vector form as

$$\rho_c^m = \|\mathbf{x} - \mathbf{x}^m\| + b_r + \tilde{\varepsilon}_\rho^m \tag{3.79}$$

where $b_r = c\delta t_r$ is the error in range (in meters) due to the receiver's clock bias.

To linearize Eq. (3.79) we will use a Taylor series expansion and discard the higher order terms (HOT) in order to extract only the linear terms.

For any function $f(x, y, z)$ the Taylor series expansion around (x_i, y_i, z_i) is

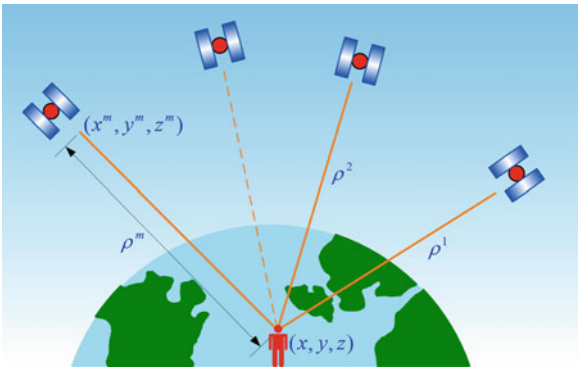


Fig. 3.20 Position estimation requires range measurements to three satellites to calculate the position of the receiver. A fourth satellite is needed to estimate the clock bias in the receiver common to all pseudo-range measurements

$$f(x,y,z) = f(x_i,y_i,z_i) + \left.\frac{\partial f}{\partial x}\right|_{x_i,y_i,z_i}(x-x_i) + \left.\frac{\partial f}{\partial y}\right|_{x_i,y_i,z_i}(y-y_i) + \left.\frac{\partial f}{\partial z}\right|_{x_i,y_i,z_i}(z-z_i) + H.O.T \tag{3.80}$$

Linearizing Eq. (3.79) around the current best estimate (Misra and Enge 2001) of

$$\mathbf{x}_{EST} = [x_{EST}, y_{EST}, z_{EST}]^T \tag{3.81}$$

gives

$$\rho_c^m = \sqrt{(x_{EST}-x^m)^2 + (y_{EST}-y^m)^2 + (z_{EST}-z^m)^2} + \frac{(x_{EST}-x^m)(x-x_{EST}) + (y_{EST}-y^m)(y-y_{EST}) + (z_{EST}-z^m)(z-z_{EST})}{\sqrt{(x_{EST}-x^m)^2 + (y_{EST}-y^m)^2 + (z_{EST}-z^m)^2}} + b_r + \tilde{\varepsilon}_\rho^m \tag{3.82}$$

We define the estimated range as

$$\rho_{c,EST}^m = \sqrt{(x_{EST}-x^m)^2 + (y_{EST}-y^m)^2 + (z_{EST}-z^m)^2} + b_{r,EST} \tag{3.83}$$

and so have

$$\rho_c^m - \rho_{c,EST}^m = \frac{(x_{EST}-x^m)(x-x_{EST}) + (y_{EST}-y^m)(y-y_{EST}) + (z_{EST}-z^m)(z-z_{EST})}{\sqrt{(x_{EST}-x^m)^2 + (y_{EST}-y^m)^2 + (z_{EST}-z^m)^2}} + b_r - b_{r,EST} + \tilde{\varepsilon}_\rho^m \tag{3.84}$$

which can be written more compactly as

$$\delta\rho_c^m = \left(\mathbf{1}_{EST}^m\right)^T \delta\mathbf{x} + \delta b_r + \tilde{\varepsilon}_\rho^m \tag{3.85}$$

where

$$\delta\rho_c^m = \rho_c^m - \rho_{c,EST}^m$$

$$\delta b_r = b_r - b_{r,EST}$$

$$\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_{EST} = [x, y, z]^T - [x_{EST}, y_{EST}, z_{EST}]^T$$

$$\mathbf{1}_{EST}^m = \frac{[(x_{EST} - x^m), (y_{EST} - y^m), (z_{EST} - z^m)]^T}{\sqrt{(x_{EST} - x^m)^2 + (y_{EST} - y^m)^2 + (z_{EST} - z^m)^2}} \tag{3.86}$$

with $\mathbf{1}_{EST}^m$ being the estimated line of sight unit vector from the mth satellite to the receiver's position.

For M satellites, the linearized pseudo-range measurements equations can be written as

$$\delta\boldsymbol{\rho}_c = \begin{bmatrix} \delta\rho_c^1 \\ \delta\rho_c^2 \\ \vdots \\ \delta\rho_c^M \end{bmatrix}_{M\times 1} = \begin{bmatrix} \left(\mathbf{1}_{EST}^1\right)^T & 1 \\ \left(\mathbf{1}_{EST}^2\right)^T & 1 \\ \vdots & 1 \\ \left(\mathbf{1}_{EST}^M\right)^T & 1 \end{bmatrix}_{M\times 4} \begin{bmatrix} \delta\mathbf{x} \\ \delta b_r \end{bmatrix}_{4\times 1} + \begin{bmatrix} \tilde{\varepsilon}_\rho^1 \\ \tilde{\varepsilon}_\rho^2 \\ \vdots \\ \tilde{\varepsilon}_\rho^M \end{bmatrix}_{M\times 1} \tag{3.87}$$

$$\delta\boldsymbol{\rho}_c = G_{M\times 4} \begin{bmatrix} \delta\mathbf{x} \\ \delta b_r \end{bmatrix}_{4\times 1} + \tilde{\boldsymbol{\varepsilon}}_{\rho,M\times 1} \tag{3.88}$$

where G is the *geometry* matrix with $M \times 4$ dimensions which characterizes the relative geometry of a satellite and the receiver, written as

$$G_{M\times 4} = \begin{bmatrix} \left(\mathbf{1}_{EST}^1\right)^T & 1 \\ \left(\mathbf{1}_{EST}^2\right)^T & 1 \\ \vdots & 1 \\ \left(\mathbf{1}_{EST}^M\right)^T & 1 \end{bmatrix} \tag{3.89}$$

For four satellites (M = 4) there are four equations and four unknowns and it is possible to solve Eq. (3.88) for unknowns directly as

$$\begin{bmatrix} \delta\mathbf{x} \\ \delta b_r \end{bmatrix} = G^{-1}\delta\boldsymbol{\rho}_{\boldsymbol{c}} \tag{3.90}$$

However, with the GPS constellation containing a minimum of 24 satellites, the average satellite availability is always greater than four. In the case of $M > 4$ we

have an over-determined system and the solution is sought by a least squares criterion. We therefore need to find $\left[\delta\hat{\mathbf{x}}, \delta\hat{b}_r\right]^T$ such that it minimizes

$$\left\| \delta\boldsymbol{\rho}_c - G \begin{bmatrix} \delta\hat{\mathbf{x}} \\ \delta\hat{b}_r \end{bmatrix} \right\|^2 \tag{3.91}$$

The least squares solution is given by

$$\begin{bmatrix} \delta\hat{\mathbf{x}} \\ \delta\hat{b}_r \end{bmatrix} = (G^T G)^{-1} G^T \delta\boldsymbol{\rho}_c \tag{3.92}$$

Now the improved estimates of their states are

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{x}_{EST} + \delta\hat{\mathbf{x}} \\ \hat{b}_r &= b_{r,EST} + \delta\hat{b}_r \end{aligned} \tag{3.93}$$

Up until now, it has been assumed that all pseudo-range measurements are of equal quality, which may not always be true; low elevation satellite measurements contain larger errors than higher elevation satellites. Using weighting factors that take into account the elevations of the available satellite, the least squares solution becomes

$$\begin{aligned} \begin{bmatrix} \delta\hat{\mathbf{x}} \\ \delta\hat{b}_r \end{bmatrix} &= (G^T W G)^{-1} G^T W \delta\boldsymbol{\rho}_c \\ [\delta\hat{\mathbf{y}}] &= (G^T W G)^{-1} G^T W \delta\boldsymbol{\rho}_c \end{aligned} \tag{3.94}$$

where W is the weighting matrix.

3.8.3 Satellite Geometry and Dilution of Precision

The positions of the available satellites in relation to the GPS receiver play an important role in the accuracy of the position estimation, and this affect is called geometric dilution of precision (GDOP) or simply dilution of precision (DOP). This concept is illustrated in Fig. 3.21 for three satellite-user geometries using an example in two dimensions. Ideally, the signals from the satellites should form circles which intersect at a point (the receiver's position),[3] provided the user has a precise calculation of the transit time of the signals. This is illustrated in Fig. 3.21a. However, this is never true because there is always an error in the range measurements due to clock biases, etc. With this uncertainty in measurements,

[3] There are two points of intersection, but the other point is remote and can easily be isolated and discarded.

the signals from each satellite form two concentric circles as shown in Figs. 3.21b and 3.21c. The areas between the two circles is the region of uncertainty in the range. This pair of concentric circles intersects forming a plane (instead of a point) which contains the region of possible receiver positions. The area of this region will depend on the relative geometry of the two satellites with respect to the user. In Fig. 3.21b the satellites are located at right angles relative to the user and the area of intersection is smaller and the DOP is lower; therefore the error in the position estimation will be smaller. Figure 3.21c shows the satellites almost collinear with respect to the receiver. Assuming the same ranging error (the same region of uncertainty in the circles), we see that the area of intersection is larger in this case, meaning that the computed position will be less accurate and the DOP will be greater.

In the case of GPS, dilution of precision is characterized by the geometry matrix G, which relates the parameters of the user's position and time bias errors to those of the pseudo-range errors.

The error covariance of the estimated quantities of Eq. (3.92) can be written as

$$\begin{aligned}
E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)&=E\left\{ (G^TG)^{-1}G^T\delta\boldsymbol{\rho}_c \quad \left[(G^TG)^{-1}G^T\delta\boldsymbol{\rho}_c\right]^T \right\}\\
E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)&=E\left\{ (G^TG)^{-1}G^T\delta\boldsymbol{\rho}_c \quad \delta\boldsymbol{\rho}_c^TG\left((G^TG)^{-1}\right)^T \right\}\\
E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)&=E\left\{ (G^TG)^{-1}G^T\delta\boldsymbol{\rho}_c \quad \delta\boldsymbol{\rho}_c^TG\left((G^TG)^T\right)^{-1} \right\}\\
E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)&=E\left\{ (G^TG)^{-1}G^T\delta\boldsymbol{\rho}_c \quad \delta\boldsymbol{\rho}_c^TG(G^TG)^{-1} \right\}\\
E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)&=(G^TG)^{-1}G^TE(\delta\boldsymbol{\rho}_c\delta\boldsymbol{\rho}_c^T)G(G^TG)^{-1}
\end{aligned} \tag{3.95}$$

to yield

$$E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T) = (G^TG)^{-1}G^TRG(G^TG)^{-1} \tag{3.96}$$

where R is the covariance of the pseudo-range measurements. Assuming that the measurement errors are uncorrelated (apart from having a zero mean) and have the same variance σ^2, then

$$R = \sigma^2 I \tag{3.97}$$

and the above covariance (3.96) can be expressed as

$$\begin{aligned}
E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)&=(G^TG)^{-1}G^T\sigma^2IG(G^TG)^{-1}\\
E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)&=\sigma^2(G^TG)^{-1}G^TIG(G^TG)^{-1}\\
E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)&=\sigma^2(G^TG)^{-1}G^TG(G^TG)^{-1}
\end{aligned} \tag{3.98}$$

to obtain

$$E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)=\sigma^2(G^TG)^{-1} \tag{3.99}$$

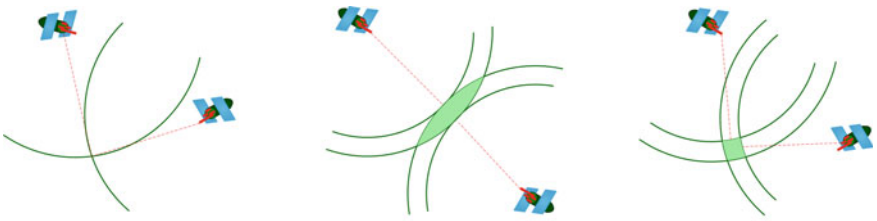


Fig. 3.21 The dilution of precision with range measurements in two dimensions. **a** The ideal position estimation with no clock errors. **b** A position estimation with a clock offset but a favorable satellite-user geometry. **c** A position estimation with same clock offset but a poor satellite-user geometry

By defining $C_x = E(\delta\hat{\mathbf{y}}\delta\hat{\mathbf{y}}^T)$ and H = $(G^TG)^{-1}$ we can rewrite this as

$$C_x = \sigma^2 H \tag{3.100}$$

where σ^2 is the variance of the user's range error. If we define σ_x^2, σ_y^2, σ_z^2 as the variances of the x, y and z positional components and σ_b^2 as the variance of the receiver's clock bias error, then Eq. (3.100) gives

$$\sigma_x^2 = \sigma^2 H_{11};\ \sigma_y^2 = \sigma^2 H_{22}; \sigma_z^2 = \sigma^2 H_{33};\ \sigma_b^2 = \sigma^2 H_{44} \tag{3.101}$$

and the RMS position error is

$$\Delta r_{RMS} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$$

$$\Delta r_{RMS} = \sigma\sqrt{H_{11} + H_{22} + H_{33}} \tag{3.102}$$

It is evident from (3.101) and (3.102) that the position estimates are affected by the variance of the user's range error as well as by the diagonal components of the H matrix, which totally depends on the geometry matrix G.

Various DOP parameters which characterize the role of user-satellite geometry can now be defined.

$$\text{Position dilution of precision (PDOP)} = \sqrt{H_{11} + H_{22} + H_{33}} \tag{3.103}$$

$$\text{Time dilution of precision(TDOP)} = \sqrt{H_{44}} \tag{3.104}$$

$$\text{Geometric dilution of precision (GDOP)} = \sqrt{H_{11} + H_{22} + H_{33} + H_{44}} \tag{3.105}$$

To get the vertical and horizontal dilution of precision (which is more intuitive to the user) we must convert the position error Δx from rectangular ECEF to ENU coordinates by multiplying with the direction cosine matrix R_e^l

$$\Delta\mathbf{x}_L = R_e^l \Delta\mathbf{x} \tag{3.106}$$

where $\Delta \mathbf{x}_L = (\Delta x_E, \Delta y_N, \Delta z_U)^T$.

Now we can also define

$$\begin{aligned}\begin{bmatrix}\Delta \mathbf{x}_L \\ \Delta b_r\end{bmatrix} &= \begin{bmatrix}R_e^l & 0 \\ 0 & 1\end{bmatrix}\begin{bmatrix}\Delta \mathbf{x} \\ \Delta b_r\end{bmatrix} \\ &= \tilde{R}_e^l \begin{bmatrix}\Delta \mathbf{x} \\ \Delta b_r\end{bmatrix}\end{aligned} \tag{3.107}$$

and the covariance is defined by the law of covariance propagation (Parkinson et al. 1996) as

$$\begin{aligned} Cov\begin{bmatrix}\Delta \mathbf{x}_L \\ \Delta b_r\end{bmatrix} &= Q_x = \tilde{R}_e^l C_x (\tilde{R}_e^l)^T \\ Q_x &= \sigma^2 \tilde{R}_e^l (G^T G)^{-1} (\tilde{R}_e^l)^T \\ Q_x &= \sigma^2 \left\{ \tilde{R}_e^l G^T G \left(\tilde{R}_e^l\right)^T \right\}^{-1} \\ Q_x &= \sigma^2 \left(\tilde{G}^T \tilde{G}\right)^{-1} \end{aligned} \tag{3.108}$$

where $\tilde{G} = G\left(\tilde{R}_e^l\right)^T$ and has a similar structure as the geometry matrix G. By now letting $\tilde{H} = (\tilde{G}^T \tilde{G})^{-1}$ we have

$$Q_x = \sigma^2 \tilde{H} \tag{3.109}$$

and $\sigma_E^2 = \sigma^2 \tilde{H}_{11}$; $\sigma_N^2 = \sigma^2 \tilde{H}_{22}$; $\sigma_U^2 = \sigma^2 \tilde{H}_{33}$; $\sigma_b^2 = \sigma^2 \tilde{H}_{44}$ where $\sigma_E, \sigma_N, \sigma_U$ are the standard deviations of the east, north and up components of the position error respectively.

The diagonal elements of $\tilde{H}$ correspond to the east, north, vertical, and time DOP as follows

$$\tilde{H} = \begin{bmatrix} EDOP^2 & & & \\ & NDOP^2 & & \\ & & VDOP^2 & \\ & & & TDOP^2 \end{bmatrix} \tag{3.110}$$

Now, we have

$$\text{Horizontal dilution of precision (HDOP)} = \sqrt{\tilde{H}_{11} + \tilde{H}_{22}} \tag{3.111}$$

$$\text{Vertical dilution of precision} = \sqrt{\tilde{H}_{33}} \tag{3.112}$$

We also obtain

$$\text{RMS horizontal position error} = \sqrt{\sigma_E^2 + \sigma_N^2} = \sigma.HDOP \tag{3.113}$$

$$\text{RMS vertical error} = \sigma_U = \sigma.VDOP \tag{3.114}$$

$$\text{RMS 3D error} = \sqrt{\sigma_E^2 + \sigma_N^2 + \sigma_U^2} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} = \sigma.PDOP \tag{3.115}$$

It is interesting to note that DOP is a function of the satellite-user geometry only, and no observations are necessary in order to calculate DOP. It can therefore be computed in advance using almanac or satellite orbital data for planning of the trajectory data collection. The satellite-user geometry improves as the number of available satellites increases, which increases the positional accuracy except when an additional satellite gives a very poor range measurement.

3.8.4 Doppler Measurements

The frequency observed by the GPS receiver differs from the L1 or L2 due to the Doppler shift caused by the relative motion of the satellite and the receiver. This is measured in the carrier tracking loop of a GPS receiver. The Doppler shift is the projection of the relative velocities onto the line of sight unit vector scaled by the transmitted frequency divided by the speed of light (Misra and Enge 2001); i.e.

$$D^m = \frac{\{(\mathbf{v}^m - \mathbf{v}).1^m\}L}{c} \tag{3.116}$$

where

$\mathbf{v}^m = [v_x^m, v_y^m, v_z^m]$ is the velocity of the mth satellite in the e-frame

$\mathbf{v} = [v_x, v_y, v_z]$ is the true velocity of the receiver in the e-frame

$\mathbf{1}^m = \frac{[(x-x^m),(y-y^m),(z-z^m)]^T}{\sqrt{(x-x^m)^2+(y-y^m)^2+(z-z^m)^2}} = \left[1_x^m, 1_y^m, 1_z^m\right]^T$ is the true line of sight unit vector from the mth satellite to the receiver

L is the satellite transmitted frequency

c is the speed of light

Given the Doppler effect, the pseudo-range rate $\dot{\rho}^m$ can be computed as

$$\dot{\rho}^m = -\frac{D^m c}{L} \tag{3.117}$$

3.8.5 Velocity Estimation from Doppler

The velocity can be estimated from the pseudo-range rate (Parkinson et al. 1996), starting by differentiating Eq. (3.77) to obtain

$$\dot{\rho}^m = \dot{r}^m + c\delta\dot{t}_r + \tilde{\varepsilon}_{\dot{\rho}}^m \tag{3.118}$$

The true pseudo-range rate is expressed as

$$\dot{r}^m = 1_x^m.(v_x - v_x^m) + 1_y^m.(v_y - v_y^m) + 1_z^m.(v_z - v_z^m) \tag{3.119}$$

and substituting this into Eq. (3.118) gives

$$\begin{aligned}\dot{\rho}^m &= 1_x^m.(v_x - v_x^m) + 1_y^m.(v_y - v_y^m) + 1_z^m.(v_z - v_z^m) + c\delta\dot{t}_r + \varepsilon_{\dot{\rho}}^m \\ &= 1_x^m.(v_x - v_x^m) + 1_y^m.(v_y - v_y^m) + 1_z^m.(v_z - v_z^m) + c\delta\dot{t}_r + \varepsilon_{\dot{\rho}}^m\end{aligned} \tag{3.120}$$

where

$\delta\dot{t}_r$ is the receiver's clock drift (sec/sec)
$\varepsilon_{\dot{\rho}}^m$ is the error in observation (meters/sec)

The term $c\delta\dot{t}_r$ is the receiver clock drift in meters/sec and represented as d_r in subsequent equations.

Since we can obtain the velocity of the satellite from the data contained in the ephemeris file, the known velocity terms can be transferred to the left-hand side

$$\dot{\rho}^m + 1_x^m.v_x^m + 1_y^m.v_y^m + 1_z^m v_z^m = 1_x^m.v_x + 1_y^m.v_y + 1_z^m.v_z + d_r + \varepsilon_{\dot{\rho}}^m \tag{3.121}$$

and this can be written more compactly as

$$\begin{aligned}\dot{\rho}^m + \mathbf{1}^m.\mathbf{v}^m &= \mathbf{1}^m.\mathbf{v} + d_r + \varepsilon_{\dot{\rho}}^m \\ \bar{\dot{\rho}}^m &= \mathbf{1}^m.\mathbf{v} + d_r + \varepsilon_{\dot{\rho}}^m\end{aligned} \tag{3.122}$$

where

$$\bar{\dot{\rho}}^m = \dot{\rho}^m + \mathbf{1}^m.\mathbf{v}^m$$

and the line of sight unit vector $\mathbf{1}^m$ is obtained from the estimated position of the user.

For M satellites, the pseudo-range rate measurements model can be written as

$$\bar{\dot{\boldsymbol{\rho}}} = \begin{bmatrix} \bar{\dot{\rho}}^1 \\ \bar{\dot{\rho}}^2 \\ \vdots \\ \bar{\dot{\rho}}^M \end{bmatrix}_{M\times 1} = \begin{bmatrix} (\mathbf{1}^1)^T & 1 \\ (\mathbf{1}^2)^T & 1 \\ \vdots & \vdots \\ (\mathbf{1}^M)^T & 1 \end{bmatrix}_{M\times 4} \begin{bmatrix} \mathbf{v} \\ d_r \end{bmatrix}_{4\times 1} + \begin{bmatrix} \varepsilon_{\dot{\rho}}^1 \\ \varepsilon_{\dot{\rho}}^2 \\ \vdots \\ \varepsilon_{\dot{\rho}}^M \end{bmatrix}_{M\times 1} \tag{3.123}$$

$$\bar{\dot{\boldsymbol{\rho}}}_{M\times 1} = G_{M\times 4}\begin{bmatrix}\mathbf{v}\\ d_r\end{bmatrix}_{4\times 1} + \boldsymbol{\varepsilon}_{\dot{\rho}M\times 1} \tag{3.124}$$

Assuming that the position of the receiver is known, we can find the velocity and the clock drift $\left[\hat{\mathbf{v}}, \hat{d}_r\right]^T$ by minimizing

$$\begin{bmatrix}\hat{\mathbf{v}}\\ \hat{d}_r\end{bmatrix} = (G^T G)^{-1} G^T \bar{\dot{\boldsymbol{\rho}}} \tag{3.125}$$

whose least squares solution is

$$\begin{bmatrix}\hat{\mathbf{v}}\\ \hat{d}_r\end{bmatrix} = (G^T G)^{-1} G^T \bar{\dot{\boldsymbol{\rho}}} \tag{3.126}$$

The position estimate for the line of sight unit vector **1** for next epochs can be estimated by using the velocity information.

In the above discussion, mathematical details of the velocity estimation were presented assuming that the initial position of the receiver was known in advance, but in practice the initial position of the receiver is usually unknown and therefore its position and velocity are estimated simultaneously.

3.8.6 Position and Velocity Estimation

Position and velocity can be estimated simultaneously using the procedure below. In fact this procedure is better, since it does not require projecting the position for the next epoch as was done in the previous section.

The pseudo-range rate can be modeled as

$$\dot{\rho}^m = 1_x^m.(v_x - v_x^m) + 1_y^m.(v_y - v_y^m) + 1_z^m.(v_z - v_z^m) + d_r + \varepsilon_{\dot{\rho}}^m \tag{3.127}$$

and the estimated pseudo-range rate is

$$\dot{\rho}_{EST}^m = 1_{x,EST}^m.(v_{x,EST} - v_x^m) + 1_{y,EST}^m.(v_{y,EST} - v_y^m) + 1_{z,EST}^m.(v_{z,EST} - v_z^m) + d_{r,EST} \tag{3.128}$$

where $v_{x,EST}, v_{y,EST}, v_{z,EST}$ are the estimated velocity components of the receiver in the ECEF frame.

The error in the pseudo-range rate can then be expressed as

$$\dot{\rho}^m - \dot{\rho}_{EST}^m = 1_{x,EST}^m.(v_x - v_{x,EST}) + 1_{y,EST}^m.(v_y - v_{y,EST}) + 1_{z,EST}^m.(v_z - v_{z,EST}) + d_r - d_{r,EST} + \varepsilon_{\dot{\rho}}^m \tag{3.129}$$

$$= \left(\mathbf{1}_{EST}^{m}\right)^{T}(\mathbf{v} - \mathbf{v}_{EST}) + d_r - d_{r,EST} + \varepsilon_{\dot{\rho}}^{m} \tag{3.130}$$

$$\delta\dot{\rho}^{m} = (\mathbf{1}_{EST}^{m})^{T}.\delta\mathbf{v} + \delta d_r + \varepsilon_{\dot{\rho}}^{m} \tag{3.131}$$

where

$$\delta\dot{\rho}^{m} = \dot{\rho}^{m} - \dot{\rho}_{EST}^{m}$$

$$\delta d_r = d_r - d_{r,EST}$$

$$\delta\mathbf{v} = \mathbf{v} - \mathbf{v}_{EST} = \left[v_x, v_y, v_z\right]^{T} - \left[v_{x,EST}, v_{y,EST}, v_{z,EST}\right]^{T}$$

For M satellites, the linearized pseudo-range rate measurements model can be written as

$$\delta\dot{\boldsymbol{\rho}} = \begin{bmatrix} \delta\dot{\rho}^{1} \\ \delta\dot{\rho}^{2} \\ \vdots \\ \delta\dot{\rho}^{M} \end{bmatrix}_{M\times 1} = \begin{bmatrix} (\mathbf{1}_{EST}^{1})^{T} & 1 \\ (\mathbf{1}_{EST}^{2})^{T} & 1 \\ \vdots & \vdots \\ (\mathbf{1}_{EST}^{M})^{T} & 1 \end{bmatrix}_{M\times 4} \begin{bmatrix} \delta\mathbf{v} \\ \delta d_r \end{bmatrix}_{4\times 1} + \begin{bmatrix} \varepsilon_{\dot{\rho}}^{1} \\ \varepsilon_{\dot{\rho}}^{2} \\ \vdots \\ \varepsilon_{\dot{\rho}}^{M} \end{bmatrix}_{M\times 1} \tag{3.132}$$

$$\delta\dot{\boldsymbol{\rho}}_{M\times 1} = G_{M\times 4}\begin{bmatrix} \delta\mathbf{v} \\ \delta d_r \end{bmatrix}_{4\times 1} + \boldsymbol{\varepsilon}_{\boldsymbol{\rho M\times 1}} \tag{3.133}$$

Equations (3.87) for the pseudo-ranges and (3.132) for the pseudo-range rates can be combined into a single model as

$$\begin{bmatrix} \delta\rho_c^{1} \\ \vdots \\ \delta\rho_c^{M} \\ \delta\dot{\rho}^{1} \\ \vdots \\ \delta\dot{\rho}^{M} \end{bmatrix}_{2M\times 1} = \begin{bmatrix} (\mathbf{1}_{EST}^{1})^{T} & 1 & 0_{3\times 1} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ (\mathbf{1}_{EST}^{M})^{T} & 1 & 0_{3\times 1} & 0 \\ 0_{3\times 1} & 0 & (\mathbf{1}_{EST}^{1})^{T} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0_{3\times 1} & 0 & (\mathbf{1}_{EST}^{M})^{T} & 1 \end{bmatrix}_{2M\times 8} \begin{bmatrix} \delta\mathbf{x} \\ \delta b_r \\ \delta\mathbf{v} \\ \delta d_r \end{bmatrix}_{8\times 1} + \begin{bmatrix} \tilde{\varepsilon}_{\rho}^{1} \\ \vdots \\ \tilde{\varepsilon}_{\rho}^{M} \\ \varepsilon_{\dot{\rho}}^{1} \\ \vdots \\ \varepsilon_{\dot{\rho}}^{M} \end{bmatrix}_{2M\times 1} \tag{3.134}$$

$$\delta\mathbf{z}_{2M\times 1} = \tilde{G}_{2M\times 8}\delta\mathbf{S}_{8\times 1} + \boldsymbol{\varepsilon}_{\boldsymbol{2M\times 1}} \tag{3.135}$$

where

$$\delta\mathbf{z} = \left[\rho_c^{1}, \cdots, \rho_c^{M}, \dot{\rho}^{1} \cdots \dot{\rho}^{M}\right]^{T}$$
$$\delta\mathbf{S} = \left[\delta\mathbf{x}, \delta b_r, \delta\mathbf{v}, \delta d_r\right]^{T}$$
$$\boldsymbol{\varepsilon} = \left[\tilde{\varepsilon}_{\rho}^{1}, \cdots, \tilde{\varepsilon}_{\rho}^{M}, \varepsilon_{\dot{\rho}}^{1} \cdots, \varepsilon_{\dot{\rho}}^{M}\right]^{T}$$

$$\tilde{G} = \begin{bmatrix} (\mathbf{1}_{EST}^{1})^T & 1 & 0_{3\times1} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ (\mathbf{1}_{EST}^{M})^T & 1 & 0_{3\times1} & 0 \\ 0_{3\times1} & 0 & (\mathbf{1}_{EST}^{1})^T & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0_{3\times1} & 0 & (\mathbf{1}_{EST}^{M})^T & 1 \end{bmatrix}. \tag{3.136}$$

In the case of four or more satellites $(M \geq 4)$, the least squares solution to Eq. (3.135) is

$$\delta\hat{\mathbf{S}} = \begin{bmatrix} \delta\hat{\mathbf{x}} \\ \delta\hat{b}_r \\ \delta\hat{\mathbf{v}} \\ \delta\hat{d}_r \end{bmatrix} = \left(\tilde{G}^T\tilde{G}\right)^{-1}\tilde{G}^T\delta\mathbf{z} \tag{3.137}$$

Finally, the improved estimates of receiver's position and clock bias are

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{x}_{EST} + \delta\hat{\mathbf{x}} \\ \hat{b} &= b_{EST} + \delta\hat{b} \end{aligned} \tag{3.138}$$

and for the velocity and clock drift

$$\begin{aligned} \hat{\mathbf{v}} &= \mathbf{v}_{EST} + \delta\hat{\mathbf{v}} \\ \hat{d}_r &= d_{r,EST} + \delta\hat{d} \end{aligned} \tag{3.139}$$

If measurement weighting is available, then a weighted least squares solution is used. Also, if a priori estimates have large errors then the least squares solution will be iterated until the change in the estimate is sufficiently small.

3.9 Carrier Phase Positioning

Carrier phase measurements are much more precise than the code phase, but they are ambiguous owing to integer ambiguities in the measurements. However, if the integer ambiguities can be estimated the positioning accuracy can be considerably improved. As various methods have been proposed to resolve these ambiguities, it is appropriate to discuss carrier phase positioning. As discussed above, the carrier phase measurements can be modeled as

$$\Phi = r + c\delta t - c\delta t_s - I + T + \lambda N + \varepsilon_\Phi \tag{3.140}$$

There is a subtle point here that must be mentioned: the geometric range r is not the same for code and carrier due to the difference in the antenna phase center but we shall neglect this in our treatment. There are some other error terms which are

usually hard to estimate and may safely be ignored for medium accuracy. But for applications needing higher accuracy these have to be included. By taking into account the pertinent errors, Eq. (3.140) can be written as

$$\Phi = r + c(\delta t_r - \delta t_s) - I + T + \lambda N + c(\Delta t_r + \Delta t_s) + \lambda(\phi_{r0} - \phi_{s0}) + \Delta m + \varepsilon_\Phi \tag{3.141}$$

where

$\Delta t_r, \Delta t_s$ are the equipment delays that occur at the receiver and the satellite respectively (sec)

ϕ_{r0}, ϕ_{s0} are initial phases of the receiver and the satellite carrier (radians)

Δm is the error due to multipath at the receiver and the satellite (meters)

The equipment delay (also called the hardware delay) is due to the time it takes for the electrical signal to pass from the signal generator (processor) to the antenna phase center of the satellite, and in the case of the receiver from the antenna phase center to the processor. It differs for the carrier and the code phase measurements and also for L1 and L2. Similarly, the delay due to multipath is different for the carrier phase and the code phase. Small errors like tidal effects and relativity have still been neglected but these can also be modeled and removed.

Equation (3.141) contains a lot of nuisance parameters considering that we are interested only in the measurements and the position coordinates contained in the geometric ranges. We should either estimate the nuisance parameters or eliminate them by relative positioning techniques. The objective of relative positioning is to eliminate the nuisance parameters and simplify the equations that relate the carrier phase measurements to the position vector between the reference receivers and the roving receiver.

3.9.1 *Relative Positioning and Linear Combinations of GPS Observables*

Some of the errors which impair the accuracy of GPS exhibit certain correlations amongst signals received at several stations from several satellites at the same epochs. Linear combinations of these measurements are formed to take advantage of these correlations and improve the accuracy of the relative positions. In relative positioning, the difference of carrier or code measurements taken at two receivers allows the effect of errors that are common to the measurements to be reduced or removed.

Here it may be noted that there is a difference between relative and differential positioning. In differential positioning the *errors* are transmitted to rover stations and in relative positioning the *measurements* are conveyed to the rover. In many

applications relative positioning suffices (e.g. formation flying) but knowledge of the position of the reference stations is required for absolute positioning.

3.9.2 Relative Positioning

In relative positioning, the user's position is determined with respect to a known reference point or base station. Unlike differential positioning where the reference station broadcasts the corrections in the pseudo-ranges, in relative positioning the reference receiver broadcasts the time-tagged measurements of the pseudo-ranges. The user forms the difference of its own pseudo-ranges with those received from the reference station and then estimates its position relative to that. Basically, in relative positioning the vector between the user and the reference station (called the baseline vector, or just the baseline) is determined. If we know the coordinates of the reference receiver, the absolute position of the user can also be obtained.

As illustrated in Fig. 3.22, for a reference receiver j and a user receiver (or rover) i we can write

$$\mathbf{x}_i = \mathbf{x}_j + \mathbf{x}_{ji} \tag{3.142}$$

Therefore, the baseline vector becomes

$$\mathbf{x}_{ji} = \mathbf{x}_i - \mathbf{x}_j \tag{3.143}$$

and hence

$$\mathbf{x}_{ji} = \begin{bmatrix} x_i - x_j \\ y_i - y_j \\ z_i - z_j \end{bmatrix} = \begin{bmatrix} \Delta x_{ji} \\ \Delta y_{ji} \\ \Delta y_{ji} \end{bmatrix} \tag{3.144}$$

We can use code or phase ranges to obtain the relative position of the user. As phase ranges are more accurate they are usually used for relative positioning. For this purpose, we can form linear combinations of these ranges in order to reduce or remove errors that are common to the measurements. These combinations are formed by taking single, double, or triple differences of the measurements. They can be across satellites, receiver or time, or indeed any of their combinations.

3.9.3 Linear Combinations of GPS Measurements

We have seen that in the measurement equations for code and carrier phase, there are many nuisance parameters in addition to the position coordinates buried in the geometric range r. These nuisance parameters are ionospheric delay, tropospheric delay, satellite and receiver clock errors, ephemeris errors and integer ambiguities. We need to either significantly reduce or completely eliminate these. One way to

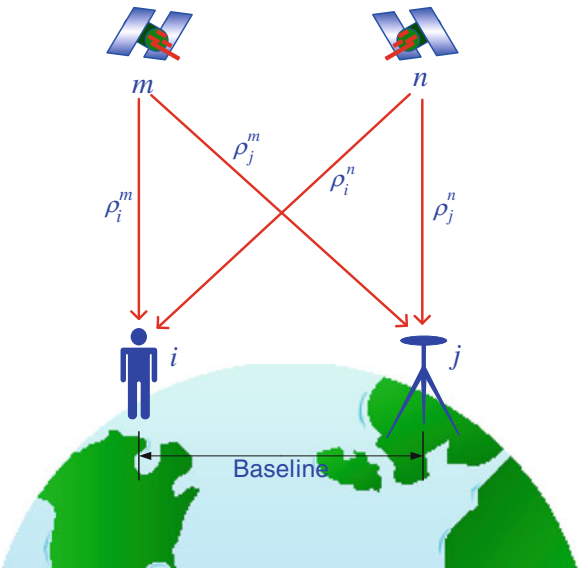


Fig. 3.22 An illustration of relative positioning. Relative positioning

accomplish this is by taking linear combinations of the measurements from a fixed known receiver (called either a reference station or a base station) and the moving receiver (a rover) whose coordinates are to be calculated. Linear combinations are formed by taking the difference of the measurements between receivers, satellites and epochs or any of their combinations. In this relative positioning technique, the known parameters are the satellite coordinates (from an ephemeris), the reference stations coordinates, and the pseudo-ranges. The only unknown is the rover whose coordinates are determined either absolutely or in relation to the base station. The most common linear combinations are

a. Between receiver single difference (RSD)

According to Fig. 3.22, two RSDs can be written as

$$\begin{aligned} RSD_{ij}^{m} &= \rho_i^m - \rho_j^m \\ RSD_{ij}^{n} &= \rho_i^n - \rho_j^n \end{aligned} \tag{3.145}$$

b. Between satellite single difference (SSD)

Similarly, two SSDs can be expressed as

$$\begin{aligned} SSD_{i}^{mn} &= \rho_i^m - \rho_i^n \\ SSD_{j}^{mn} &= \rho_j^m - \rho_j^n \end{aligned} \tag{3.146}$$

c. Double difference (DD)

This is difference of the receiver (or satellite) single difference

$$\begin{aligned} DD_{ij}^{m,n} &= RSD_{ij}^{m} - RSD_{ij}^{n} \\ DD_{i,j}^{mn} &= SSD_{i}^{mn} - SSD_{j}^{mn} \end{aligned} \tag{3.147}$$

d. Triple difference (TD)

This is the double difference at two different epochs

$$\begin{aligned} TD(t)_{ij}^{m,n} &= DD_{ij}^{m,n}(t_2) - DD_{ij}^{m,n}(t_1) \\ TD(t)_{i,j}^{mn} &= DD_{i,j}^{mn}(t_2) - DD_{i,j}^{mn}(t_1) \end{aligned} \tag{3.148}$$

RSD removes the receiver clock bias and mitigates the ionospheric errors, and DD further removes the ephemeris and satellite clock errors. TD also removes the integer ambiguities. Usually DD is used for carrier phase positioning because it removes most of the errors but it requires at least one base station. Let's examine these linear combinations in more detail.

3.9.3.1 Receiver Single Difference

Suppose two receivers i and j are tracking a certain satellite m. The carrier phase measurements from satellite k at receivers i and j would be modeled as

$$\begin{aligned} \Phi_i^m(t_i) =& r_i^m(t_i) + c\left[\delta t_r^i(t_i) - \delta t_s^m(t_i)\right] - I_i^m + T_i^m + \lambda N_i^m + c\left[\Delta t_r^i(t_i) + \Delta t_s^m(t_i)\right] + \\ & \lambda\left[\phi_r^i(t_0) - \phi_s^m(t_0)\right] + \Delta m_i^m + \varepsilon_\Phi \end{aligned} \tag{3.149}$$

$$\begin{aligned} \Phi_j^m(t_j) =& r_j^m(t_j) + c\left[\delta t_r^j(t_j) - \delta t_s^m(t_j)\right] - I_j^m + T_j^m + \lambda N_j^m + c\left[\Delta t_r^i(t_j) + \Delta t_s^m(t_j)\right] + \\ & \lambda\left[\phi_r^j(t_0) - \phi_s^m(t_0)\right] + \Delta m_j^m + \varepsilon_\Phi \end{aligned} \tag{3.150}$$

Ideally, these two measurements should be simultaneous but due to different clock errors they are generally not. Therefore the time arguments t_i, t_j for the two receivers are different. There will also be a different travel time from the reference receiver to the user receiver. However, with continuous receiver clock updates the measurements can be treated as simultaneous and the time argument is deleted. Furthermore, measurements of the reference station may be extrapolated to match the epoch of the user's receiver.

$$\Phi_i^m = r_i^m + c\left(\delta t_r^i - \delta t_s^m\right) - I_i^m + T_i^m + \lambda N_i^m + c\left(\Delta t_r^i + \Delta t_s^m\right) + \lambda\left[\phi_r^i(t_0) - \phi_s^m(t_0)\right] + \Delta m_i^m + \varepsilon_{\Phi i}^m \tag{3.151}$$

$$\Phi_j^m = r_j^m + c\left(\delta t_r^j - \delta t_s^m\right) - I_j^m + T_j^m + \lambda N_j^m + c\left(\Delta t_r^j + \Delta t_s^m\right) + \lambda\left[\phi_r^j(t_0) - \phi_s^m(t_0)\right] + \Delta m_j^m + \varepsilon_{\Phi j}^m \tag{3.152}$$

The difference between these two measurements is

$$\begin{aligned}\Phi_j^m - \Phi_i^m = &\left\{r_j^m - r_i^m\right\} + \left\{c\left(\delta t_r^j - \delta t_s^m\right) - c\left(\delta t_r^i - \delta t_s^m\right)\right\} - \left\{I_j^m - I_i^m\right\} + \\ &\left\{T_j^m - T_i^m\right\} + \left\{\lambda N_j^m - \lambda N_i^m\right\} + \left\{c\left(\Delta t_r^j + \Delta t_s^m\right) - c\left(\Delta t_r^i + \Delta t_s^m\right)\right\} + \\ &\left\{\lambda\left[\phi_r^j(t_0) - \phi_s^m(t_0)\right] - \lambda\left[\phi_r^i(t_0) - \phi_s^m(t_0)\right]\right\} + \left\{\Delta m_j^m - \Delta m_i^m\right\} + \left\{\varepsilon_{\Phi j}^m - \varepsilon_{\Phi i}^m\right\}\end{aligned} \tag{3.153}$$

The common satellite errors, which are the clock error, the satellite equipment delay, and the satellite's initial phase offset are all canceled out. We can combine the remaining difference of similar terms and denote them as

$$\begin{array}{ll} \Phi_{ji}^m = \Phi_j^m - \Phi_i^m & N_{ji}^m = N_j^m - N_i^m \\ r_{ji}^m = r_j^m - r_i^m & \Delta t_r^{ji} = \Delta t_r^j - \Delta t_r^i \\ \delta t_r^{ji} = \delta t_r^j - \delta t_r^i & \phi_r^{ji}(t_0) = \phi_r^j(t_0) - \phi_r^i(t_0) \\ I_{ji}^m = I_j^m - I_i^m & \Delta m_{ji}^m = \Delta m_j^m - \Delta m_i^m \\ T_{ji}^m = T_j^m - T_i^m & \varepsilon_{\Phi ji}^m = \varepsilon_{\Phi j}^m - \varepsilon_{\Phi i}^m \end{array} \tag{3.154}$$

Substituting the above terms into Eq. (3.153) yields

$$\Phi_{ji}^m = r_{ji}^m + c\delta t_r^{ji} - I_{ji}^m + T_{ji}^m + \lambda N_{ji}^m + c\Delta t_r^{ji} + \phi_r^{ji}(t_0) + \Delta m_{ji}^m + \varepsilon_{\Phi ji}^m \tag{3.155}$$

The ambiguity difference is still integer (and can be positive or negative). The measurement noise difference $\varepsilon_{\Phi ji}^k$ is larger by a factor of $\sqrt{2}$ than either of the individual noise terms. The common part of the satellite ephemeris errors (hidden in the geometric ranges r) is also eliminated. The common part is the difference in the projection of ephemeris error vector onto the ranges r_i and r_j. The ionospheric and tropospheric difference terms depend upon the length of the baseline, and for short baselines tend to zero as compared to the measurement and multipath errors. Taking account of these factors and ignoring equipment delay, initial phase offset and multipath, Eq. (3.155) can be simplified for a short baseline as

$$\Phi_{ji}^m = r_{ji}^m + c\delta t_r^{ji} + \lambda N_{ji}^m + + \varepsilon_{\Phi ji}^m \tag{3.156}$$

3.9.3.2 Double difference

Taking the RSD from another satellite n, we get an equation similar to (3.155)

$$\Phi_{ji}^{n} = r_{ji}^{n} + c\delta t_{r}^{ji} - I_{ji}^{n} + T_{ji}^{n} + \lambda N_{ji}^{n} + c\Delta t_{r}^{ji} + \phi_{r}^{ji}(t_0) + \Delta m_{ji}^{n} + \varepsilon_{\Phi ji}^{n} \tag{3.157}$$

and the difference between Eqs. (3.157) and (3.155) is

$$\begin{aligned}\Phi_{ji}^{m} - \Phi_{ji}^{n} = &\left[r_{ji}^{m} + c\delta t_{r}^{ji} - I_{ji}^{m} + T_{ji}^{m} + \lambda N_{ji}^{m} + c\Delta t_{r}^{ji} + \phi_{r}^{ji}(t_0) + \Delta m_{ji}^{m} + \varepsilon_{\Phi ji}^{m}\right] - \\ &\left[r_{ji}^{n} + c\delta t_{r}^{ji} - I_{ji}^{n} + T_{ji}^{n} + \lambda N_{ji}^{n} + c\Delta t_{r}^{ji} + \phi_{r}^{ji}(t_0) + \Delta m_{ji}^{n} + \varepsilon_{\Phi ji}^{n}\right]\end{aligned} \tag{3.158}$$

The difference terms for the clocks offset, equipment delays and initial offsets of the receiver have all canceled out. Combining the remaining terms gives

$$\Phi_{ji}^{mn} = r_{ji}^{mn} - I_{ji}^{mn} + T_{ji}^{mn} + \lambda N_{ji}^{mn} + \Delta m_{ji}^{mn} + \varepsilon_{\Phi ji}^{mn} \tag{3.159}$$

It can be verified that the double difference formed by taking the difference *between satellite single difference* (SSD) yields the same form as Eq. (3.159). For a short baseline and a relatively clear site that has no reflecting objects, the double difference terms relating to ionospheric, tropospheric terms and multipath reduce to zero (as tropospheric delays differ for even short baselines, it is usually modeled). Therefore, we obtain

$$\Phi_{ji}^{mn} = r_{ji}^{mn} + \lambda N_{ji}^{mn} + \varepsilon_{\Phi ji}^{mn} \tag{3.160}$$

It is worth emphasizing that each term in (3.160) which has two subscripts and two superscripts actually contains four terms, as follows

$$\begin{aligned}\Phi_{ji}^{mn} &= \Phi_{j}^{m} - \Phi_{j}^{n} - \Phi_{i}^{m} + \Phi_{i}^{n} \\ r_{ji}^{mn} &= r_{j}^{m} - r_{j}^{n} - r_{i}^{m} + r_{i}^{n} \\ N_{ji}^{mn} &= N_{j}^{m} - N_{j}^{n} - N_{i}^{m} + N_{i}^{n} \\ \varepsilon_{\Phi ji}^{mn} &= \varepsilon_{\Phi j}^{m} - \varepsilon_{\Phi j}^{n} - \varepsilon_{\Phi i}^{m} + \varepsilon_{\Phi i}^{n}\end{aligned} \tag{3.161}$$

3.9.3.3 Triple difference

The integer ambiguity terms are the nuisance parameters that still linger on both single and double differenced equations. These can also be eliminate. The integer ambiguities remain fixed as long as the various receivers maintain the carrier lock (i.e. there is no cycle slip), making them time independent. In triple difference we form the difference between the DD at two measurement epochs to eliminate these time independent parameters. For two measurement epochs t_1 and t_2, the DD equations for these epochs can be written as

$$\begin{aligned}\Phi_{ji}^{mn}(t_1) &= r_{ji}^{mn}(t_1) + \lambda N_{ji}^{mn} + \varepsilon_{\Phi ji}^{mn}(t_1)\\ \Phi_{ji}^{mn}(t_2) &= r_{ji}^{mn}(t_2) + \lambda N_{ji}^{mn} + \varepsilon_{\Phi ji}^{mn}(t_2)\end{aligned} \tag{3.162}$$

and the triple difference between these epochs is

$$\Phi_{ji}^{mn}(t_2) - \Phi_{ji}^{mn}(t_1) = r_{ji}^{mn}(t_2) + \lambda N_{ji}^{mn} + \varepsilon_{\Phi ji}^{mn}(t_2) - \left[r_{ji}^{mn}(t_1) + \lambda N_{ji}^{mn} + \varepsilon_{\Phi ji}^{mn}(t_1)\right] \tag{3.163}$$

$$\Phi_{ji}^{mn}(t_{21}) = r_{ji}^{mn}(t_{21}) + \varepsilon_{\Phi ji}^{mn}(t_{21}) \tag{3.164}$$

It should be kept in mind that each term in this equation comprises eight terms as follows

$$\begin{aligned}\Phi_{ji}^{mn}(t_{21}) = \Phi_{ji}^{mn}(t_2) - \Phi_{ji}^{mn}(t_1) =& \Phi_j^m(t_2) - \Phi_j^n(t_2) - \Phi_i^m(t_2) + \Phi_i^n(t_2)\\ &- \Phi_j^m(t_1) + \Phi_j^n(t_1) + \Phi_i^m(t_1) - \Phi_i^n(t_1)\\ r_{ji}^{mn}(t_{21}) = r_{ji}^{mn}(t_2) - r_{ji}^{mn}(t_1) =& r_j^m(t_2) - r_j^n(t_2) - r_i^m(t_2) + r_i^n(t_2)\\ &- r_j^m(t_1) + r_j^n(t_1) + r_i^m(t_1) - r_i^n(t_1)\\ \varepsilon_{\Phi ji}^{mn}(t_{21}) = \varepsilon_{\Phi ji}^{mn}(t_2) - \varepsilon_{\Phi ji}^{mn}(t_1) =& \varepsilon_{\Phi j}^m(t_2) - \varepsilon_{\Phi j}^n(t_2) - \varepsilon_{\Phi i}^m(t_2) + \varepsilon_{\Phi i}^n(t_2)\\ &- \varepsilon_{\Phi j}^m(t_1) + \varepsilon_{\Phi j}^n(t_1) + \varepsilon_{\Phi i}^m(t_1) - \varepsilon_{\Phi i}^n(t_1)\end{aligned} \tag{3.165}$$

Equation (3.164) for the triple difference does not contain the ambiguity terms, which is a great bonus. However, the dilution of precision will be large and the position estimates are usually less accurate than when using the single or double differences. The triple difference approach is helpful in finding the discontinuities in carrier tracking, and can therefore identify the cycle slips.

3.9.3.4 Notes on RSD and DD Linear Combinations

In RSD and DD the ambiguity term is still an integer and we must estimate it for each satellite-receiver pair before we can estimate the baseline vector. In the case of RSD, for M visible satellites we have three baseline coordinates, receiver clock bias and $M - 1$ integer ambiguity parameters that are unknown. Hence the total number of unknowns is $4 + M - 1 = M + 3$ for the first epoch. As the ambiguities remain fixed until a cycle slip occurs, each new epoch will add only one unknown: the receiver's clock bias. Thus by assuming four visible satellites and a stationary user, we need at least three epochs to solve the nine unknowns: three ambiguities, three coordinates and three clock biases. For a similar configuration with DD the total number of unknowns is $M + 2$ for the first epoch, and subsequent epochs do not add unknowns. So we need at least two epochs to calculate six unknowns. In kinematic positioning the user's coordinates change for each epoch and for SD each additional epoch adds

four unknowns, with the result that we need at least six satellites and five epochs to get a solution. Similarly, for DD we need a minimum of five satellites and they should be tracked for at least four epochs. Therefore, in kinematic positioning a different approach is used in which ambiguities are fixed beforehand (e.g. with initialization) so that the solution simplifies to a code phase model.

3.9.4 Position Estimation from Carrier Phase Measurements

Usually carrier phase measurements are utilized in relative positioning where DD is used to cancel out or mitigate the nuisance parameters. The details are beyond the scope of this book, so for these the reader is referred to (Gleason and Gebre-Egziabher 2009). To illustrate carrier phase positioning, a linear model for point positioning based on the carrier phase measurements (Hofmann-Wellenhof et al. 2008; Misra and Enge 2001; Leick 2004) is presented.

As shown earlier, the carrier phase measurement equation can be written as

$$\Phi = r + c\delta t_r - c\delta t_s - I + T + \lambda N + \varepsilon_\Phi \tag{3.166}$$

It may be noted that because we are dealing with point positioning (not relative positioning) we have ignored equipment delay, initial phase offsets and multipath and their effects can either be modeled or ignored. After compensating for satellite clock bias, ionospheric and tropospheric errors, and introducing time epoch t and a superscript for the mth satellite we can write the corrected pseudo-range as

$$\Phi^m(t) = r^m(t) + c\delta t_r(t) + \lambda N^m + \varepsilon_\Phi^m(t) \tag{3.167}$$

where the geometric range is the actual distance between the satellite and the user, which is non-linear and given as

$$r^m = \sqrt{(x - x^m)^2 + (y - y^m)^2 + (z - z^m)^2} \tag{3.168}$$

where

$\mathbf{x} = [x, y, z]^T$ is the receiver position in ECEF frame

$\mathbf{x}^m = [x^m, y^m, z^m]^T$ is theposition of the m^{th} satellite in ECEF frame.

By a similar mathematical treatment to that in Sect. 3.8.2, we can linearize Eq. (3.168) around the current best estimate $\mathbf{x}_{EST} = [x_{EST}, y_{EST}, z_{EST}]^T$ as

$$r^m = \sqrt{(x_{EST} - x^m)^2 + (y_{EST} - y^m)^2 + (z_{EST} - z^m)^2} + \frac{(x_{EST} - x^m)(x - x_{EST}) + (y_{EST} - y^m)(y - y_{EST}) + (z_{EST} - z^m)(z - z_{EST})}{\sqrt{(x_{EST} - x^m)^2 + (y_{EST} - y^m)^2 + (z_{EST} - z^m)^2}} \tag{3.169}$$

The estimated geometric range and the receiver's position can be defined as

$$r_{EST}^{m} = \sqrt{(x_{EST} - x^{m})^{2} + (y_{EST} - y^{m})^{2} + (z_{EST} - z^{m})^{2}} \tag{3.170}$$

$$x = x_{EST} + \delta x\,,\; y = y_{EST} + \delta y\,,\; z = z_{EST} + \delta z \tag{3.171}$$

Substituting the above definitions into Eq. (3.169) gives

$$r^{m} = r_{EST}^{m} + \frac{(x_{EST} - x^{m})\delta x + (y_{EST} - y^{m})\delta y + (z_{EST} - z^{m})\delta z}{r_{EST}^{m}} \tag{3.172}$$

and inserting this into Eq. (3.167) provides

$$\begin{aligned}\Phi^{m}(t) =& r_{EST}^{m}(t) + \frac{[x_{EST} - x^{m}(t)]\delta x + [y_{EST} - y^{m}(t)]\delta y + [z_{EST} - z^{m}(t)]\delta z}{r_{EST}^{m}(t)} \\ &+ c\delta t_{r}(t) + \lambda N^{m} + \varepsilon_{\Phi}^{m}(t)\end{aligned} \tag{3.173}$$

By transferring the known quantities over to the left-hand side and defining $b_r = c\delta t_r(t)$

$$\begin{aligned}\Phi^{m}(t) - r_{EST}^{m}(t) =& \frac{[x_{EST} - x^{m}(t)]\delta x + [y_{EST} - y^{m}(t)]\delta y + [z_{EST} - z^{m}(t)]\delta z}{r_{EST}^{m}(t)} \\ &+ b_{r}(t) + \lambda N^{m} + \varepsilon_{\Phi}^{m}(t)\end{aligned} \tag{3.174}$$

The estimated line of sight unit vector from the mth satellite to the receiver can be introduced as

$$\mathbf{1}_{EST}^{m} = \frac{[(x_{EST} - x^{m}), (y_{EST} - y^{m}), (z_{EST} - z^{m})]^{T}}{\sqrt{(x_{EST} - x^{m})^{2} + (y_{EST} - y^{m})^{2} + (z_{EST} - z^{m})^{2}}} \tag{3.175}$$

enabling us to rewrite Eq. (3.174) as

$$\Phi^{m}(t) - r_{EST}^{m}(t) = \left(\mathbf{1}_{EST}^{m}\right)^{T}\delta\mathbf{x} + b_{r}(t) + \lambda N^{m} + \varepsilon_{\Phi}^{m}(t) \tag{3.176}$$

Notice that in comparison to the pseudo-range, the number of unknowns has increased by the integer ambiguities.

By defining $y^{m} = \Phi^{m}(t) - r_{EST}^{m}(t)$ and $\mathbf{x} = [\delta x, \delta y, \delta z, N^{m}, b_{r}(t)]$ we get

$$y^{m} = \begin{bmatrix} 1_{x,EST}^{m} & 1_{y,EST}^{m} & 1_{z,EST}^{m} & \lambda & 1 \end{bmatrix}\mathbf{x} + \varepsilon_{\Phi}^{m}(t) \tag{3.177}$$

and by further defining $\mathbf{a} = \begin{bmatrix} 1^m_{x,EST} & 1^m_{y,EST} & 1^m_{z,EST} & \lambda & 1 \end{bmatrix}$ we get

$$y^m = \mathbf{a}\mathbf{x} + \boldsymbol{\varepsilon}^{\boldsymbol{m}}_{\boldsymbol{\Phi}}(\boldsymbol{t}) \tag{3.178}$$

and for multiple satellites

$$\mathbf{y}^m = \mathrm{A}\mathbf{x} + \boldsymbol{\varepsilon}^{\boldsymbol{m}}_{\boldsymbol{\Phi}}(\boldsymbol{t}) \tag{3.179}$$

where $m = 1, 2, \ldots, M$.

For four visible satellites, the terms in this equation can be written as

$$\mathbf{y} = \begin{bmatrix} \Phi^1(t) - r^1_{EST}(t) \\ \Phi^2(t) - r^2_{EST}(t) \\ \Phi^3(t) - r^3_{EST}(t) \\ \Phi^4(t) - r^4_{EST}(t) \end{bmatrix} \tag{3.180}$$

$$\mathbf{A} = \begin{bmatrix} 1^1_{x,EST}(t) & 1^1_{y,EST}(t) & 1^1_{z,EST}(t) & \lambda & 0 & 0 & 0 & 1 \\ 1^2_{x,EST}(t) & 1^2_{y,EST}(t) & 1^2_{z,EST}(t) & 0 & \lambda & 0 & 0 & 1 \\ 1^3_{x,EST}(t) & 1^3_{y,EST}(t) & 1^3_{z,EST}(t) & 0 & 0 & \lambda & 0 & 1 \\ 1^4_{x,EST}(t) & 1^4_{y,EST}(t) & 1^4_{z,EST}(t) & 0 & 0 & 0 & \lambda & 1 \end{bmatrix} \tag{3.181}$$

$$\mathbf{x} = \left[\delta x, \delta y, \delta z, N^1, N^2, N^3, N^4, b_r(t)\right]^T \tag{3.182}$$

In the above system of equations, there are eight known values and only four equations that cannot be solved for a single epoch. For a stationary user, another epoch will add a further receiver clock bias term, yielding eight equations and nine unknowns, so we need at least three epochs to obtain a dozen equations with ten unknowns that can be solved by using the least squares approach. In that case the dimensions of matrix A will be 12 × 10.

For a moving user, each epoch will introduce four unknowns (i.e. three for the user's coordinates and one for the receiver's clock bias) and we need at least five satellites and five epochs to obtain 25 equations with 25 unknowns which can be resolved.

3.10 Integer Ambiguity

With GPS, positioning accuracy at meter to centimeter level is possible depending on the type of measurements and the methods used. It is widely accepted that the carrier phase observable is the most precise measurement, with a resolution of 0.2–1 mm. Use of the carrier phase measurements with DGPS enables centimeter level positioning accuracy. This is only possible when the IAs are accurately resolved to their integer values. Efficient resolution of IA requires that the errors in the range measurements be made as small as possible. As a rule of thumb, instant

ambiguity resolution (AR) can be achieved if the range measurements are accurate to half the wavelength of the ambiguities being resolved. For the kinematic applications, this is required to be done as rapidly as possible. As a result of ongoing research many methods have been developed to resolve IA. These include the *ambiguity function method* (AFM) (Hofmann-Wellenhof et al. 2008), *least squares ambiguity search technique* (LSAST) (Counselman and Gourevitch 1981), *fast ambiguity search filter* (FASF) (Hatch 1990), and *least squares ambiguity decorrelation adjustment* (LAMBDA) (Chen 1994). The difficulty of resolving IA increases with the length of the baseline between the base and rover receivers because the measurements become decorrelated and errors increase.

Once the initial IA has been resolved, it stays constant and GPS can keep track of the partial phase as well as the total number of cycles that have passed since the GPS began to track the satellite. But if the signal becomes weaker or is obstructed and the tracking is lost (a condition known as *cycle slip*) then the receiver loses its count of the number of complete cycles and must derive IA afresh. If this occurs while using a long baseline the residual measurement errors increase, as does the uncertainty in the GPS position. If these errors can be mitigated or the estimate of the position can be improved, then IA can be resolved relatively quickly and more reliably.

The detail of actual AR methods is beyond the scope of this book; however, a general summary of the steps involved will be given. For a more detailed account the reader should consult a dedicated text such as (Teunissen and Tiberius 1994). Usually DD carrier phase measurements are used for relative positioning, as they remove or mitigate most of the nuisance parameters. Hence it is the ambiguity in the DD measurements that must be resolved.

3.10.1 Integer Ambiguity Resolution

In general, for ambiguity resolution (AR) the GNSS data processing is performed in three different steps (Hofmann-Wellenhof et al. 2008; Kleusberg and Teunissen 1998; Leick 2004; Misra and Enge 2001).

Step I: Obtaining the float solution

In this step no distinction is made between ambiguities and other parameters (e.g. baseline coordinates and atmospheric delays). Parameters are estimated without taking into account the special integer nature of the ambiguities. This solution is often referred to as the *float* solution. The parameters are estimated using either a least squares algorithm or a Kalman filter because the inconsistencies in the data are due to forms of noise. It may be noted that the *integerness* of the ambiguities has not yet been exploited. After the fixed ambiguities are found, it is important to determine whether they are the proper values or not. The process of assessing the correctness of integer values is called ambiguity validation.

Step II: Obtaining fixed ambiguities

In the second step, the integer ambiguities are estimated from the float ambiguities by exploiting their integer nature. There are numerous ways of finding the integer ambiguities. These range from simple rounding schemes to advanced methods that use clever integer search algorithms. One popular and well documented method is LAMBDA.

Step III: Obtaining the fixed solution

Finally, the estimated IA values are used to improve the float solution of the first-step where other parameters (e.g. baseline coordinates and atmospheric delays) are recomputed to obtain better accuracy. This final solution is usually called the fixed solution because it is much more precise than the float solution.

3.10.2 Ambiguity Dilution of Precision

The covariance matrix of the least squares ambiguities provides insight into their stochastic nature. However, it is desired to have a simpler quantity which sums up the important characteristics of these ambiguities. For example, DOP measures are used to show how the satellite-receiver geometry affects a specific parameter (e.g. vertical position, horizontal position, time) with scalar quantities which are simple functions of the relevant variance–covariance matrices. In (Joosten and Tiberius 2000) Teunissen proposed using the geometric mean of the ambiguity conditional standard deviations as a measure of dilution of precision and this is referred to as the ambiguity DOP (ADOP).

The DOP measures utilized for GPS positioning are based on the trace of the relevant variance covariance matrix, but the same could not be done because (1) the trace does not account for the invariance properties and (2) it uses only the diagonal entries of the variance–covariance matrix. Hence the ADOP parameter is based on the determinant of the covariance matrix $\Lambda_{\hat{N}}$ and is given as

$$ADOP = \sqrt{\left|\Lambda_{\hat{N}}\right|}^{1/n} \text{ (cycle)} \tag{3.183}$$

where n is the order of ambiguity covariance matrix. ADOP is invariant under volume, preserving ambiguity transformations and choice of reference satellite for DD, so it can be used to compute the volume of the ambiguity search space. For detailed properties of ADOP the reader is referred to (Teunissen 1997).

References

Langley RB (Jun 1990) Why is GPS Signal So Complex. GPS World

Klobuchar JA (1987) Ionospheric time-delay algorithm for single-frequency GPS users. IEEE Trans Aerosp Electron Syst AES-23 (3):325–331

GPS Augmentations U.S. Coast Guard Navigation Center (2009) www.gps.gov/systems/augmentations/index.html. Accessed 22 Aug 2009
Bate RR, Mueller DD, White JE (1971) Fundamentals of astrodynamics. Dover Publications, New York
Battin RH (1987) An introduction to the mathematics and methods of astrodynamics. AIAA, New York
Chao CC (1971) The Tropospheric calibration model for Mariner mars. Jet Propulsion Laboratory, California Institute of Technology, Pasandena
Chen D (1994) Development of a fast ambiguity search filtering/FASF) method for GPS carrier phase ambiguity resolution. Ph D. dissertation, University, Calgary
Counselman CG, Gourevitch SA (1981) Miniature interferometer terminals for Earth surveying: ambiguity and multipath with global positioning system. IEEE Trans Geosci Remote Sens GE-19(4):244–252
Crosby GK, Ely WBS, McPherson KW, Stewart JM, Kraus DK, Cashin TP, Bean KW, Elrod BD (2000) A ground-based regional augmentation system (GRAS)—The Australian proposal. In: 13th International technical meeting of the satellite division of the institute of navigation (ION GPS 2000), Salt Lake City, Sep 19–22, 2000 Sep 2000
El-Rabbany A (2002) Introduction to GPS—The global positioning system. Mobile communication series. Artech House Inc, Norwood
Essentials of Satellite Navigation (2007) A Compendium by Ublox
Farrell JA (1998) The global positioning system & inertial navigation. McGraw-Hill, New York
Farrell JA (2008) Aided navigation: GPS with high rate sensors. McGraw-Hill, New York
Diggelen FV (2009) A-GPS assisted GPS, GNSS, and SBAS. Artech House, Boston
Gleason S, Gebre-Egziabher D (2009) GNSS applications and methods. Artech House, Boston
Government US (2012) The Global Positioning System. U.S. Government. www.gps.gov/systems/gps/. Accessed Feb 2012 2012
Hatch R (1990) Instantaneous ambiguity resolution. In: Kinematic systems in geodesy, surveying, and remote sensing, KIS symposium, Banff, 1990. pp 299–308
Hofmann-Wellenhof B, Lichtenegger H, Wasle E (2008) GNSS-global navigation satellite systems: GPS, GLONASS, Galileo, and more. Springer, New York
Hopfield HS (1969) Two-Quartic tropospheric refractivity profile for correcting satellite data. J Geophys Res J Geophys Res 74(18):4487–4499
IS-GPS-200F (2011) Interface Specification IS-GPS-200F. ARINC Engineering Systems Center, El Segundo, CA
Joosten P, Tiberius C (2000) Fixing the ambiguities: are you sure they're right? GPS World, vol 11
Kaplan ED, Hegarty CJ (2006) Understanding GPS principles and applications, 2nd edn. Artech House, Boston
Kleusberg A, Teunissen PJG (1998) GPS for geodesy, 2nd edn. Springer, Berlin
Leick A (2004) GPS satellite surveying. Wiley, Hoboken
McLellan JA (1992) Design and analysis of a low cost GPS aided navigation system. M. Eng. Dissertation, University of Calgary, Calgary, AB
McPherson K (2001) A ground-based regional augmentation system (GRAS). Navigation 49(193):19–25
Misra P, Enge P (2001) Global positioning system: signals, measurements, and performance. Ganga-Jamuna Press, Lincoln
Parkinson BW, Spilker JJ, Axelrad P, Enge P (1996) Global positioning system: theory and application, vol I. American Institute of Aeronautics and Astronautics, Washington
Prasad R, Ruggieri M (2005) Applied Satellite Navigation using GPS, GALILEO, and Augmentation systems. Artech House
Seeber G (2003) Satellite geodesy, 2nd edn. de Gruyter, Berlin

Teunissen PJG (1997) A canonical theory for short GPS baselines. Part IV: precision versus reliability. J Geodesy 71(9):513–525. doi:10.1007/s001900050119

Teunissen PJG (1994) Tiberius CCJM Integer least-squares estimation of the GPS phase ambiguities. In: Proceedings of international symposium on kinematic systems in geodesy, Geomatics and navigation KIS'94, Banff, pp 221–231

Chapter 4
Inertial Navigation System

An inertial navigation system is an autonomous system that provides information about position, velocity and attitude based on the measurements by inertial sensors and applying the dead reckoning (DR) principle. DR is the determination of the vehicle's current position from knowledge of its previous position and the sensors measuring accelerations and angular rotations. Given specified initial conditions, one integration of acceleration provides velocity and a second integration gives position. Angular rates are processed to give the attitude of the moving platform in terms of pitch, roll and yaw, and also to transform navigation parameters from the body frame to the local-level frame.

4.1 Principle of Inertial Navigation

The principle of inertial navigation is based upon Newton's first law of motion, which states

> A body continues in its state of rest, or uniform motion in a straight line, unless it is compelled to change that state by forces impressed on it.

Put simply, this law says that a body at rest tends to remain at rest and a body in motion tends to remain in motion unless acted upon by an outside force. The full meaning of this is not easily visualized in the Earth's reference frame. For it to apply, the body must be in an inertial reference frame (a non-rotating frame in which there are no inherent forces such as gravity).

Newton's second law of motion shares importance with his first law in the inertial navigation system, and states

> Acceleration is proportional to the resultant force and is in the same direction as this force.

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_4,

This can be expressed mathematically as

$$F = ma \tag{4.1}$$

where

F is the force
m is the mass of the body
a is the acceleration of the body due to the applied force F.

The physical quantity pertinent to an inertial navigation system is acceleration, because both velocity v and displacement s can be derived from acceleration by the process of integration. Conversely, velocity and acceleration can be estimated by differentiation from displacement, written mathematically

$$v = \frac{ds}{dt}; a = \frac{dv}{dt} = \frac{d^2s}{dt^2} \tag{4.2}$$

Differentiation is the process of determining how one physical quantity varies with respect to another. Integration, the inverse of differentiation, is the process of summing all rate-of-change that occurs within the limits being investigated, which can be written mathematically as

$$v = \int adt; s = \int vdt = \int \int adtdt \tag{4.3}$$

An inertial navigation system is an integrating system consisting of a detector and an integrator. It detects acceleration, integrates this to derive the velocity and then integrates that to derive the displacement. By measuring the acceleration of a vehicle in an inertial frame of reference and then transforming it to the navigation frame and integrating with respect to time, it is possible to obtain velocity, attitude and position differences. Measurement of the vehicle's rotation is needed for the transformation from the inertial to the navigation frame and for the computation of the attitude of the vehicle.

4.2 Physical Implementation of an INS

There are two implementation approaches to an INS: (1) a stable platform system also known as a gimbaled system, and (2) a strapdown system. The components of these systems are shown in Fig. 4.1. In the stable platform, the inertial sensors are mounted on a set of gimbals such that the platform always remains aligned with the navigation frame. This is done by having a set of torque motors rotate the platform in response to rotations sensed by the gyroscopes. Thus the output of the accelerometers is directly integrated for velocity and position in the navigation frame. Since gimbaled systems are mechanically complex and expensive, their use is limited.

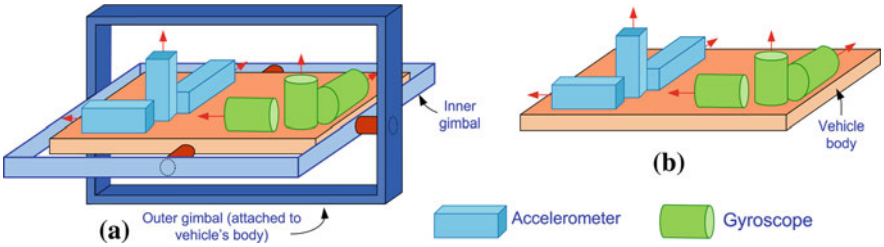


Fig. 4.1 Arrangement of the components of a gimbaled IMU (*left*) and a strapdown IMU (*right*)

Table 4.1 Comparison of gimbaled platform and strapdown navigation systems

Characteristics	Strapdown systems	Gimbaled systems
Size	Relatively small	Bigger
Weight	Relatively lighter	Heavy
Performance	High accuracy	Superior performance
Robustness	Highly reliable, immune to shocks and vibrations	High reliability, low immunity to shocks and vibrations

Advances in electronics gave rise to strapdown systems. In these, the inertial sensors are rigidly mounted onto the body of the moving platform and the gimbals are replaced by a computer that simulates the rotation of the platform by software frame transformation. Rotation rates measured by the gyroscopes are applied to continuously update the transformation between the body and navigation frames. The accelerometer measurements are then passed through this transformation to obtain the acceleration in the navigation frame. Strapdown systems are favored for their reliability, flexibility, low power usage, being lightweight and less expensive than stable platforms. The transition to strapdown systems was facilitated by the introduction of optical gyros to replace rotor gyros, and by the rapid development of the processor technology required to perform the computations. Table 4.1 gives a comparison of the major characteristics of the two systems.

An INS can be thought of as consisting of three principal modules: an inertial measurement unit (IMU), a pre-processing unit, and a mechanization module. An IMU uses three mutually orthogonal accelerometers and three mutually orthogonal gyroscopes. The signals are pre-processed by some form of filtering to eliminate disturbances prior to the mechanization algorithm which converts the signals into positional and attitude information. The three major modules of an INS are shown in Fig. 4.2.

4.3 Inertial Measurement Unit

The measurements of the acceleration and the rotation of the vehicle are made by a suite of inertial sensors mounted in a unit called the inertial measurement unit (IMU). This holds two orthogonal sensor triads, one with three accelerometers and

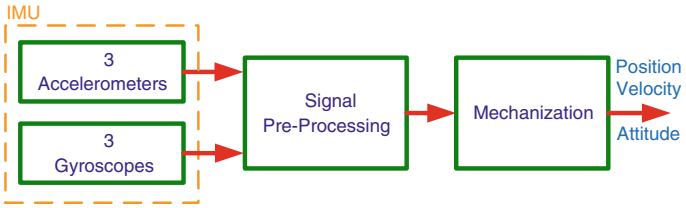


Fig. 4.2 The principal modules of an inertial navigation system

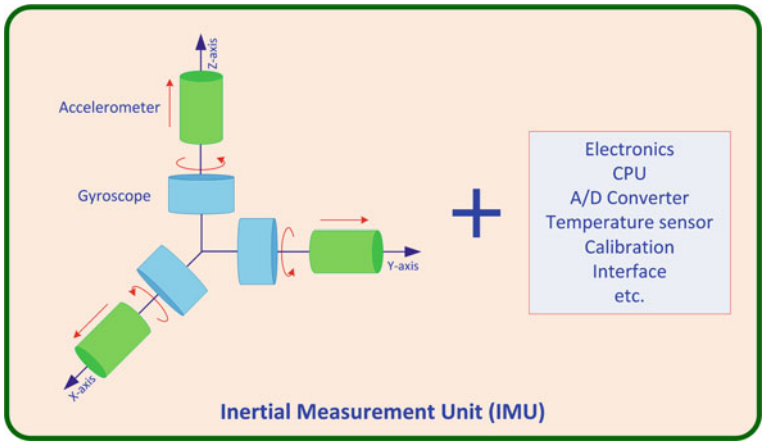


Fig. 4.3 The components of a typical inertial measurement unit (*IMU*)

the other with three gyroscopes. Accelerometers measure linear motion in three mutually orthogonal directions, whereas gyroscopes measure angular motion in three mutually orthogonal directions. Nominally, the axes of these two triads are parallel, sharing the origin of the accelerometer triad. The sensor axes are fixed in the body of the IMU, and are therefore called the body axes or body frame. Apart from the inertial sensors, the IMU also contains related electronics to perform self-calibration, to sample the inertial sensor readings and then to convert them into the appropriate form for the navigation equipment and algorithms. Figure 4.3 shows the components of a typical IMU.

4.4 Inertial Sensors

A brief description of the two main kinds of inertial sensors, accelerometers and gyroscopes, now follows.

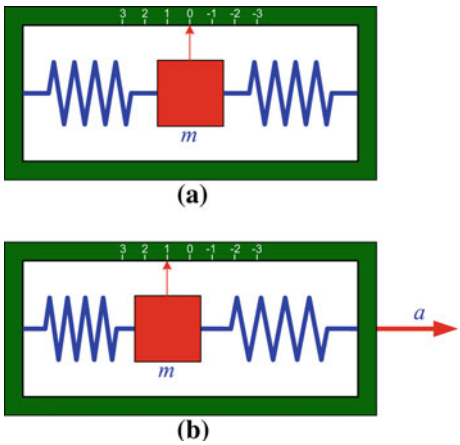


Fig. 4.4 **a** An accelerometer in the null position with no force acting on it, **b** the same accelerometer measuring a linear acceleration of the vehicle in the positive direction (*to the right*)

4.4.1 Accelerometers

An accelerometer consists of a proof mass, m, connected to a case by a pair of springs as shown in Fig. 4.4. In this case the sensitive axis of the accelerometer is along the spring in the horizontal axis. Acceleration will displace the proof mass from its equilibrium position, with the amount of displacement proportional to the acceleration. The displacement from the equilibrium position is sensed by a pick-off and is then scaled to provide an indication of acceleration along this axis. The equilibrium position is calibrated for zero acceleration. Acceleration to the right will cause the proof mass to move left in relation to the case and (as shown by the scale) indicates positive acceleration.

If the accelerometer is stood on a bench with its sensitive axis vertical in the presence of a gravitational field, the proof mass will be displaced downward with respect to the case, indicating positive acceleration. The fact that the gravitational acceleration is downward, in the *same direction* as the displacement as shown in Fig. 4.5, is sometimes a cause of confusion for the beginners in navigation.

The explanation for this lies in the equivalence principle, according to which, in the terrestrial environment it is not possible to separate inertia and navigation by the accelerometer measurement in a single point. Therefore, the output of an accelerometer due to a gravitational field is the negative of the field acceleration. The output of an accelerometer is called the specific force, and is given by

$$f = a - g \tag{4.4}$$

where

f is the specific force
a is the acceleration with respect to the inertial frame
g is the gravitational acceleration which is +9.8 m/s^2.

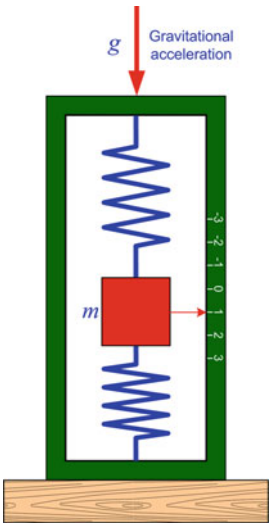


Fig. 4.5 An accelerometer resting on a bench with gravitational acceleration acting on it

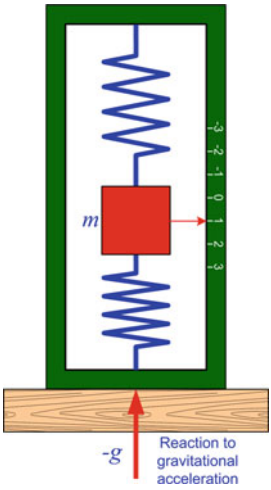


Fig. 4.6 An accelerometer resting on a bench where reaction to the gravitational acceleration is acting on it

It is this which causes confusion. The easy way to remember this relation is to think of one of two cases. If the accelerometer is sitting on a bench it is at rest so acceleration a is zero. The force on the accelerometer is the force of reaction of the bench against the case, which is the negative of g along the positive (upward) direction and therefore causes the mass to move downward (Fig. 4.6).

Or imagine dropping the accelerometer in a vacuum. In this case the specific force read by the accelerometer f is zero and the actual acceleration is $a = g$. To navigate with respect to the inertial frame we need a, therefore in the navigation equations we convert the output of an accelerometer from f to a by adding g.

4.4.1.1 Accelerometer Measurements

An accelerometer measures translational acceleration (less the gravity component) along its sensitive axis typically by sensing the motion of a proof mass relative to the case. From Eq. (4.4) the output of an accelerometer triad is

$$\mathbf{f} = \mathbf{a} - \mathbf{g} \tag{4.5}$$

where $\mathbf{f}$ is the specific force vector, $\mathbf{a}$ is the acceleration vector of the body, and $\mathbf{g}$ is the gravitational vector. The acceleration $\mathbf{a}$ can be expressed as the double derivative of the position vector $\mathbf{r}$, as

$$\mathbf{a} = \left.\frac{d^2\mathbf{r}}{dt^2}\right|_i = \ddot{\mathbf{r}} \tag{4.6}$$

The gravitational field vector was earlier shown to be related to the gravity vector as

$$\bar{\mathbf{g}} = \mathbf{g} + \Omega_{ie}\Omega_{ie}\mathbf{r} \tag{4.7}$$

where Ω_{ie} is the skew-symmetric matrix representing the rotation of the Earth in the inertial frame.

Substituting Eqs. (4.6) and (4.7) into Eq. (4.5) provides

$$\mathbf{f} = \left.\frac{d^2\mathbf{r}}{dt^2}\right|_i - \mathbf{g} - \Omega_{ie}\Omega_{ie}\mathbf{r} \tag{4.8}$$

4.4.2 Gyroscopes

To fully describe the motion of a body in 3-D space, rotational motion as well as translational motion must be measured. Sensors which measure angular rates with respect to an inertial frame of reference are called gyroscopes. If the angular rates are mathematically integrated this will provide the change in angle with respect to an initial reference angle. Traditionally, these rotational measurements are made using the angular momentum of a spinning rotor. The gyroscopes either output angular rate or attitude depending upon whether they are of the rate sensing or rate integrating type. It is customary to use the word *gyro* as a short form of the word *gyroscope*, so in the ensuing treatment these words are used interchangeably.

4.4.2.1 Gyroscope Measurements

Gyros measure the angular rate of a body with respect to the navigation frame, the rotation of the navigation frame with respect to the Earth-fixed frame (as it traces

the curvature of the Earth), and the rotation of the Earth as it spins on its axis with respect to inertial space. These quantities are all expressed in the body frame and can be given as

$$\omega_{ib}^{b} = \omega_{ie}^{b} + \omega_{en}^{b} + \omega_{nb}^{b} \tag{4.9}$$

where

ω_{ib}^{b} is the rotation rate of the body with respect to the i-frame
ω_{nb}^{b} is the rotation rate of the body with respect to the navigation frame (also referred to as the n-frame)
ω_{en}^{b} is the rotation rate of the navigation frame with respect to the e-frame
ω_{ie}^{b} is the rotation rate of the Earth with respect to the i-frame.

Traditional gyroscopes were mechanical and based on angular momentum, but more recent ones are based on either the Coriolis effect on a vibrating mass or the Sagnac interference effect. There are three main types of gyroscope (Lawrence 1998): mechanical gyroscopes, optical gyroscopes, and micro-electro-mechanical system (MEMS) gyroscopes.

4.5 Basics of Inertial Navigation

As mentioned before, inertial positioning is based on the simple principle that differences in position can be determined by a double integration of acceleration, sensed as a function of time in a well-defined and stable coordinate frame. Mathematically, we can express this as

$$\Delta P(t) = P(t) - P(t_o) = \int_{t_0}^{t} \int_{t_0}^{t} a(t) dt dt \tag{4.10}$$

where

$P(t_o)$ is the initial point of the trajectory
$a(t)$ is the acceleration along the trajectory obtained from inertial sensor measurements in the coordinate frame prescribed by $P(t)$.

Next, we shall consider examples of navigation in one and two dimensions. An overview of three-dimensional navigation will be given as a preview of the more detailed treatment provided in later chapters.

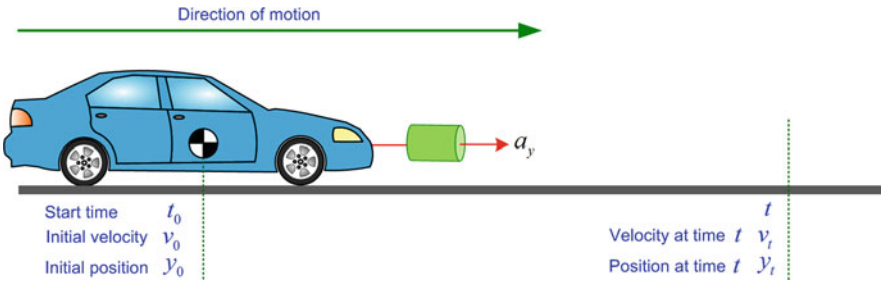


Fig. 4.7 One-dimensional (*1D*) inertial navigation, with the green cylinder depicting the accelerometer

4.5.1 Navigation in One Dimension

To comprehend the full scale three-dimensional inertial system it is easier to start with an example of a one-dimensional (1D) inertial system with a single axis. For this, consider a vehicle moving in a straight line (i.e. in a fixed direction) as shown in Fig. 4.7. To calculate its velocity and position, which are the only unknowns in this case, we need only a single accelerometer mounted on the vehicle that has its sensitive axis along the direction of motion.

With prior knowledge of the initial position $y = y_0$ and initial velocity $v = v_0$ of the vehicle, we are able to calculate its velocity v_t at any time t by integrating the output of the accelerometer a_y as follows

$$v_t = \int a_y dt = a_y t + v_0 \tag{4.11}$$

A second integration will yield the position y_t of the vehicle at time t

$$\begin{aligned} y_t &= \int v_t dt \\ y_t &= \int (a_y t + v_0) dt \\ y_t &= \frac{1}{2} a_y t^2 + v_0\, t + y_0 \end{aligned} \tag{4.12}$$

4.5.2 Navigation in Two Dimensions

Extending the concept of navigation from the simple 1D example to 2D makes the implementation more complex, mainly because we need the acceleration to be in the same frame as the coordinate system. This requires the transformation of the acceleration measured by the accelerometers from the INS frame to a stable Earth-

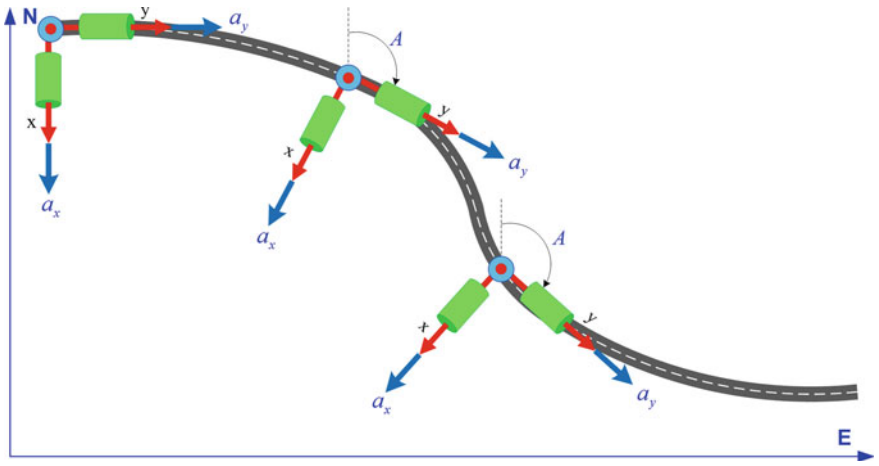


Fig. 4.8 Inertial navigation using a 2D strapdown system

fixed coordinate frame. The stable Earth-fixed coordinate frame is often chosen as a local-level frame that is referred to as the navigation frame. As stated earlier, the transformation can either be established mechanically inside the INS by a stable platform or numerically as in the strapdown concept.

In 2D it is necessary to monitor both the translational motion of the vehicle in two directions and also its rotational motion, manifested as a change in direction. Two accelerometers are required to detect the acceleration in two directions. One gyroscope is required to detect the rotational motion in a direction perpendicular to the plane of motion (for simplicity, we neglect the Earth's rotation which would also be detected). Based on the advantages provided by a strapdown system, from this point on we shall limit our discussion to this type of system.

Strapdown systems mathematically transform the output of the accelerometers attached to the body into the east-north coordinate system (the 2D form of ENU) prior to performing the mathematical integration. These systems use the output of the gyroscope attached to the body to continuously update the transformation that is utilized to convert from body coordinates to east-north coordinates. Figure 4.8 shows the concept of inertial navigation in 2D as a platform makes turns, rotating through an angle A (called the azimuth angle[1]) measured from north. The blue cylindrical objects depict the accelerometers, the gyroscope is a blue disc whose sensitive axis depicted by a red dot points out of the paper towards the reader.

The accelerometers measure the acceleration of the body axes (X and Y) but we need the acceleration in the east-north coordinate system. This is accomplished using a transformation matrix which can be explained with the help of the diagram shown in Fig. 4.9.

[1] The terms azimuth angle and yaw angle are both used to represent the deviation from north. The difference lies in the direction of measurement: the azimuth angle is measured clockwise from north whereas the yaw angle is measured counter clockwise.

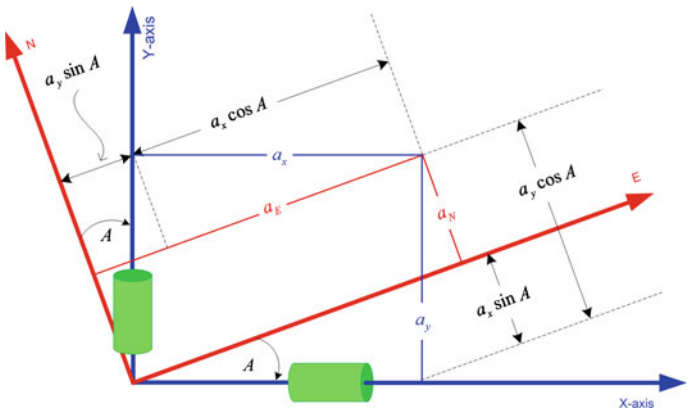


Fig. 4.9 Transformation from the vehicle frame (X-Y) to the navigation frame (E-N)

The vehicle axes X and Y make an angle A with the east and north directions respectively, and the accelerations along east direction a_E and the north direction a_N can be written as

$$a_E = a_y \sin A + a_x \cos A \tag{4.13}$$

$$a_N = a_y \cos A - a_x \sin A \tag{4.14}$$

which in the matrix form is

$$\begin{bmatrix} a_E \\ a_N \end{bmatrix} = \begin{bmatrix} \cos A & \sin A \\ -\sin A & \cos A \end{bmatrix} \begin{bmatrix} a_x \\ a_y \end{bmatrix} \tag{4.15}$$

and can be expressed more compactly as

$$\mathbf{a}^n = R_b^n \mathbf{a}^b \tag{4.16}$$

where

- $\mathbf{a}^n$ is the acceleration in the navigation frame (E-N)
- $\mathbf{a}^b$ is the acceleration in the body frame measured by the accelerometers
- R_b^n is the rotation matrix which rotates $\mathbf{a}^b$ to the navigation frame.

Given the accelerations in the navigation frame, we can integrate to obtain the velocities

$$\begin{aligned} v_E &= \int \left(a_x \cos A + a_y \sin A\right) dt \\ v_N &= \int \left(a_y \cos A - a_x \sin A\right) dt \end{aligned} \tag{4.17}$$

and again to obtain the position in the navigation frame

$$\begin{aligned} x_E &= \int\int \left(a_x \cos A + a_y \sin A\right) dtdt \\ x_N &= \int\int \left(a_y \cos A - a_x \sin A\right) dtdt \end{aligned} \tag{4.18}$$

which in the matrix form is

$$\begin{pmatrix} x_E \\ x_N \end{pmatrix} = \int\int \begin{pmatrix} \cos A & \sin A \\ -\sin A & \cos A \end{pmatrix} \begin{pmatrix} a_x \\ a_y \end{pmatrix} dtdt \tag{4.19}$$

It may be noted that this whole process is dependent on knowing the azimuth angle A which is calculated from the measurement by the gyroscope that monitors angular changes of the orientation of the accelerometers from the local E-N frame. These angular changes resolve the accelerometer measurements from the sensor axes into the local E-N axes. This angular change also determines the direction of motion of the moving platform defined by the azimuth angle, which is also known as the heading angle because it is the deviation from the north direction in the E-N plane. This is based on mathematically integrating the gyroscope angular velocity measurements relative to the initial azimuth angle A_0 as follows

$$A(t) = \int \omega_{gyro} dt + A_o \tag{4.20}$$

In this equation it should be noted (as pointed out previously) that the Earth's rotation components have been neglected for simplicity and ease of understanding of the basic concept of navigation.

4.6 Navigation in Three Dimensions

Inertial navigation in three dimensions (3D) requires three gyroscopes to measure the attitude angles of the body (pitch, roll and azimuth) and three accelerometers to measure accelerations along the three axes (in the east, north and up directions). Another complication is the involvement of gravity in the accelerations. The total acceleration encountered by the body is what is measured by the accelerometers, a combination of the acceleration due to gravity and that due to all other external forces. In order to remove the component of acceleration due to gravity, the tilt (or attitude) of the accelerometer with respect to the local vertical must be supplied by the gyroscope. At this point we will summarize the general concept of 3D inertial navigation. The mathematical details will be presented in a later chapter, where knowledge of 2D navigation will assist in understanding the subject.

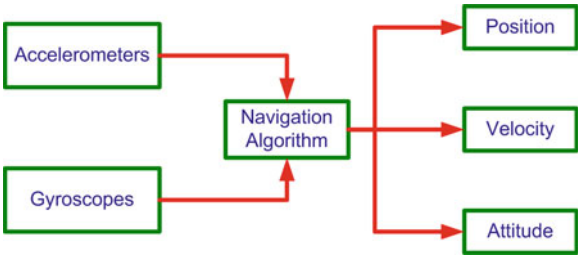


Fig. 4.10 The concept of inertial navigation

4.7 Overview of an Inertial Navigation System in 3D

The operation of an INS is based on processing the inertial sensor measurements received at its input and yielding a set of navigation parameters (position, velocity and attitude) of the moving platform at its output. In general these parameters are determined in a certain reference frame. Figure 4.10 shows the general concept of inertial navigation system.

The accelerometers are attached to the moving platform in order to monitor its accelerations in three mutually orthogonal directions. The gyroscopes provide the attitude (pitch, roll and azimuth) of the moving platform, and their measurements are used to rotate the data from the accelerometers into the navigation frame. The time integral of each acceleration component gives a continuous estimate of the corresponding velocity component of the platform relative to the initial velocities. A second integration yields the position with respect to a known starting point in a given frame of reference. This principle is outlined in Fig. 4.11.

4.8 Theoretical Measurements of the Inertial Sensor

Before delving into the details of inertial navigation and the errors associated with inertial sensors, it is important to look rather closely at the measurements taken by the accelerometer and gyroscope triads. To assist understanding, we will deal with the stationary and moving cases separately. Since the l-frame is more commonly used for everyday navigation (for reasons that will be described in Chap. 5) the ENU frame (a type of l-frame) will be used in this section where required.

4.8.1 Theoretical Measurements of a Stationary Accelerometer Triad

Consider the case where the accelerometer triad is stationary and level with the ground, as shown in Fig. 4.12. Since the accelerometers are stationary the only acting force will be the Earth's gravity (or more correctly the reaction to the force of gravity).

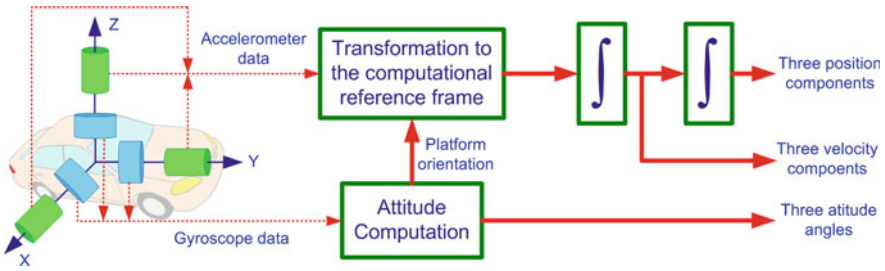


Fig. 4.11 The general principle of inertial navigation in 3D

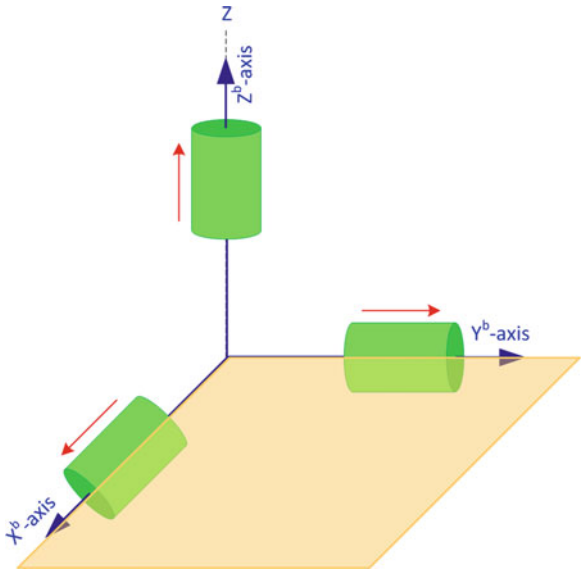


Fig. 4.12 An accelerometer triad that is level with the ground

In this case the accelerometers pointing in the x, y directions will not measure anything and the accelerometer in z direction measures the reaction to the gravity vector g. The nominal measurements will therefore be

$$f_x = 0; f_y = 0; f_z = g$$

It should be noted that the actual measurements will also have some errors (as we shall see later in Sect. 4.11.)

Now consider the case where the accelerometer triad is stationary but this time has been rotated about its x-axis to make an angle p with the ground, as shown in Fig. 4.13. As a consequence, the z-axis is inclined at the same angle from its previous position denoted as a dotted line Z into a new position shown as a dotted line Z'. As was defined in Chap. 2, the angle p is called the pitch angle. In this orientation, the accelerometers in the y, z directions will each measure a portion of the gravity vector.

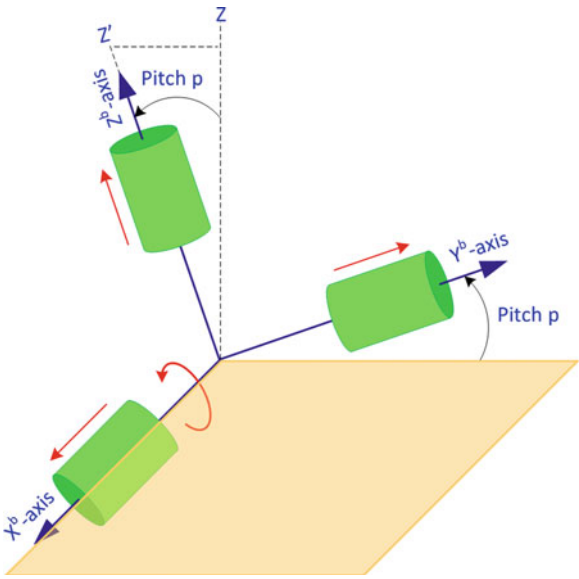


Fig. 4.13 An accelerometer triad with the y-axis making an angle p (*called pitch*) with the level ground

According to the geometry of the Fig. 4.13, the measurements of the x and y accelerometers are

$$\begin{aligned} f_x &= 0 \\ f_y &= g\sin(p) \\ f_z &= g\cos(p) \end{aligned} \tag{4.21}$$

Now rotate the sensor triad about its y-axis so that its x-axis makes an angle r with its previous position and its z-axis makes the same angle r with its previous position Z'. This new orientation of the sensor triad is depicted in Fig. 4.14. In this orientation, all the accelerometers will be measuring some part of the gravity vector

$$\begin{aligned} f_x &= -g\cos(p)\sin(r) \\ f_y &= g\sin(p) \\ f_z &= g\cos(p)\cos(r) \end{aligned} \tag{4.22}$$

4.8.2 Theoretical Measurements of a Stationary Gyro Triad

Now consider a gyroscope triad which is stationary on the Earth's surface. Since the triad is stationary, the only rotational motion acting on the sensors will be the Earth's rotation rate ω_e. Now assume that the body frame (in this case is the triad

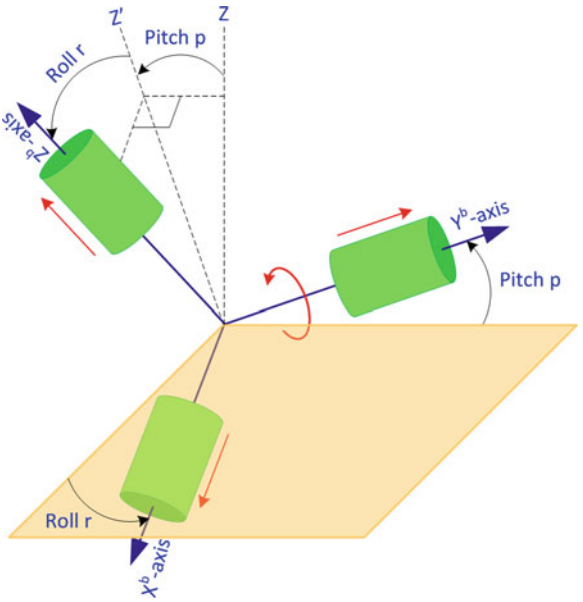


Fig. 4.14 An accelerometer triad with the y-axis making an angle p (*called pitch*) with the level ground and with the y-axis then rotated through an angle r (*called roll*)

itself) coincides with the navigation frame, which is the ENU local-level frame, implying that

$$\begin{aligned} &\text{x} - \text{axis: east} \\ &\text{y} - \text{axis: north} \\ &\text{z} - \text{axis: upward} \end{aligned}$$

as shown in Fig. 4.15.

In this stationary case the gyroscope measurements depend on the latitude of the gyro triad because they are merely measuring the Earth's rotation rate. Their measurements at the equator and at the poles will either be zero or ω_e, depending on the direction of the sensitive axis of the gyroscope, but at any other point they will measure a quantity that lies between these limiting values in accordance with Table 4.2.

For an arbitrary point P on Earth and assuming that the b-frame is aligned with navigation frame, the measurements of the gyroscopes depend on latitude. This is depicted in Fig. 4.16.

According to Fig. 4.16, the components of the Earth's rotation measured by the gyro triad are

$$\omega_N = \omega_y = \omega_e \sin(90 - \varphi) = \omega_e \cos\varphi \tag{4.23}$$

$$\omega_U = \omega_z = \omega_e \cos(90 - \varphi) = \omega_e \sin\varphi \tag{4.24}$$

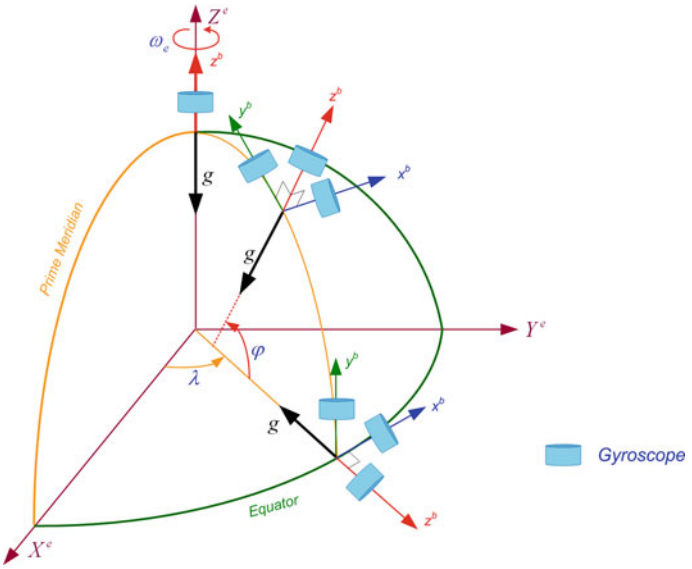


Fig. 4.15 Gyroscope triad at the Earth's surface with its axis aligned with the ENU frame at latitudes 0, 90° and an arbitrary latitude in that range

Table 4.2 The gyroscope measurements at various points on Earth

	Equator	Arbitrary position	North pole
ω_x	0	0	0
ω_y	ω_e	0–ω_e	0
ω_z	0	0–ω_e	ω_e

and it is also evident that $\omega_e = \sqrt{\omega_N^2 + \omega_U^2}$. Also, because the x-gyro is located in a plane perpendicular to the Earth's rotation axis, it will not sense part of the Earth's rotation rate and hence $\omega_E = 0$.

These stationary measurements of the gyro triad represent the Earth's rotation interpreted in the local-level frame, denoted by the angular velocity vector ω_{ie}^l as

$$\boldsymbol{\omega}_{ie}^{\ell} = \begin{bmatrix} \omega_E \\ \omega_N \\ \omega_U \end{bmatrix} = \begin{bmatrix} 0 \\ \omega_e \cos\varphi \\ \omega_e \sin\varphi \end{bmatrix} \tag{4.25}$$

4.8.3 Theoretical Measurements of a Moving Gyro Triad

On a moving platform using the local-level frame as a navigation frame, the gyro triad will monitor two rotational components: the stationary part discussed in the previous section and the non-stationary part caused by the change of orientation of the local-level frame.

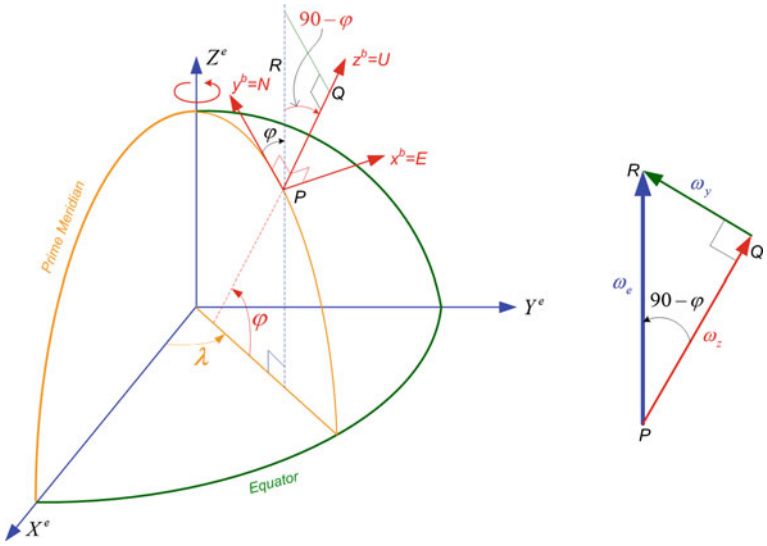


Fig. 4.16 The geometry of the Earth's surface with a gyroscope triad at an arbitrary point P (*in the ENU frame*)

For simplicity, we will assume a vehicle that is moving with velocities v_e and v_n in the east and north directions respectively, and that the b-frame is aligned with the navigation frame. To generalize the measurement we will assume that the vehicle is moving across the Earth at an altitude h. In this case we will derive an expression for the rate of change of latitude φ and longitudeλ.

Figure 4.17 shows the navigation frame (l-frame) as it moves over the surface of the Earth, viewed from the meridian plane.

The rate of change of the latitude φ is

$$\frac{\Delta L_N}{\Delta t} = \frac{\Delta \varphi}{\Delta t} \times (R_M + h) \tag{4.26}$$

where ΔL_N is the small segment of the arc covered during a small time Δt.

When Δt approaches zero

$$v_n = \dot{\varphi} \times (R_M + h) \tag{4.27}$$

$$\dot{\varphi} = \frac{v_n}{R_M + h} \tag{4.28}$$

Similarly, Fig. 4.18 shows the same l-frame as viewed from the top (z-axis) of the e-frame. Since the vehicle is moving in E-N, there will be a component of velocity along the east direction as well, and we can compute the rate of change of the longitude λ as

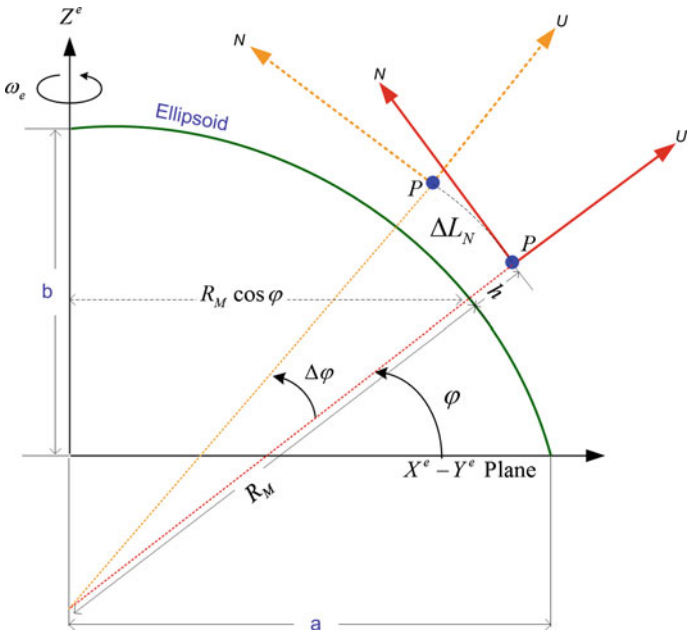


Fig. 4.17 A depiction of the LLF as it moves over the Earth's surface, viewed from the meridian plane

$$\frac{\Delta L_E}{\Delta t} = \frac{\Delta\lambda}{\Delta t} \times (R_N + h)\cos\varphi \tag{4.29}$$

When Δt approaches zero

$$v_e = \dot{\lambda} \times (R_N + h)\cos\varphi \tag{4.30}$$

$$\dot{\lambda} = \frac{v_e}{(R_N + h)\cos\varphi} \tag{4.31}$$

and

$$\dot{h} = v_u \tag{4.32}$$

The gyroscopes of a moving triad will measure the stationary component due to the Earth's rotation as well the non-stationary component caused by the rate of change of latitude and longitude as the vehicle travels. According to the geometry of Fig. 4.16, the angular velocity of the local-level frame as measured by the gyroscope along the x-axis (the E direction) is

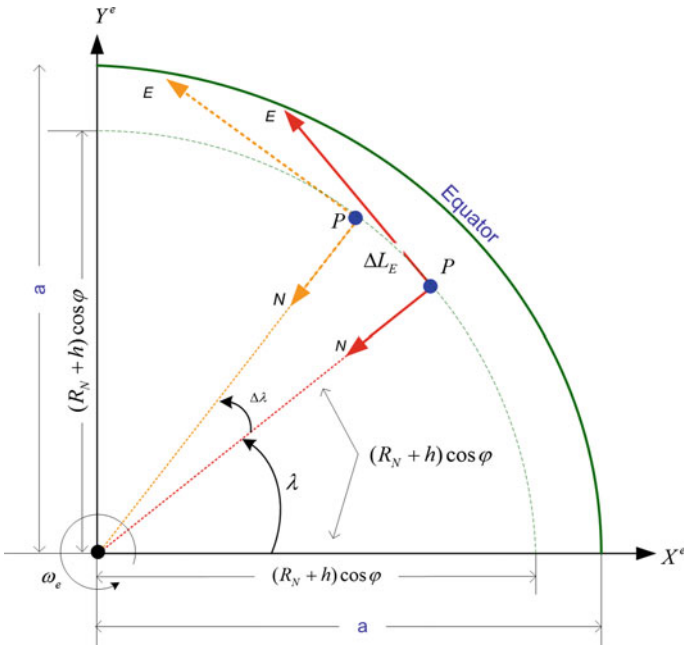


Fig. 4.18 An illustration of the LLF over the Earth's surface, viewed from the z-axis

$$\omega_E = -\dot{\varphi} = \underbrace{-\frac{v_n}{R_M + h}}_{\textit{Non Stationary Component}} \tag{4.33}$$

and that for the y-axis (the N direction) is

$$\omega_N = \dot{\lambda}\cos\varphi + \omega^e \cos\varphi \tag{4.34}$$

Substituting $\dot{\lambda}$ from Eq. (4.31) gives

$$\begin{aligned} \omega_N &= \frac{v_e}{(R_N + h)\cos\varphi}(\cos\varphi) + \omega^e \cos\varphi \\ \omega_N &= \underbrace{\frac{v_e}{R_N + h}}_{\textit{Non Stationary Component}} + \underbrace{\omega^e \cos\varphi}_{\textit{Stationary Component}} \end{aligned} \tag{4.35}$$

And the angular velocity measured by the gyroscope along the z-axis (the U direction) is

$$\begin{aligned}\omega_U &= \dot{\lambda}\sin\varphi + \omega^e \sin\varphi \\ \omega_U &= \frac{v_e}{(R_N + h)\cos\varphi}\sin\varphi + \omega^e \sin\varphi \\ \omega_U &= \underbrace{\frac{v_e}{R_N + h}\tan\varphi}_{\text{Non Stationary Component}} + \underbrace{\omega^e \sin\varphi}_{\text{Stationary Component}}\end{aligned} \tag{4.36}$$

The angular velocity of the local-level frame with respect to the e-frame as expressed in the local level frame ω^l_{el} consists of the non-stationary components

$$\boldsymbol{\omega}^{\ell}_{e\ell} = \begin{bmatrix}\omega_E \\ \omega_N \\ \omega_U\end{bmatrix}_{\text{Non}-\text{stationarycomponent}} = \begin{bmatrix}-\dot{\varphi} \\ \dot{\lambda}\cos\varphi \\ \dot{\lambda}\sin\varphi\end{bmatrix} = \begin{bmatrix}-\dfrac{v_n}{R_M + h} \\ \dfrac{v_e}{R_N + h} \\ \dfrac{v_e \tan\varphi}{R_N + h}\end{bmatrix} \tag{4.37}$$

4.9 Notes on Inertial Sensor Measurements

In the examples above, it was assumed that the INS body frame was aligned with the navigation frame. But for strapdown systems the b-frame can take essentially any arbitrary direction because the accelerometers and gyros are strapped onto the vehicle, which can adopt any orientation with respect to the navigation frame. The establishment of the relationship between the INS body frame and the local level (navigation) frame is usually done at the beginning of the survey by a stationary alignment process. If continuous external velocity information is available (e.g. from GPS) this can be done in kinematic mode. In this process, the initial attitude angles (pitch, roll and azimuth) between the b-frame and the n-frame require to be estimated. The attitude angles are used in generating the rotation matrix R^n_b for the transformation from the b-frame to the n-frame. The rotation rates measured by the gyros are used to constantly update this matrix. Once this transformation has been made, the process of integrating an acceleration measurement twice will provide the IMU's position difference relative to the initial point.

However, as noted earlier, accelerometers cannot separate the total platform acceleration from that caused by the presence of gravity. In fact, accelerometers provide the sum of the platform's acceleration in space and its acceleration due to gravity. The accelerometer measurements must be combined with knowledge of the ambient gravitational field in order to determine the acceleration of the vehicle with respect to a non-inertial reference frame. Obviously, the inertial navigation is fundamentally dependent on an accurate specification of the position, velocity and attitude of the moving platform prior to the start of navigation.

4.10 Inertial Sensor Performance Characteristics

To assess an inertial sensor for a particular application, numerous characteristics must be considered. But first we will introduce some general terms.

a. **Repeatability**: The ability of a sensor to provide the same output for repeated applications of the same input, presuming all other factors in the environment remain constant. It refers to the maximum variation between repeated measurements in the same conditions over multiple runs.
b. **Stability**: This is the ability of a sensor to provide the same output when measuring a constant input over a period of time. It is defined for single run.
c. **Drift**: The term drift is often used to describe the change that occurs in a sensor measurement when there is no change in the input. It is also used to describe the change that occurs when there is zero input.

The performance characteristics of inertial sensors (either accelerometers or gyroscopes) are usually described in terms of the following principal parameters: sensor bias, sensor scale factor, noise and bandwidth. These parameters (among others) will be discussed in the next section, which deals with the errors of inertial sensors.

4.11 Inertial Sensor Errors

Inertial sensors are prone to various errors which get more complex as the price of the sensor goes down. The errors limit the accuracy to which the observables can be measured. They are classified according to two broad categories of systematic and stochastic (or random) errors.

4.11.1 Systematic Errors

These types of errors can be compensated by laboratory calibration, especially for high-end sensors. Some common systematic sensor errors (Grewal et al. 2007) are described below.

4.11.1.1 Systematic Bias Offset

This is a bias offset exhibited by all accelerometers and gyros. It is defined as the output of the sensor when there is zero input, and is depicted in Fig. 4.19. It is independent of the underlying specific force and angular rate.

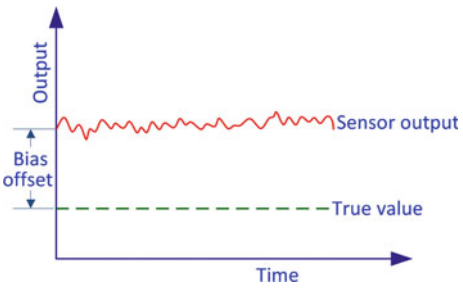


Fig. 4.19 Inertial sensor bias error

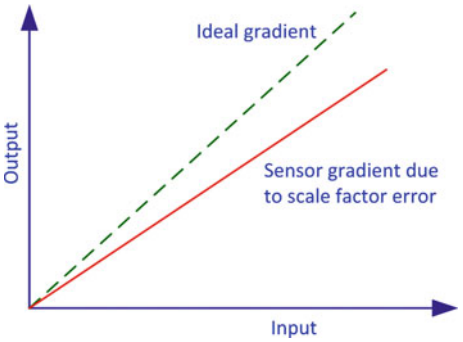


Fig. 4.20 Inertial sensor scale factor error

4.11.1.2 Scale Factor Error

This is the deviation of the input–output gradient from unity. The accelerometer output error due to scale factor error is proportional to the true specific force along the sensitive axis, whereas the gyroscope output error due to scale factor error is proportional to the true angular rate about the sensitive axis. Figure 4.20 illustrates the effect of the scale factor error.

4.11.1.3 Non-linearity

This is non-linearity between the input and the output, as shown in Fig. 4.21.

4.11.1.4 Scale Factor Sign Asymmetry

This is due to the different scale factors for positive and negative inputs, as shown in Fig. 4.22.

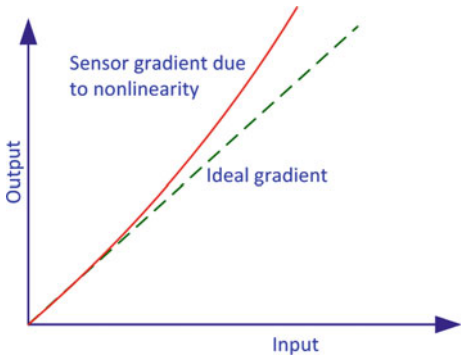


Fig. 4.21 Non-linearity of inertial sensor output

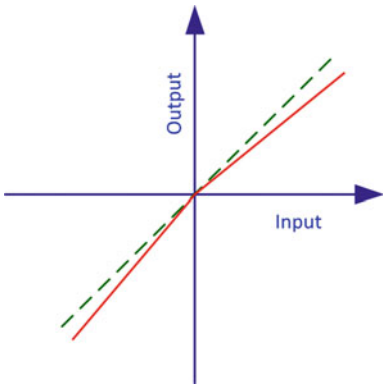


Fig. 4.22 Scale factor sign asymmetry

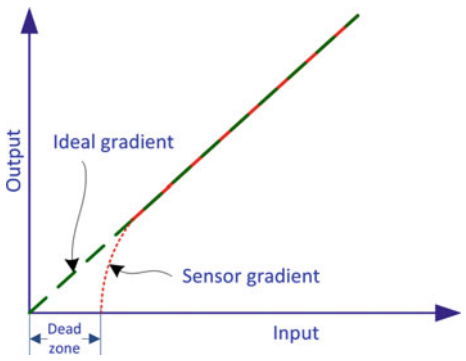


Fig. 4.23 Dead zone in the output of an inertial sensor

4.11.1.5 Dead Zone

This is the range where there is no output despite the presence of an input, and it is shown in Fig. 4.23.

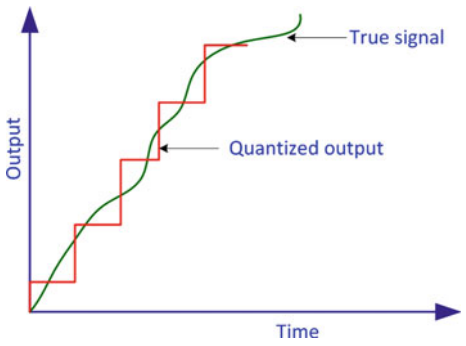


Fig. 4.24 The error due to quantization of an analog signal to a digital signal

4.11.1.6 Quantization Error

This type of error is present in all digital systems which generate their inputs from analog signals, and is illustrated in Fig. 4.24.

4.11.1.7 Non-orthogonality Error

As the name suggests, non-orthogonality errors occur when any of the axes of the sensor triad depart from mutual orthogonality. This usually happens at the time of manufacturing. Figure 4.25 depicts the case of the z-axis being misaligned by an angular offset of θ_{zx} from xz-plane and $\theta_{zy}$from the yz-plane.

4.11.1.8 Misalignment Error

This is the result of misaligning the sensitive axes of the inertial sensors relative to the orthogonal axes of the body frame as a result of mounting imperfections. This is depicted in Fig. 4.26 for a sensor frame misalignment (using superscript 's') with respect to the body in a 2D system in which the axes are offset by the small angle $\delta\theta$.

4.11.2 Random Errors

Inertial sensors suffer from a variety of random errors which are usually modeled stochastically in order to mitigate their effects.

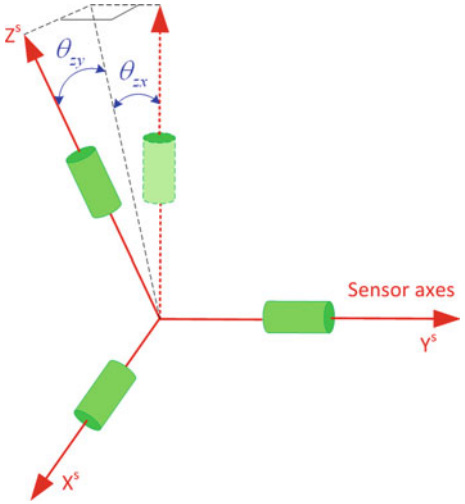


Fig. 4.25 Sensor axes non-orthogonality error

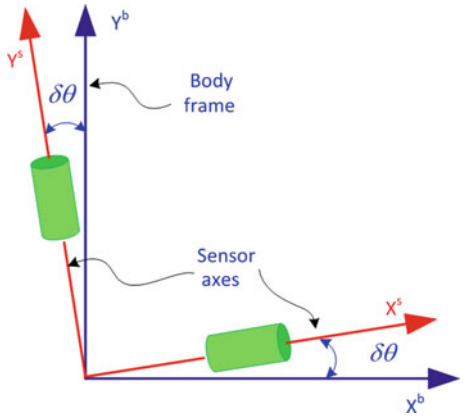


Fig. 4.26 Misalignment error between the body frame and the sensor axes

4.11.2.1 Run-to-Run Bias Offset

If the bias offset changes for every run, this falls under the bias repeatability error, and is called the run-to-run bias offset.

4.11.2.2 Bias Drift

This is a random change in bias over time during a run. It is the instability in the sensor bias for a single run, and is called bias drift. It is illustrated in Fig. 4.27. Bias is deterministic but bias drift is stochastic. One cause of bias drift is a change in temperature.

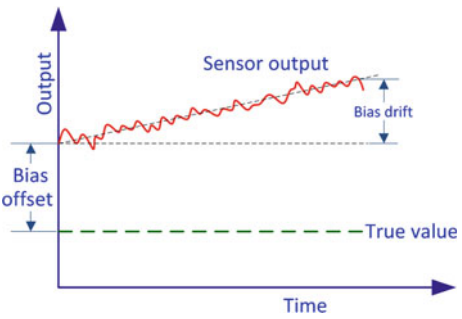


Fig. 4.27 Error in sensor output due to bias drift

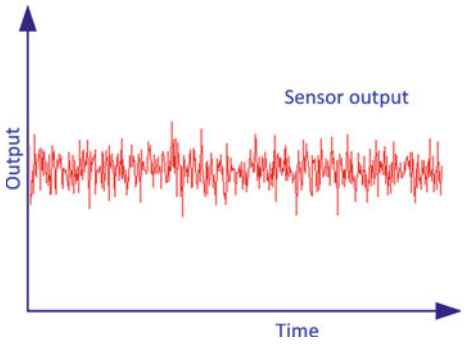


Fig. 4.28 A depiction of white noise error

4.11.2.3 Scale Factor Instability

Random changes in scale factor during a single run. This is usually the result of temperature variations. The scale factor can also change from run to run, but stay constant during a particular run. This demonstrates the repeatability of the sensor and is also called the run-to-run scale factor.

4.11.2.4 White Noise

This is an uncorrelated noise that is evenly distributed in all frequencies. This type of noise can be caused by power sources but can also be intrinsic to semiconductor devices. White noise is illustrated in Fig. 4.28.

4.11.3 Notes on Random Errors

Most manufacturers express the randomness associated with their inertial sensors by the concept of random walk. The angle random walk (ARW) for gyroscopes is usually specified in terms of deg/hr/$\sqrt{\text{Hz}}$ or deg/$\sqrt{\text{hr}}$, and the velocity random

Table 4.3 Performance specification of various KVH gyroscopes

	KVH DSP-300 (single axis FOG)	KVH DSP-3100 (single axis FOG)	DSP-3400 single axis FOG
Bandwidth	100 Hz	1000 Hz	1000 Hz
Bias drift	$<3^\circ/\text{h}$	$<1^\circ/\text{h}$	$<1^\circ/\text{h}$
ARW	$<6^\circ/\text{h}/\sqrt{\text{Hz}}$ $\left(0.1^\circ/\sqrt{\text{h}}\right)$	$<4^\circ/\text{h}/\sqrt{\text{Hz}}$ $\left(0.0667^\circ/\sqrt{\text{h}}\right)$	$<4^\circ/\text{h}/\sqrt{\text{Hz}}$ $\left(0.0667^\circ/\sqrt{\text{h}}\right)$
Scale factor	$<0.05\,\%$	$<0.05\,\%$	$<0.05\,\%$

walk (VRW) for accelerometers is given in terms of $\mu\text{g}/\sqrt{\text{Hz}}$ or $\text{m/s}/\sqrt{\text{hr}}$. This definition requires knowledge of the data rate (sampling frequency) at which the sensor measurements are acquired by the data acquisition systems. The data rate is related to the bandwidth of the sensor, which is another important parameter. The inertial sensor bandwidth (specified in Hz) defines the range of frequencies that can be monitored by the sensor. For example a gyroscope with 100 Hz bandwidth is capable of monitoring the dynamics of frequencies less than 100 Hz. Any higher frequencies will not be detected. For this, the sensor has to sample the signal with at least double the maximum frequency; in this case 200 Hz. But whilst increasing the data rate will broaden the bandwidth to facilitate monitoring higher frequency dynamics, the measurements will also be noisier. Table 4.3 shows some important performance specifications for various KVH gyroscopes (KVH 2012).

4.11.4 Mathematical Models of Inertial Sensor Errors

The performance of an INS can be described in terms of its two major groups of sensors, namely gyroscopes and accelerometers.

4.11.4.1 Gyroscope Measurement Model

Gyroscopes are angular rate sensors that provide either angular rate or attitude depending on whether they are of the rate sensing or rate integrating type.

Measurements of angular rate can be modeled by the observation equation

$$\tilde{\boldsymbol{\omega}}_{ib}^{b} = \boldsymbol{\omega}_{ib}^{b} + \mathbf{b}_g + S\boldsymbol{\omega}_{ib}^{b} + N\boldsymbol{\omega}_{ib}^{b} + \boldsymbol{\varepsilon}_g \tag{4.38}$$

where

$\tilde{\boldsymbol{\omega}}_{ib}^{b}$ is the gyroscope measurement vector (deg/h)
$\boldsymbol{\omega}_{ib}^{b}$ is the true angular rate velocity vector (deg/h)

$\mathbf{b}_g$ is the gyroscope instrument bias vector (deg/h)
S_g is a matrix representing the gyro scale factor
N_g is a matrix representing non-orthogonality of the gyro triad
$\boldsymbol{\varepsilon}_g$ is a vector representing the gyro sensor noise (deg/h).

The matrices N_g and $S_g$are given as

$$N_g = \begin{bmatrix} 1 & \theta_{g,xy} & \theta_{g,xz} \\ \theta_{g,yx} & 1 & \theta_{g,yz} \\ \theta_{g,zx} & \theta_{g,zy} & 1 \end{bmatrix}$$

$$S_g = \begin{bmatrix} s_{g,x} & 0 & 0 \\ 0 & s_{g,y} & 0 \\ 0 & 0 & s_{g,z} \end{bmatrix}$$

where $\theta_{(.),(.)}$ are the small angles defining the misalignments between the different gyro axes and $s_{(.),(.)}$are the scale factors for the three gyros.

4.11.4.2 Accelerometer Measurement Model

Performance factors describing accelerometer accuracy are similar to those which characterize the gyro accuracy bias uncertainty, scale factor stability, and random noise. Measurements of the specific force can be modeled by the observation equation

$$\tilde{\mathbf{f}}^b = \mathbf{f^b} + \mathbf{b_a} + \mathbf{S_1 f} + \mathbf{S_2 f^2} + \mathbf{N_a f} + \delta\mathbf{g} + \boldsymbol{\varepsilon}_\mathbf{a} \tag{4.39}$$

where

$\tilde{\mathbf{f}}^b$ is the accelerometer measurement vector (m/s^2)
$\mathbf{f}^b$ is the true specific force vector (i.e. observable) (m/s^2)
$\mathbf{b}_a$ is the accelerometer instrument bias vector (m/s^2)
S_1 is a matrix of the linear scale factor error
S_2 is a matrix of the non-linear scale factor error
N_a is a matrix representing non-orthogonality of the accelerometer triad
$\delta\mathbf{g}$ is the anomalous gravity vector (i.e. the deviation from the theoretical gravity value) (m/s^2)
$\boldsymbol{\varepsilon}_a$ is a vector representing the accelerometer sensor noise (m/s^2).

The matrices N_a $S_1$and S_2 are

$$N_a = \begin{bmatrix} 1 & \theta_{a,xy} & \theta_{a,xz} \\ \theta_{a,yx} & 1 & \theta_{a,yz} \\ \theta_{a,zx} & \theta_{a,zy} & 1 \end{bmatrix}$$

Table 4.4 Classification of inertial measurement units

Performance	Strategic grade	Navigation grade	Tactical grade	Commercial grade[a]
Positional error	30 m/ h < 100 m/h	1 nmi[b]/h or .5 m/s	10–20 nmi/h	Large variation
Gyroscope drift	0.0001–0.001	<0.01 °/h	1–10°/h	0.1°/s
Gyroscope random walk	–	$<0.002^\circ/\sqrt{h}$	0.05– $<0.02^\circ/\sqrt{h}$	Several $^\circ/\sqrt{h}$
Accelerometer bias	0.1–1	<100 μg	1–5 mg	100–1,000 μg
Applications	Submarines Intercontinental ballistic missile	General navigation high precision georeferencing mapping	Integrated with GPS for mapping Weapons (short time))	Research Low cost navigation pedometers, Antilock breaking active suspension, airbags

[a] Also called automotive grade
[b] 1 nautical mile (nmi) ≈6,076 ft ≈1,851 m

$$S_1 = \begin{bmatrix} s_{1,x} & 0 & 0 \\ 0 & s_{1,y} & 0 \\ 0 & 0 & s_{1,z} \end{bmatrix}$$

$$S_2 = \begin{bmatrix} s_{2,x} & 0 & 0 \\ 0 & s_{2,y} & 0 \\ 0 & 0 & s_{2,z} \end{bmatrix}$$

where $\theta_{(.),(.)}$ are the small angles defining the misalignments between the different accelerometer axes and $s_{(.),(.)}$are the scale factors for the three accelerometers.

For both the inertial sensors, the scale factors and biases are usually considered to be constant (over a certain time) but unknown quantities which are uncorrelated between the different sensors. In principle these errors can be eliminated by the calibration techniques described in Sect. 4.13. The sensor noise ε consists of white noise, correlated noise and random walk, etc. In principle these errors can be minimized by the estimation techniques described in Chap. 7.

4.12 Classification of Inertial Sensors

No universally agreed definition exists for categorizing inertial sensors. However, a rough comparison of different inertial navigation sensors/systems is outlined in Table 4.4 with data obtained from (Groves Dec 2007), (Petovello et al. Jun 2007), (Barbour and Schmidt 2001) and (Wang and Williams 2008).

Usually the cost of an IMU is dictated by the type of inertial sensors. IMUs are categorized according to their intended applications, which mainly depend on the gyroscope bias expressed in deg/hour. A secondary measure of performance is the gyroscope random walk, which is usually expressed in terms of deg/root-hour and accelerometer bias.

4.12.1 Gyroscope Technologies and their Applications

There are several gyroscope technologies, including ring laser gyroscopes (RLG), dynamically tuned gyroscopes (DTG), hemispherical resonant gyroscopes (HRG), and interferometric fiber-optic gyroscopes (IFOG). Spinning mass and ring laser gyroscopes offer high performance but at high cost, and find their application in strategic/tactical and submarine navigation systems. DTG offer a medium level of performance and share some applications with RLG (Prasad and Ruggieri 2005). IFOG and Coriolis-based gyroscopes are of lower performance but cost less and are typically used in torpedoes, tactical missile midcourse guidance, flight control and smart munitions guidance and robotics. There has been a recent trend towards MEMS gyroscopes that are being researched for low cost navigation applications such as car navigation and portable navigation devices. Details of all these sensor technologies can be found in (Barbour and Schmidt 2001; Lawrence 1998).

4.12.2 Accelerometer Technologies and their Applications

The main accelerometer technologies are mechanical pendulous force-rebalance accelerometers, vibrating beam accelerometers (VBAs) and gravimeters. The best performance is provided by mechanically floated instruments, and these are used in strategic missiles. Mechanical pendulous rebalance accelerometers are used in submarine navigation, land and aircraft navigation and space applications. Quartz resonator accelerometers are low grade sensors typically found in tactical missile midcourse guidance.

4.13 Calibration of Inertial Sensors

Calibration is defined as the process of comparing instrument outputs with known reference information to determine coefficients that force the output to agree with the reference information across the desired range of output values. Calibration is used to compute deterministic errors of sensors in the laboratory. The calibration parameters to be determined can change according to the specific technology in an IMU. To accurately determine all of the parameters, special calibration devices are

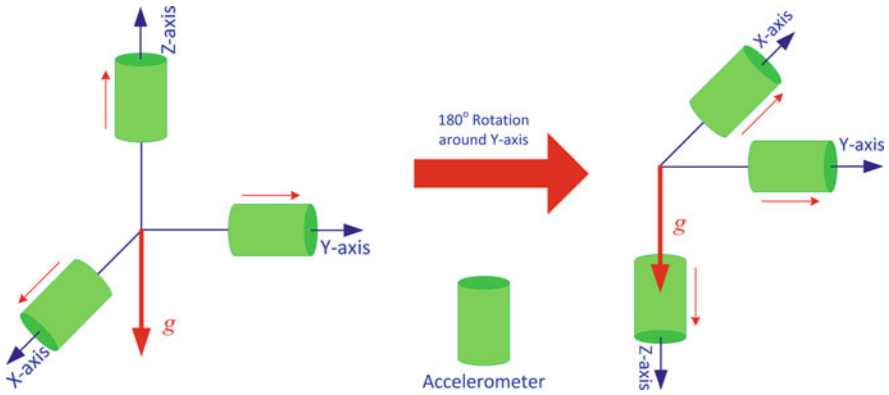


Fig. 4.29 Calibrating an accelerometer, with the sensitive axis facing upward on the *left* and downward on the *right*

needed, such as three-axial turntables, to perform either a six-position static test or an angle rate test.

4.13.1 Six-Position Static Test

In this procedure for sensor calibration, the inertial system is mounted on a level table with each sensitive axis pointing alternately up and down (six positions for three axes). Using these sensor readings it is possible to extract estimates of the accelerometer bias and scale factor by summing and differencing combinations of the inertial system measurements.

4.13.1.1 Accelerometer Calibration

Accelerometers are normally calibrated by sensing gravity. Each accelerometer of the triad is placed on a calibrated rate table with its sensitive axis facing up. After taking about 10–15 min of data, the mean f_{up} is computed. Then a similar reading is taken for f_{down} with sensitive axis pointed downwards. To reiterate the point made earlier, the accelerometers will measure the reaction to gravity.

By Fig. 4.29, the measurements with the sensitive axis of the accelerometer up and down can be expressed as

$$f_{up} = b_a + (1 + S_a)g \tag{4.40}$$

$$f_{down} = b_a - (1 + S_a)g \tag{4.41}$$

The bias b_a is computed by adding these two reading

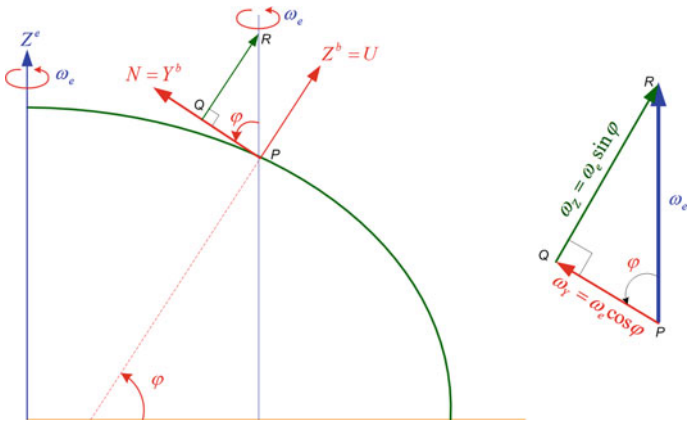


Fig. 4.30 Geometry explaining the component of the Earth's rotation measured by gyros aligned with the north and up directions

$$b_a = \frac{f_{up} + f_{down}}{2} \tag{4.42}$$

and the scale factor S_a is obtained by subtracting the two reading

$$S_a = \frac{f_{up} - f_{down} - 2g}{2g} \tag{4.43}$$

where g represents gravity.

This procedure is repeated for each of the three accelerometers to obtain their individual bias and scale factors.

4.13.1.2 Gyroscope Calibration

A similar procedure is employed for gyroscopes, but this time the Earth's rotation rate is measured instead of gravity. According to Fig. 4.30, for a body frame located at a latitude φ the theoretical projections of the Earth's rotation rate on the body axes are

$$\omega_x = 0;\ \omega_y = \omega_e \cos\varphi;\ \omega_z = \omega_e \sin\varphi$$

Therefore, a vertical gyroscope (with its sensitive axis pointing up) will sense a component of gravity that is $\omega_e \sin\varphi$, and this is used in the calibration.

In accordance with Fig. 4.31, the measurements of the gyroscope with its sensitive axis up and down are

$$\omega_{up} = b_{go} + (1 + S_g)\omega_e \sin\varphi \tag{4.44}$$

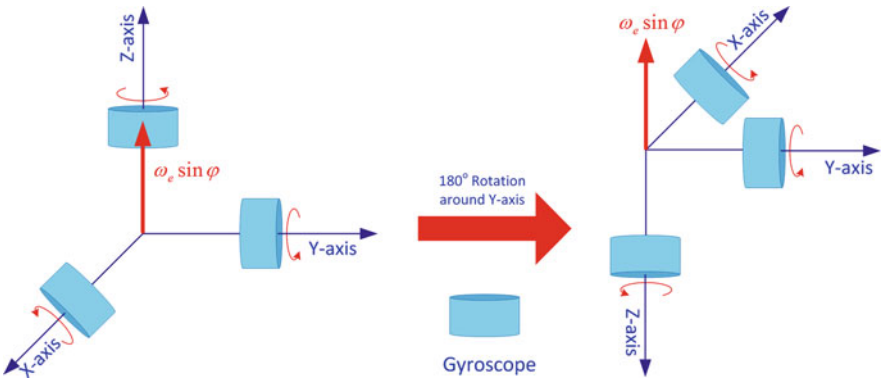


Fig. 4.31 Calibrating a gyroscope, with the sensitive axis facing upward on the *left* and downward on the *right*

$$\omega_{down} = b_{go} - (1 + S_g)\omega_e \sin\varphi \tag{4.45}$$

where $\omega_e \sin\varphi$ is the vertical component of the Earth's rotation rate, ω_e is the Earth's rotation rate about its spin axis, and φ is the latitude of the location of the gyroscope.

The bias and scale factor are obtained in a similar way as for accelerometers

$$b_g = \frac{\omega_{up} + \omega_{down}}{2} \tag{4.46}$$

$$S_g = \frac{\omega_{up} - \omega_{down} - 2\omega_e \sin\varphi}{2\omega_e \sin\varphi} \tag{4.47}$$

where b_{gyro} is the bias, S_{gyro} is the scale factor, ω_e is the Earth's rotation rate and φ is the latitude of the location of the gyroscope.

This procedure is repeated for each of the three gyroscopes to obtain their individual bias and scale factors.

For low cost gyroscopes that cannot detect the Earth's rotation rate, the table can be rotated at a constant rate of $\omega_t = 60^\circ/\text{s}$ (or any indeed other rate that is above the detection threshold of the gyroscopes) and $\omega_e \sin\varphi$ is replaced with the value of ω_t in Eqs. (4.44), (4.45) and (4.47).

4.13.2 Angle Rate Tests

Angle rate tests are utilized to calibrate gyroscope biases, scale factors and non-orthogonalities. In this type of calibration the IMU is mounted on a precision rate table (Fig. 4.32 shows one) which is rotated through a set of very accurately defined angles. By comparing these known rotations with the estimates of the

Fig. 4.32 A precision rate table with a gyroscope mounted on top

angles obtained by integrating the rate outputs provided by the gyros, it is possible to estimate the various errors in the gyros measurements. For instance, if the table is rotated clockwise and counterclockwise through the same angle then the biases and scale factors errors of the gyros can be estimated.

4.14 Importance of Calibration of Inertial Sensors

Calibration of the inertial sensors plays an important role in the ultimate accuracy of an INS. Any residual flaws in the sensors cause errors which, as we shall see, tend to grow with time.

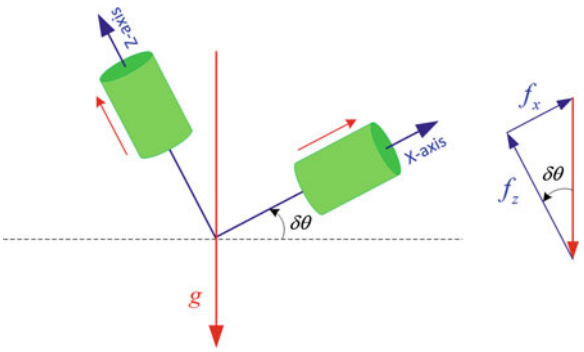


Fig. 4.33 A stationary accelerometer placed on a *tilted plane*

Consider Fig. 4.33, which shows a stationary accelerometer on a plane tilted from horizontal by the small angle $\delta\theta$. We can prove that this tilt is observable because the accelerometers are sitting in a gravitational field.

The accelerometers will measure a component of g as follows

$$f_z = g\cos\delta\theta \text{ in the z direction} \tag{4.48}$$

$$f_x = g\sin\delta\theta \text{ in the x direction} \tag{4.49}$$

where g is the magnitude of the gravitational field.

For small angles where cos $\delta\theta$ is near unity and sin $\delta\theta$ is $\delta\theta$, the measured values can be approximated as

$$f_z = g \tag{4.50}$$

$$f_x = g\delta\theta \tag{4.51}$$

Therefore the output of the x accelerometer gives us a direct measurement of the tilt about the y axis. Similarly, the output of the y accelerometer provides a measure of the tilt about the x axis. This discussion provides a simple example on correlation of errors. If the x accelerometer has a bias error of b_x, the output of the accelerometer will be

$$f_x = b_x + g\delta\theta \tag{4.52}$$

If we are trying to level the platform (which involves determining its attitude because we do not actually rotate the sensors in a strapdown system) employing the accelerometer output, we cannot tell the difference between the contribution of accelerometer bias or the tilt. Hence, the value of the sensor bias determines the accuracy to which we can estimate the initial tilt alignment of the INS (as will be discussed in greater detail later). This can be clearly explained in the case of small tilt angles where f_x will be approximately zero and

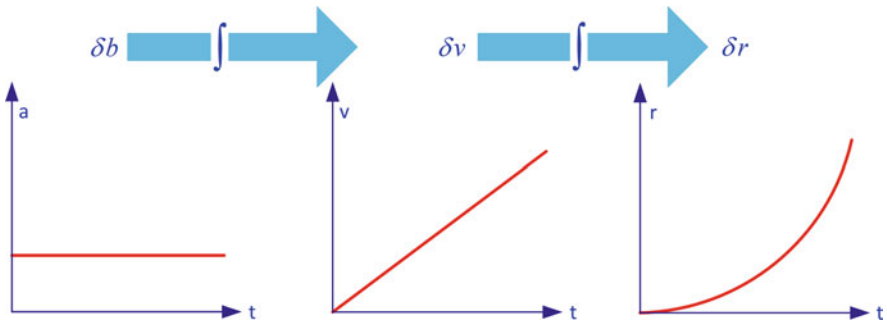


Fig. 4.34 Effect of a bias error in acceleration, velocity and position

$$\delta\theta \approx -\frac{b_x}{g} \tag{4.53}$$

4.14.1 Case-I: Bias Error in an Accelerometer

An uncompensated accelerometer bias error b_f (expressed in terms of either g or m/s^2) will introduce an error proportional to t in the velocity measurement and an error proportional to t^2 in the position measurement due to the integrations to obtain the velocity v and the position r from the gyroscope output. Taking these into account

$$v = \int b_f dt \tag{4.54}$$

$$r = \int v dt = \int \int b_f t dt = \frac{1}{2} b_f t^2 \tag{4.55}$$

Therefore an accelerometer bias introduces a first-order error in velocity and a second-order error in the position, as shown in Fig. 4.34.

4.14.2 Case-II: Bias Error in the Gyroscope

An uncompensated gyro bias (expressed in terms of deg/h or radians/s) in the gyro error b_ω will introduce an angle error (in roll or pitch) that is proportional to time t, and hence

$$\delta\theta = \int b_\omega dt = b_\omega t \tag{4.56}$$

This misalignment of the INS will result in the measured acceleration vector being incorrectly projected. This in turn will introduce acceleration in one of the horizontal channels (as mentioned previously) with a value of

$$a = g\sin(\delta\theta) \approx g\delta\theta \approx gb_\omega t \tag{4.57}$$

Now when we integrate this acceleration to obtain the velocity and position, it will introduce an error proportional to t^2 in the velocity of

$$v = \int b_\omega g t dt = \frac{1}{2} b_\omega g t^2 \tag{4.58}$$

and an error proportional to t^3 in the position of

$$r = \int v dt = \int \int \frac{1}{2} b_\omega g t^2 dt = \frac{1}{6} b_\omega g t^3 \tag{4.59}$$

Since, a gyro bias introduces a second-order error in velocity and a third-order error in position, the gyroscope tends to dictate the quality of an IMU and thus the accuracy of the output of navigation algorithms.

4.15 Initialization and Alignment of Inertial Sensors

An INS takes acceleration and rotation rates from sensors to calculates velocity and attitudes by integrating them once, and then integrates the velocity in order to obtain the position. The navigation equations require starting values for position, velocity and attitude. These are readily available from the last epoch of an ongoing iteration, but for the first epoch the INS must be specifically provided with this information before it can begin to function. This process is called initialization for position and velocity, and is called alignment for attitude (Groves Dec 2007).

4.15.1 Position and Velocity Initialization

Position can be initialized using a vehicle's last known position before it started to move. For a system where the INS is integrated with other systems, typically GPS, a position can easily be provided by the external navigation system. In some cases the starting point is known a priori (for example a pre-surveyed location). If the vehicle is stationary then the velocity initialization can be made with zero input. If it is moving, the initial velocity can be provided by an external navigation source such as GPS, Doppler or an odometer.

4.15.2 Attitude Alignment

Attitude alignment involves two steps. First, the platform is leveled by initializing the pitch (p) and roll (r) angles, and then gyro-compassing to provide an initial value of the heading (alternatively known as the yaw angle 'y' or azimuth 'A').

4.15.2.1 Accelerometer Leveling

With the vehicle held stationary, accelerometers measure the components of the reaction to gravity due to the pitch and roll angles (i.e. the tilt with respect to the horizontal plane). The accelerometer measurements are in the body frame and can be expressed as

$$\begin{aligned}\mathbf{f}^b &= R_l^b\left(-\mathbf{g}^l\right)\\ &= \left(R_b^l\right)^T\left(-\mathbf{g}^l\right)\end{aligned} \tag{4.60}$$

where the rotation matrix R_b^l is defined as

$$R_b^l = \begin{bmatrix} \cos y\cos r - \sin y\sin p\sin r & -\sin y\cos p & \cos y\sin r + \sin y\sin p\cos r \\ \sin y\cos r + \cos y\sin p\sin r & \cos y\cos p & \sin y\sin r - \cos y\sin p\cos r \\ -\cos p\sin r & \sin p & \cos p\cos r \end{bmatrix} \tag{4.61}$$

and the gravity vector $\mathbf{g}^l$ is defined as

$$\mathbf{g}^l = \begin{bmatrix} 0 & 0 & -g \end{bmatrix}^T$$

Substituting these values into Eq. (4.60) gives

$$\mathbf{f}^b = \left(\begin{bmatrix} \cos(y)\cos(r) - \sin(y)\sin(p)\sin(r) & -\sin(y)\cos(p) & \cos(y)\sin(r) + \sin(y)\sin(p)\cos(r) \\ \sin(y)\cos(r) + \cos(y)\sin(p)\sin(r) & \cos(y)\cos(p) & \sin(y)\sin(r) - \cos(y)\sin(p)\cos(r) \\ -\cos(p)\sin(r) & \sin(p) & \cos(p)\cos(r) \end{bmatrix}\right)^T \left(-\begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix}\right) \tag{4.62}$$

$$\mathbf{f}^b = \begin{bmatrix} \cos(y)\cos(r) - \sin(y)\sin(p)\sin(r) & \sin(y)\cos(r) + \cos(y)\sin(p)\sin(r) & -\cos(p)\sin(r) \\ -\sin(y)\cos(p) & \cos(y)\cos(p) & \sin(p) \\ \cos(y)\sin(r) + \sin(y)\sin(p)\cos(r) & \sin(y)\sin(r) - \cos(y)\sin(p)\cos(r) & \cos(p)\cos(r) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \tag{4.63}$$

$$\begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} = \begin{bmatrix} -g\cos(p)\sin(r) \\ g\sin(p) \\ g\cos(p)\cos(r) \end{bmatrix} \tag{4.64}$$

From Eq. (4.64) we can calculate the pitch and roll angles as

$$p = \tan^{-1}\left(\frac{f_y}{\sqrt{f_x^2 + f_z^2}}\right) \tag{4.65}$$

$$r = \tan^{-1}\left(\frac{-f_x}{f_z}\right) \tag{4.66}$$

4.15.2.2 Gyrocompassing

As well as potentially being sensitive to the Earth's rotation rate, a gyroscope will measure the rotation of the body frame with respect to the e-frame, hence

$$\boldsymbol{\omega}_{ib}^b = \boldsymbol{\omega}_{ie}^b + \boldsymbol{\omega}_{eb}^b \tag{4.67}$$

Because the sensor triad is required to be stationary for the calibration process, the second term on the right-hand side of this equations is zero, and

$$\begin{aligned} \boldsymbol{\omega}_{ib}^b &= \boldsymbol{\omega}_{ie}^b + 0 \\ \boldsymbol{\omega}_{ib}^b &= \boldsymbol{R}_e^b \boldsymbol{\omega}_{ie}^e \\ \boldsymbol{\omega}_{ib}^b &= \boldsymbol{R}_l^b \boldsymbol{R}_e^l \boldsymbol{\omega}_{ie}^e \\ \boldsymbol{\omega}_{ib}^b &= \left(\boldsymbol{R}_b^l\right)^T \left(\boldsymbol{R}_l^e\right)^T \boldsymbol{\omega}_{ie}^e \end{aligned} \tag{4.68}$$

where, the rotation matrix $\boldsymbol{R}_b^l$ is defined in Eq. (4.61) and the rotation matrix $\boldsymbol{R}_l^e$ is given as

$$\boldsymbol{R}_l^e = \begin{bmatrix} -\sin\lambda & -\sin\varphi\cos\lambda & \cos\varphi\cos\lambda \\ \cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi\sin\lambda \\ 0 & \cos\varphi & \sin\varphi \end{bmatrix}$$

and

$$\boldsymbol{\omega}_{ie}^e = \begin{bmatrix} 0 & 0 & \omega_e \end{bmatrix}^T$$

The term ω_e is the Earth's rotation about its spin axis, which is approximately 15.04 deg/h.

By substituting the expressions for $\boldsymbol{R}_b^l, \boldsymbol{R}_l^e$ and $\boldsymbol{\omega}_{ie}^e$ into Eq. (4.68), we get

$$\boldsymbol{\omega}_{ib}^b = \begin{bmatrix} \cos(y)\cos(r) - \sin(y)\sin(p)\sin(r) & \sin(y)\cos(r) + \cos(y)\sin(p)\sin(r) & -\cos(p)\sin(r) \\ -\sin(y)\cos(p) & \cos(y)\cos(p) & \sin(p) \\ \cos(y)\sin(r) + \sin(y)\sin(p)\cos(r) & \sin(y)\sin(r) - \cos(y)\sin(p)\cos(r) & \cos(p)\cos(r) \end{bmatrix} \begin{bmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\sin\varphi\cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi \\ \cos\varphi\cos\lambda & \cos\varphi\sin\lambda & \sin\varphi \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \omega_e \end{bmatrix} \tag{4.69}$$

$$\boldsymbol{\omega}_{ib}^{b} = \begin{bmatrix} \cos(y)\cos(r) - \sin(y)\sin(p)\sin(r) & \sin(y)\cos(r) + \cos(y)\sin(p)\sin(r) & -\cos(p)\sin(r) \\ -\sin(y)\cos(p) & \cos(y)\cos(p) & \sin(p) \\ \cos(y)\sin(r) + \sin(y)\sin(p)\cos(r) & \sin(y)\sin(r) - \cos(y)\sin(p)\cos(r) & \cos(p)\cos(r) \end{bmatrix} \begin{bmatrix} 0 \\ \omega_e \cos\varphi \\ \omega_e \sin\varphi \end{bmatrix} \tag{4.70}$$

$$\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \begin{bmatrix} \{\sin(y)\cos(r) + \cos(y)\sin(p)\sin(r)\}\omega_e \cos\varphi - \omega_e \cos(p)\sin(r)\sin\varphi \\ \omega_e cos\varphi \cos(y)\cos(p) + \omega_e \sin\varphi \sin(p) \\ \{\sin(y)\sin(r) - \cos(y)\sin(p)\cos(r)\}\omega_e \cos\varphi + \omega_e \sin\varphi \cos p \cos(r) \end{bmatrix} \tag{4.71}$$

Since the pitch angle p and roll angle r were already obtained during the accelerometer leveling process using Eqs. (4.65) and (4.66) respectively, we will now use the three gyro measurements to obtain the yaw angle y. It is evident from Eq. (4.71) that if the gyro measurement ω_x is multiplied by $\cos(r)$ and the gyro measurement ω_z is multiplied by $\sin(r)$ then

$$\omega_x(\cos r) + \omega_z(\sin r) = (\omega_e \cos\varphi)\sin y \tag{4.72}$$

Similarly

$$\omega_y(\cos p) + \omega_x(\sin p \sin r) - \omega_z(\cos r \sin p) = (\omega_e \cos\varphi)\cos y \tag{4.73}$$

Consequently

$$\tan y = \frac{\omega_x(\cos r) + \omega_z(\sin r)}{\omega_y(\cos p) + \omega_x(\sin p \sin r) - \omega_z(\cos r \sin p)} \tag{4.74}$$

Finally the yaw angle is

$$y = \tan^{-1}\left[\frac{\omega_x(\cos r) + \omega_z(\sin r)}{\omega_y(\cos p) + \omega_x(\sin p \sin r) - \omega_z(\cos r \sin p)}\right] \tag{4.75}$$

For land vehicle applications, where the vehicle travels is almost a horizontal plane, the pitch and roll angles are close to zero and the small angle approximation yields

$$\begin{aligned} \cos r = \cos p &\approx 1 \\ \sin r = \sin p &\approx 0 \end{aligned} \tag{4.76}$$

Equation (4.75) now reduces to

$$y = \tan^{-1}\left(\frac{\omega_x}{\omega_y}\right) \tag{4.77}$$

It should be noted that gyro-compassing is not feasible for low grade IMUs, which cannot detect the Earth's rotation because their noise threshold exceeds the signal for the Earth's rotation. For these IMUs a heading (yaw angle) is obtained externally using either a compass or a magnetometer. For systems integrated with

GPS, the east and north velocities supplied by this system provide a heading after the host platform has started to move. In this case the heading is

$$y = -\tan^{-1}\left(\frac{v_e}{v_n}\right) \tag{4.78}$$

where v_e is east velocity and $v_n$is north velocity.

References

Barbour N, Schmidt G (2001) Inertial sensor technology trends. Sens J IEEE 1(4):332–339

Grewal MS, Weill LR, Andrews AP (2007) Global positioning systems, inertial navigation, and integration, 2nd edn. Wiley, New york

Groves PD (2007) Principles of GNSS inertial and multi-sensor integrated navigation systems. Artech House, USA

KVH (2012) www.kvh.com/Commercial-and-OEM/Gyros-and-Inertial-Systems-and-Compasses/Gyros-and-IMUs-and-INS.aspx. Accessed 28 Mar 2012

Lawrence A (1998) Modern inertial technology: navigation, guidance, and control, 2nd edn. Springer, New York

Petovello M, Lachapelle G, Scherzinger B (2007) What are the important considerations when selecting the type of quality of IMU for integration with GNSS. InsideGNSS

Prasad R, Ruggieri M (2005) Applied satellite navigation using GPS GALILEO and augmentation systems. Artech House, USA

Wang, HG, Williams TC (2008) Strategic inertial navigation systems-high-accuracy inertially stabilized platforms for hostile environments. IEEE Control system Magazine 28(1):65–85

Chapter 5
Inertial Navigation System Modeling

Modeling requires representing real world phenomena by mathematical language. To keep the problem tractable the goal is not to produce the most comprehensive descriptive model but to produce the simplest possible model which incorporates the major features of the phenomena of interest. The model is also restricted by the ability of mathematics to describe a phenomenon. This book deals with models which describe the motion of an object on or near the surface of the Earth. This kind of motion is greatly influenced by the geometry of the Earth. There are two broad categories for modeling motion: dynamic and kinematic.

5.1 Dynamic Modeling

Dynamic modeling deals with the description of motion due to forces or other laws of interaction to describe the effect of one physical system on another. Newton's second law of motion is an example of this modeling, and expresses the change of motion resulting from the force model $F\{\mathbf{r}, \mathbf{v}, t\}$

$$m\frac{d^2\mathbf{r}}{dt^2} = m\ddot{r} = F\{\mathbf{r}, \mathbf{v}, t\} \tag{5.1}$$

where $\mathbf{r}$ is the position vector, $\mathbf{v}$ is the velocity vector, these being referred to as state vectors, and time t is an independent variable. This equation can be solved for the velocity vector by a single integration or for the position vector by a double integration. This solution represents the motion of a particle with respect to an inertial frame of reference.

A unique analytical solution to the above equation can be given as

$$\mathbf{r}(t) = \mathbf{r}(t, \mathbf{a}, \mathbf{b}) \tag{5.2}$$

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_5,

where **a** and **b** are constant vectors which can usually be determined from initial conditions such as

$$\mathbf{r}(0, \mathbf{a}, \mathbf{b}) = \mathbf{r}(t_0) \tag{5.3}$$

$$\mathbf{v}(0, \mathbf{a}, \mathbf{b}) = \mathbf{v}(t_0) \tag{5.4}$$

As a result, if the position and velocity of a particle are specified for an initial time, they can be uniquely determined for any later time. Applications of dynamic modeling include the rotation of the Earth and describing the motion of a satellite orbiting the Earth. This kind of modeling is usually not applied to the motion of an object on or near the surface of the Earth because the model would become very involved and essentially non-analytical.

5.2 Kinematic Modeling

Kinematic modeling is often used where dynamic modeling of a moving object is impractical. It deals with studying the motion of an object without consideration of its mass or the forces acting upon it, obviating the force model. In kinematic modeling the motion of a body is determined solely based on position, velocity or acceleration measurements, collectively known as *observables*. Depending on the observable available, the following time-variable vectors are used for the position vector

$$\mathbf{r}(t) = \{x(t), y(t), z(t)\} \tag{5.5}$$

the velocity vector

$$\dot{\mathbf{r}}(t) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) = \left(v_x, v_y, v_z\right) \tag{5.6}$$

and the acceleration vector

$$\ddot{\mathbf{r}}(t) = \left(\frac{d^2x}{dt^2}, \frac{d^2y}{dt^2}, \frac{d^2z}{dt^2}\right) = \left(a_x, a_y, a_z\right) = \left(\dot{v}_x, \dot{v}_y, \dot{v}_z\right) \tag{5.7}$$

which describe the motion of a particle in an inertial frame of reference.

The type of observable that is available dictates which of these three equations to use; for example a GPS can measure either $\mathbf{r}(t)$ or $\mathbf{v}(t)$ and an accelerometer or a multi-antenna GPS system can measure acceleration.

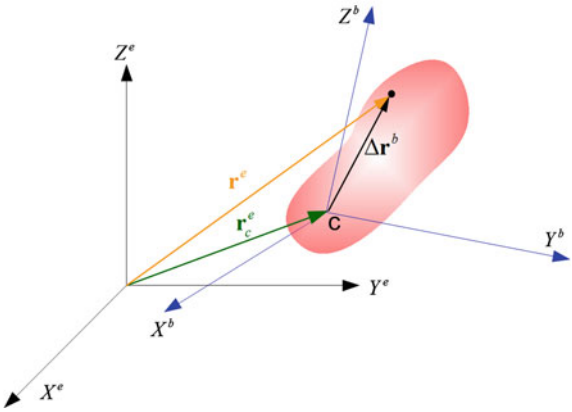


Fig. 5.1 The concept of rigid body motion depicting two reference frames

5.2.1 Rigid Body Motion Modeling

The models previously discussed deal with the motion of a *particle*, whereas real life applications involve 3D objects. For such objects a rigid body motion model is used. A rigid body is a system of particles in which the distances between the particles do not vary. Two systems of coordinates are used to describe the motion of a rigid body; a space-fixed system (in this case the ECEF frame) and a moving system that is rigidly fixed in the body (i.e. the body frame) and participates in its motion. The orientation of the axes of that system relative to the axes of the space-fixed system is specified by three independent angles. As illustrated in Fig. 5.1, rigid body motion can be expressed as the sum of two vectors.

The first vector points from the origin of the ECEF frame to the center of mass c of the body, which is denoted by $\mathbf{r}_c^e$. This represents the translational motion of the body in 3D. The second vector points from the center of mass to any particle in the body and is denoted by $\Delta\mathbf{r}^b$. This denotes the position of an arbitrary point in the body frame. If the b-frame changes its orientation relative to the e-frame, we need the rotation matrix R_b^e to transform the orientation of the rigid body back to the e-frame. Hence the position of the body in the e-frame is

$$\mathbf{r}^e(t) = \mathbf{r}_c^e(t) + R_b^e(t)\Delta\mathbf{r}^b \tag{5.8}$$

This equation is mainly used to describe the position and the attitude of a rigid body with respect to an Earth-fixed reference system.

5.2.2 Observables

The observables are measurements of the quantities of interest, excluding errors. It is very important to determine the appropriate observables for accurate kinematic modeling. The field of *estimation* deals with the optimal extraction of

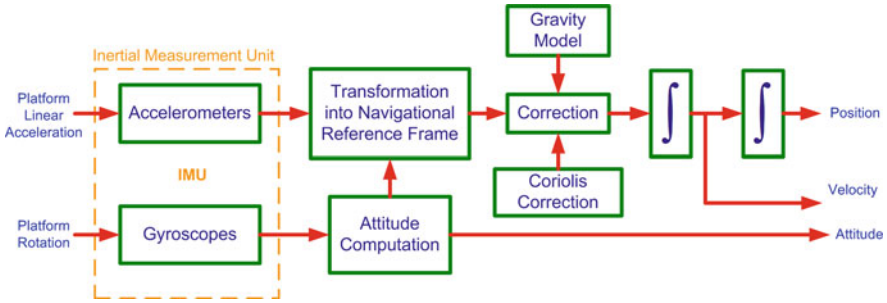


Fig. 5.2 A block diagram depicting the mechanization process of an INS

modeling parameters from the noise inherent in measurements. Measurements fall into two basic categories.

- ***Autonomous measurements*** are self-contained (no external reference is required). Accelerometer and gyroscope measurements are examples of this kind. They have good short term accuracy but are not very accurate in the long run.
- ***Non-autonomous measurements*** need some sort of external reference. They include signals from beacons, satellites, and measurements of road markers. Their accuracy is time-independent but they are constrained by the environments due to their dependency on external signals.

It is evident from Eq. (5.8) that the time-dependent quantities, namely the position vector $\mathbf{r}_c^e(t)$ and the rotation matrix $R_b^e(t)$, are the two observables for rigid body kinematic modeling. So, depending upon the type of accuracy required, autonomous measurements as well as non-autonomous measurements can be used. Being complementary in nature they can both be optimally utilized and integrated together by utilizing an estimation technique. Autonomous measurements include acceleration, angular velocity and tilt, whereas ranges, range rates, directions and angles belong to non-autonomous measurements.

5.3 INS Mechanization

Mechanization is the process of converting the output of an IMU into position, velocity and attitude information. The outputs include rotation rates about three body axes ω_{ib}^b measured by the gyroscopes triad and three specific forces f^b along the body axes measured by the accelerometer triad, all of which are with respect to the inertial frame. Mechanization is a recursive process that starts with a specified set of initial values and iterates on the output. A general diagram of INS mechanization is shown in Fig. 5.2.

5.3.1 INS Mechanization in an Inertial Frame of Reference

The output of an accelerometer is called the *specific force*, and is given as

$$\mathbf{f}^i = \mathbf{a}^i - \bar{\mathbf{g}}^i \tag{5.9}$$

where $\mathbf{f}$ is the specific force, $\mathbf{a}$ is the acceleration of the body, and $\bar{\mathbf{g}}$ is the gravitational vector.

By letting $\mathbf{a}^i = \ddot{\mathbf{r}}^i$ this can be rewritten as

$$\ddot{\mathbf{r}}^i = \mathbf{f}^i + \bar{\mathbf{g}}^i \tag{5.10}$$

where

- $\ddot{\mathbf{r}}^i$ is the second derivative of the position vector measured from the origin of the inertial frame to the moving platform
- $\mathbf{f}^i$ is the specific force
- $\bar{\mathbf{g}}^i$ is the gravitational vector

For ease of solution, the set of three second-order differential equations can be transformed to a set of first-order differential equations as follows

$$\dot{\mathbf{r}}^i = \mathbf{v}^i \tag{5.11}$$

$$\dot{\mathbf{v}}^i = \mathbf{f}^i + \bar{\mathbf{g}}^i \tag{5.12}$$

The measurements are usually made in the body frame. By assuming that the body frame coincides with the sensor triad frame (which is usually so to eliminate a further transformation) these measurements can be transformed into the inertial frame using the transformation matrix R_b^i between the body frame and the inertial frame

$$\mathbf{f}^i = R_b^i \mathbf{f}^b \tag{5.13}$$

where R_b^i is a 3×3 rotation matrix which transforms the measurement from the b-frame to the i-frame.

Since the gravitational vector is usually expressed in either the e-frame or the l-frame, it can be transformed to the i-frame through a rotation matrix R_e^i or R_l^i. The gravitational vector transformation from the e-frame to the i-frame is

$$\bar{\mathbf{g}}^i = R_e^i \bar{\mathbf{g}}^e \tag{5.14}$$

Substituting Eqs. (5.13) and (5.14) into Eq. (5.12) gives

$$\dot{\mathbf{v}}^i = R_b^i \mathbf{f}^b + R_e^i \bar{\mathbf{g}}^e \tag{5.15}$$

As discussed in Chap. 2, the rate of change of a transformation matrix is

$$\dot{R}_b^i = R_b^i \Omega_{ib}^b \tag{5.16}$$

where Ω_{ib}^b is the skew-symmetric matrix form of the vector of angular velocities $\boldsymbol{\omega}_{ib}^b$ output by the gyroscope. This can be represented as

$$\underbrace{\boldsymbol{\omega}_{ib}^b = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}}_{\text{vector}} \rightarrow \underbrace{\Omega_{ib}^b = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}}_{\text{skew-symmeticmatirx}} \tag{5.17}$$

where ω_x, ω_y and ω_z are the gyroscope measurements in the b-frame.

Solving Eq. (5.16) yields the desired transformation matrix R_b^i. Once the elements of the rotation matrix are known, it is possible to calculate the attitude of the body using Euler angles in a similar fashion to Chap. 2.

The mechanization equations for the i-frame can therefore be summarized as

$$\begin{bmatrix} \dot{\mathbf{r}}^i \\ \dot{\mathbf{v}}^i \\ \dot{R}_b^i \end{bmatrix} = \begin{bmatrix} \mathbf{v}^i \\ R_b^i \mathbf{f}^b + R_e^i \bar{\mathbf{g}}^e \\ R_b^i \Omega_{ib}^b \end{bmatrix} \tag{5.18}$$

where the specific force vector $\mathbf{f}^b$ and the angular velocity vector $\boldsymbol{\omega}_{ib}^b$ are sensor measurements. The gravity model in the e-frame $\bar{\mathbf{g}}^e$ is known in advance. The navigation parameters provided by this system are

$\mathbf{r}^i = \begin{bmatrix} x^i & y^i & z^i \end{bmatrix}$ is a 3D position vector in the i-frame

$\mathbf{v}^i = \begin{bmatrix} v_x^i & v_y^i & v_z^i \end{bmatrix}$ is a 3D velocity vectorin the i-frame

R_b^i is a 3×3 matrix containing the information for the three Euler angles.

Figure 5.3 illustrates the mechanization process in the inertial frame.

5.3.2 INS Mechanization in ECEF Frame

A position vector $\mathbf{r}^e$ in the e-frame can be transformed into the i-frame by using the rotation matrix R_e^i as follows

$$\mathbf{r}^i = R_e^i \mathbf{r}^e \tag{5.19}$$

After differentiating twice and rearranging the terms (in accordance with the derivation in Chap. 2) we get

$$\ddot{\mathbf{r}}^i = R_e^i \left(\ddot{\mathbf{r}}^e + 2\Omega_{ie}^e \dot{\mathbf{r}}^e + \dot{\Omega}_{ie}^e \mathbf{r}^e + \Omega_{ie}^e \Omega_{ie}^e \mathbf{r}^e \right) \tag{5.20}$$

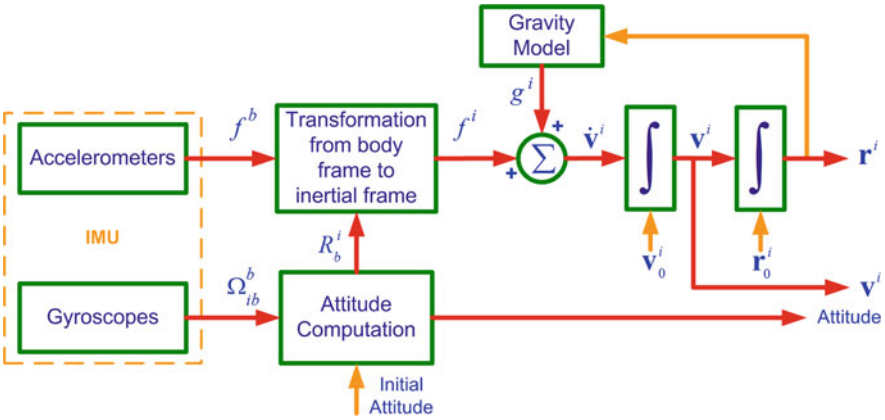


Fig. 5.3 A block diagram depicting the mechanization of an INS in the inertial frame

and substituting this into Eq. (5.10) gives

$$R_e^i\left(\ddot{\mathbf{r}}^e + 2\Omega_{ie}^e\dot{\mathbf{r}}^e + \dot{\Omega}_{ie}^e\mathbf{r}^e + \Omega_{ie}^e\Omega_{ie}^e\mathbf{r}^e\right) = \mathbf{f}^i + \bar{\mathbf{g}}^i \tag{5.21}$$

Further substituting quantities from the right-hand side of Eqs. (5.13) and (5.14) yields

$$R_e^i\left(\ddot{\mathbf{r}}^e + 2\Omega_{ie}^e\dot{\mathbf{r}}^e + \dot{\Omega}_{ie}^e\mathbf{r}^e + \Omega_{ie}^e\Omega_{ie}^e\mathbf{r}^e\right) = R_b^i\mathbf{f}^b + R_e^i\bar{\mathbf{g}}^e \tag{5.22}$$

This can be simplified by letting $R_b^i = R_e^i R_b^e$ and $\dot{\Omega}_{ie}^e\mathbf{r}^e = 0$, because the Earth's rotation rate ω_{ie} is approximately constant

$$R_e^i\left(\ddot{\mathbf{r}}^e + 2\Omega_{ie}^e\dot{\mathbf{r}}^e + \Omega_{ie}^e\Omega_{ie}^e\mathbf{r}^e\right) = R_e^i R_b^e\mathbf{f}^b + R_e^i\bar{\mathbf{g}}^e \tag{5.23}$$

$$\ddot{\mathbf{r}}^e = R_b^e\mathbf{f}^b - 2\Omega_{ie}^e\dot{\mathbf{r}}^e + \bar{\mathbf{g}}^e - \Omega_{ie}^e\Omega_{ie}^e\mathbf{r}^e \tag{5.24}$$

and because the gravity vector is defined as $\mathbf{g}^e = \bar{\mathbf{g}}^e - \Omega_{ie}^e\Omega_{ie}^e\mathbf{r}^e$, this can be further reduced to

$$\ddot{\mathbf{r}}^e = R_b^e\mathbf{f}^b - 2\Omega_{ie}^e\dot{\mathbf{r}}^e + \mathbf{g}^e \tag{5.25}$$

which is second-order and can be broken into the following first-order equations

$$\dot{\mathbf{r}}^e = \mathbf{v}_e \tag{5.26}$$

$$\dot{\mathbf{v}}_e = R_b^e\mathbf{f}^b - 2\Omega_{ie}^e\mathbf{v}^e + \mathbf{g}^e \tag{5.27}$$

The rate of change of the rotation matrix R_b^e can be given as

$$\dot{R}_b^e = R_b^e\Omega_{eb}^b \tag{5.28}$$

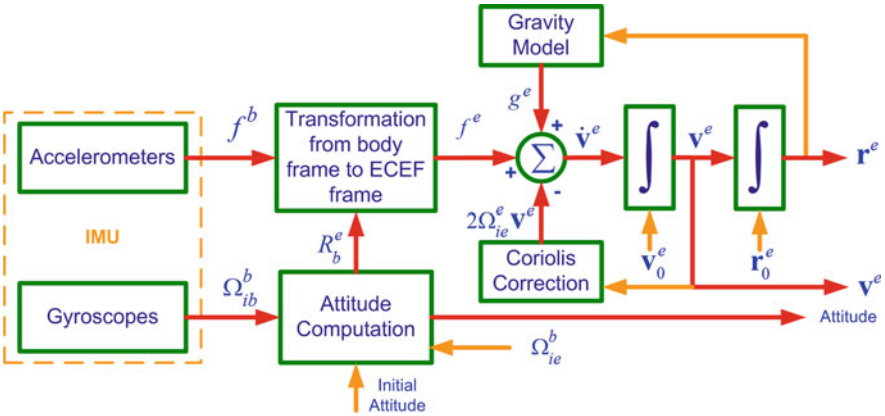


Fig. 5.4 A block diagram depicting the mechanization of an INS in the ECEF frame

As we wish to write the above expression in terms of sensed angular rate ω_{ib}^b monitored by the gyroscopes then we use the following relationship

$$\begin{aligned} \Omega_{ib}^b &= \Omega_{ie}^b + \Omega_{eb}^b \\ \Omega_{eb}^b &= -\Omega_{ie}^b + \Omega_{ib}^b \end{aligned} \tag{5.29}$$

$$\Omega_{eb}^b = \Omega_{ei}^b + \Omega_{ib}^b \tag{5.30}$$

Substituting Eq. (5.30) into Eq. (5.28) yields

$$\dot{R}_b^e = R_b^e \left(\Omega_{ei}^b + \Omega_{ib}^b\right) \tag{5.31}$$

The e-frame mechanization equations can be summarized as

$$\begin{bmatrix} \dot{\mathbf{r}}^e \\ \dot{\mathbf{v}}^e \\ \dot{R}_b^e \end{bmatrix} = \begin{bmatrix} \mathbf{v}^e \\ R_b^e \mathbf{f}^b - 2\Omega_{ie}^e \mathbf{v}^e + \mathbf{g}^e \\ R_b^e \left(\Omega_{ie}^b + \Omega_{ib}^b\right) \end{bmatrix} \tag{5.32}$$

which represents the mechanization equations in the e-frame where the inputs are the sensed accelerations $\mathbf{f}^b$ from the accelerometers and rotation rates ω_{ib}^b from the gyroscopes. The outputs are the position vector $\mathbf{r}$, the velocity vector $\mathbf{v}$, and the Euler angles, all expressed in the e-frame. Figure 5.4 is a block diagram of the mechanization in the ECEF frame.

5.3.3 INS Mechanization in the Local-Level Frame

In many applications the mechanization equations are desired in the LLF for the following reasons:

a. The navigation equations in the l-frame furnish a navigation solution that is intuitive to the user on or near the Earth's surface.
b. Since the axes of the l-frame are aligned to the local east, north and up directions, the attitude angles (pitch, roll and azimuth) can be obtained directly at the output of the mechanization equations when solved in the local-level frame.
c. The computational errors in the navigation parameters on the horizontal (E-N) plane are bound by the Schuler effect (as discussed in Chap. 6).
d. This effect stipulates that the inertial system errors of the horizontal plane components are coupled to produce the Schuler loop, and that these errors oscillate at the Schuler frequency of 1/5,000 Hz.

5.3.3.1 Position Mechanization Equations

The position vector $\mathbf{r}^l$ of a moving platform is expressed in geodetic (curvilinear) coordinates in the ECEF frame as

$$\mathbf{r}^l = [\,\varphi \quad \lambda \quad h\,]^T \tag{5.33}$$

where φ is the latitude, λ is the longitude and h is the altitude. As the platform travels on or near the surface of the Earth, the rate of change of its position is expressed in terms of the velocity in the east, north and up directions.

From Chap. 4, we know that the rate of change of the platform's latitude, longitude and altitude are

$$\dot{\varphi} = \frac{v_n}{R_M + h} \tag{5.34}$$

$$\dot{\lambda} = \frac{v_e}{(R_N + h)\cos\varphi} \tag{5.35}$$

$$\dot{h} = v_u \tag{5.36}$$

where

v_e	is the component of the velocity in the east direction
v_n	is the component of the velocity in the north direction
v_u	is the component of the velocity in the up direction
R_M	is the meridian radius of the ellipsoid
R_N	is the normal radius of the ellipsoid

Therefore using Eqs. 4.28–4.32 the time rate of change of the position components is related to the velocity components (Noureldin et al. 2009; Farrell 2008; Georgy et al. 2010; Iqbal et al. 2009) as follows

$$\begin{bmatrix} \dot{\varphi} \\ \dot{\lambda} \\ \dot{h} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{R_M+h} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_e \\ v_n \\ v_u \end{bmatrix}^l \tag{5.37}$$

$$\dot{\mathbf{r}}^l = D^{-1}\mathbf{v}^l \tag{5.38}$$

in which D^{-1} transforms the velocity vector from rectangular coordinates into curvilinear coordinates in the ECEF frame.

5.3.3.2 Velocity Mechanization Equations

The acceleration of the moving body is measured in three mutually orthogonal directions in the b-frame by a three-axis accelerometer. These measurements are known as specific force measurements and are given in the b-frame as

$$\mathbf{f}^b = \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \tag{5.39}$$

They can be transformed to the local-level frame using the rotation matrix R_b^l

$$\mathbf{f}^\ell = \begin{bmatrix} f^e \\ f^n \\ f^u \end{bmatrix} = R_b^\ell \mathbf{f}^b = R_b^\ell \begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} \tag{5.40}$$

However, for three reasons the acceleration components expressed in the local-level frame $\mathbf{f}^l$ cannot directly yield the velocity components of the moving body:

a. The rotation of the Earth about its spin axis ($\omega^e = 15^\circ/\text{hr}$) is interpreted in the local-level frame as the angular velocity vector $\boldsymbol{\omega}_{ie}^l$ and (as derived in Chap. 4) is given as

$$\boldsymbol{\omega}_{ie}^\ell = \begin{bmatrix} 0 \\ \omega^e \cos\varphi \\ \omega^e \sin\varphi \end{bmatrix} \tag{5.41}$$

b. A change of orientation of the local-level frame with respect to the Earth arises from the definition of the local north and vertical directions. The north direction is tangent to the meridian at all times, while the vertical direction is normal to the Earth's surface. This effect is interpreted by the angular velocity vector ω_{el}^l which is given in Chap. 4 as

$$\boldsymbol{\omega}_{e\ell}^{\ell} = \begin{bmatrix} -\dot{\varphi} \\ \dot{\lambda}\cos\varphi \\ \dot{\lambda}\sin\varphi \end{bmatrix} = \begin{bmatrix} -\frac{v_n}{R_M+h} \\ \frac{v_e}{R_N+h} \\ \frac{v_e\tan\varphi}{R_N+h} \end{bmatrix} \tag{5.42}$$

c. The Earth's gravity field is

$$\mathbf{g}^{\ell} = \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} \tag{5.43}$$

Taking these three factors into the consideration we can derive the expression for the time rate of change of the velocity components of the moving body.

The Earth-referenced velocity vector $\dot{\mathbf{r}}^e$ can be transformed into the local-level frame by using the rotation matrix R_e^l

$$\mathbf{v}^l = R_e^l \dot{\mathbf{r}}^e \tag{5.44}$$

where $\mathbf{v}^l = [v_e, v_n, v_u]^T$.

The time derivative is therefore

$$\dot{\mathbf{v}}^l = \dot{R}_e^l \dot{\mathbf{r}}^e + R_e^l \ddot{\mathbf{r}}^e \tag{5.45}$$

We know from Chap. 2 that the rate of change of the transformation matrix is $\dot{R}_e^l = R_e^l \Omega_{le}^e$, where Ω_{le}^e is the skew-symmetric matrix corresponding to ω_{el}^l. Substituting this into Eq. (5.45) gives

$$\dot{\mathbf{v}}^l = R_e^l \Omega_{le}^e \dot{\mathbf{r}}^e + R_e^l \ddot{\mathbf{r}}^e$$
$$\dot{\mathbf{v}}^l = R_e^l \left(\Omega_{le}^e \dot{\mathbf{r}}^e + \ddot{\mathbf{r}}^e\right)$$

and since $\Omega_{le}^e = -\Omega_{el}^e$ and $\dot{\mathbf{r}}^e = \mathbf{v}^e$

$$\dot{\mathbf{v}}^l = R_e^l \left(\ddot{\mathbf{r}}^e - \Omega_{el}^e \mathbf{v}^e\right) \tag{5.46}$$

We can transform the position vector $\mathbf{r}$ from the ECEF frame into the inertial frame by

$$\mathbf{r}^i = R_e^i \mathbf{r}^e \tag{5.47}$$

and taking the time derivative and using the relationship $\dot{R}_e^i = R_e^i \Omega_{ie}^e$ gives

$$\begin{aligned} \dot{\mathbf{r}}^i &= \dot{R}_e^i \mathbf{r}^e + R_e^i \dot{\mathbf{r}}^e \\ \dot{\mathbf{r}}^i &= R_e^i \Omega_{ie}^e \mathbf{r}^e + R_e^i \dot{\mathbf{r}}^e \\ \dot{\mathbf{r}}^i &= R_e^i \left(\Omega_{ie}^e \mathbf{r}^e + \dot{\mathbf{r}}^e\right) \end{aligned} \tag{5.48}$$

where Ω_{ie}^{e} is the skew-symmetric matrix corresponding to ω_{ie}^{e}.

Taking the time derivative a second time and rearranging the terms in a similar manner to that in Chap. 2 (except that now the rotation is between the e-frame and the i-frame, instead of the b-frame) we arrive at

$$\ddot{\mathbf{r}}^{i} = R_{e}^{i}\left(\ddot{\mathbf{r}}^{e} + 2\Omega_{ie}^{e}\dot{\mathbf{r}}^{e} + \dot{\Omega}_{ie}^{e}\mathbf{r}^{e} + \Omega_{ie}^{e}\Omega_{ie}^{e}\mathbf{r}^{e}\right) \tag{5.49}$$

and because the Earth's rotation rate is approximately constant, $\dot{\Omega}_{ie}^{e} = 0$ and hence

$$\ddot{\mathbf{r}}^{i} = R_{e}^{i}\left(\ddot{\mathbf{r}}^{e} + 2\Omega_{ie}^{e}\dot{\mathbf{r}}^{e} + \Omega_{ie}^{e}\Omega_{ie}^{e}\mathbf{r}^{e}\right) \tag{5.50}$$

From Newton's second law we also have the relationship

$$\ddot{\mathbf{r}}^{i} = \mathbf{f}^{i} + \bar{\mathbf{g}}^{i} \tag{5.51}$$

where $\mathbf{f}^{i}$ is the specific force and $\bar{\mathbf{g}}^{i}$ is the gravitational field vector, both given in the inertial frame.

Substituting the value of $\ddot{\mathbf{r}}^{i}$ from Eq. (5.51) into (5.50) gives

$$\begin{aligned}
\mathbf{f}^{i} + \mathbf{g}^{i} &= R_{e}^{i}\left(\ddot{\mathbf{r}}^{e} + 2\Omega_{ie}^{e}\dot{\mathbf{r}}^{e} + \Omega_{ie}^{e}\Omega_{ie}^{e}\mathbf{r}^{e}\right) \\
R_{i}^{e}\left(\mathbf{f}^{i} + \bar{\mathbf{g}}^{i}\right) &= \left(\ddot{\mathbf{r}}^{e} + 2\Omega_{ie}^{e}\dot{\mathbf{r}}^{e} + \Omega_{ie}^{e}\Omega_{ie}^{e}\mathbf{r}^{e}\right) \\
\mathbf{f}^{e} + \bar{\mathbf{g}}^{e} &= \left(\ddot{\mathbf{r}}^{e} + 2\Omega_{ie}^{e}\dot{\mathbf{r}}^{e} + \Omega_{ie}^{e}\Omega_{ie}^{e}\mathbf{r}^{e}\right) \\
\ddot{\mathbf{r}}^{e} &= \mathbf{f}^{e} + \bar{\mathbf{g}}^{e} - \Omega_{ie}^{e}\Omega_{ie}^{e}\mathbf{r}^{e} - 2\Omega_{ie}^{e}\dot{\mathbf{r}}^{e}
\end{aligned} \tag{5.52}$$

The gravitational field vector $\bar{\mathbf{g}}^{e}$ and the gravity field vector $\mathbf{g}^{e}$ are related by

$$\mathbf{g}^{e} = \bar{\mathbf{g}}^{e} - \Omega_{ie}^{e}\Omega_{ie}^{e}\mathbf{r}^{e} \tag{5.53}$$

and substituting both this expression for $\mathbf{g}^{e}$ and the fact that $\dot{\mathbf{r}}^{e} = \mathbf{v}^{e}$ into Eq. (5.52) gives

$$\ddot{\mathbf{r}}^{e} = \mathbf{f}^{e} + \mathbf{g}^{e} - 2\Omega_{ie}^{e}\mathbf{v}^{e} \tag{5.54}$$

Substituting this expression for $\ddot{\mathbf{r}}^{e}$ into Eq. (5.46) gives

$$\begin{aligned}
\dot{\mathbf{v}}^{l} &= R_{e}^{l}\left(\mathbf{f}^{e} + \mathbf{g}^{e} - 2\Omega_{ie}^{e}\mathbf{v}^{e} - \Omega_{el}^{e}\mathbf{v}^{e}\right) \\
\dot{\mathbf{v}}^{l} &= R_{e}^{l}\left\{\mathbf{f}^{e} + \mathbf{g}^{e} - \left(2\Omega_{ie}^{e} + \Omega_{el}^{e}\right)\mathbf{v}^{e}\right\} \\
\dot{\mathbf{v}}^{l} &= R_{e}^{l}\mathbf{f}^{e} + R_{e}^{l}\mathbf{g}^{e} - R_{e}^{l}\left(2\Omega_{ie}^{e} + \Omega_{el}^{e}\right)\mathbf{v}^{e} \\
\dot{\mathbf{v}}^{l} &= \mathbf{f}^{l} + \mathbf{g}^{l} - R_{e}^{l}\left(2R_{l}^{e}\Omega_{ie}^{l}R_{e}^{l} + R_{l}^{e}\Omega_{el}^{l}R_{e}^{l}\right)\mathbf{v}^{e} \\
\dot{\mathbf{v}}^{l} &= R_{b}^{l}\mathbf{f}^{b} + \mathbf{g}^{l} - \left(2R_{e}^{l}R_{l}^{e}\Omega_{ie}^{l} + R_{e}^{l}R_{l}^{e}\Omega_{el}^{l}\right)R_{e}^{l}\mathbf{v}^{e} \\
\dot{\mathbf{v}}^{l} &= R_{b}^{l}\mathbf{f}^{b} - \left(2\Omega_{ie}^{l} + \Omega_{el}^{l}\right)\mathbf{v}^{l} + \mathbf{g}^{l}
\end{aligned} \tag{5.55}$$

where R_{b}^{l} is the transformation matrix from the b-frame to the l-frame, $\mathbf{f}^{b}$ is the specific force measured by the accelerometers in the b-frame, and $\mathbf{g}^{l}$ is the gravity

vector in l-frame. As mentioned earlier, Ω_{ie}^{l} and Ω_{el}^{l} are the skew-symmetric matrices corresponding to ω_{ie}^{l} and ω_{el}^{l} and are expressed as

$$\boldsymbol{\omega}_{ie}^{l} = \begin{bmatrix} 0 \\ \omega^e \cos\varphi \\ \omega^e \sin\varphi \end{bmatrix} \rightarrow \Omega_{ie}^{\ell} = \begin{bmatrix} 0 & -\omega^e \sin\varphi & \omega^e \cos\varphi \\ \omega^e \sin\varphi & 0 & 0 \\ -\omega^e \cos\varphi & 0 & 0 \end{bmatrix} \tag{5.56}$$

$$\boldsymbol{\omega}_{el}^{l} = \begin{bmatrix} -\dfrac{v_n}{R_M + h} \\ \dfrac{v_e}{R_N + h} \\ \dfrac{v_e \tan\varphi}{R_N + h} \end{bmatrix} \rightarrow \Omega_{e\ell}^{\ell} = \begin{bmatrix} 0 & \dfrac{-v_e \tan\varphi}{R_N + h} & \dfrac{v_e}{R_N + h} \\ \dfrac{v_e \tan\varphi}{R_N + h} & 0 & \dfrac{v_n}{R_M + h} \\ \dfrac{-v_e}{R_N + h} & \dfrac{-v_n}{R_M + h} & 0 \end{bmatrix} \tag{5.57}$$

5.3.3.3 Attitude Mechanization Equations

The attitude (orientation) of the moving body is determined by solving the time derivative equation of the transformation matrix R_b^l that relates the body frame to the local-level frame. For local-level mechanization the following time derivative equation of the transformation matrix should be considered

$$\dot{R}_b^l = R_b^l \Omega_{lb}^b \tag{5.58}$$

where the angular velocity skew-symmetric matrix Ω_{lb}^b can be expressed as

$$\begin{aligned} \Omega_{lb}^b &= \Omega_{li}^b + \Omega_{ib}^b \\ \Omega_{lb}^b &= -\Omega_{il}^b + \Omega_{ib}^b \\ \Omega_{lb}^b &= \Omega_{ib}^b - \Omega_{il}^b \end{aligned} \tag{5.59}$$

Substituting this into Eq. (5.58) gives

$$\dot{R}_b^l = R_b^l (\Omega_{ib}^b - \Omega_{il}^b) \tag{5.60}$$

The rotation matrix can be obtained by solving Eq. (5.60) for the attitude angles.

The quantity Ω_{ib}^b, which is the rate of rotation of the b-frame with respect to the i-frame, is directly measured by the gyroscopes. However, in addition to the angular velocities of the moving body the gyroscopic measurements contain both the Earth's rotation Ω_{ie}^b and the change in orientation of the LLF. So Ω_{il}^b must be subtracted from Ω_{ib}^b to remove these affects. Ω_{il}^b is composed of two parts: Ω_{ie}^b, which is the Earth's rotation rate with respect to the i-frame expressed in the body

frame and Ω_{el}^{b}, which is the change in the orientation of the LLF with respect to the ECEF frame as expressed in body frame. Therefore

$$\Omega_{il}^{b} = \Omega_{ie}^{b} + \Omega_{el}^{b} \tag{5.61}$$

where

$$\Omega_{ie}^{b} = R_{l}^{b}\Omega_{ie}^{l}R_{b}^{l}$$

$$\Omega_{el}^{b} = R_{l}^{b}\Omega_{el}^{l}R_{b}^{l}$$

giving

$$\Omega_{il}^{b} = R_{l}^{b}\Omega_{ie}^{l}R_{b}^{l} + R_{l}^{b}\Omega_{el}^{l}R_{b}^{l} \tag{5.62}$$

$$\Omega_{il}^{b} = R_{l}^{b}(\Omega_{ie}^{l} + \Omega_{el}^{l})R_{b}^{l} \tag{5.63}$$

By substituting this into Eq. (5.60) we get

$$\dot{R}_{b}^{l} = R_{b}^{l}\left[\Omega_{ib}^{b} - R_{l}^{b}(\Omega_{ie}^{l} + \Omega_{el}^{l})R_{b}^{l}\right] \tag{5.64}$$

where Ω_{ib}^{b} is the skew-symmetric matrix corresponding to the gyroscopic measurement vector.

5.3.3.4 Summary of INS Mechanization Equations

The results of the previous sections can be summarized as follows

$$\begin{bmatrix} \dot{\mathbf{r}}^{l} \\ \dot{\mathbf{v}}^{l} \\ \dot{R}_{b}^{l} \end{bmatrix} = \begin{bmatrix} D^{-1}\mathbf{v}^{l} \\ R_{b}^{l}\mathbf{f}^{b} - (2\Omega_{ie}^{l} + \Omega_{el}^{l})\mathbf{v}^{l} + \mathbf{g}^{l} \\ R_{b}^{l}(\Omega_{ib}^{b} - \Omega_{il}^{b}) \end{bmatrix} \tag{5.65}$$

which expresses the mechanization in the local-level frame. The position output is expressed in ECEF curvilinear coordinates φ, λ, h, the velocity output is in l-frame coordinates v_e, v_n, v_u, and the attitude angles (roll, pitch and yaw) are measured with respect to the l-frame. Figure 5.5 is a block diagram of local-level frame mechanization.

5.3.4 INS Mechanization in Wander Frame

The l-frame rotates continuously as it moves over the curved surface of the Earth because its y-axis always points toward the north (tangential to the meridian). As was explained in Sect. 2.24 and is reiterated in Fig. 5.6, this rate of rotation becomes ever greater as the l-frame approaches the pole and will become infinite if the l-frame passes directly over the pole.

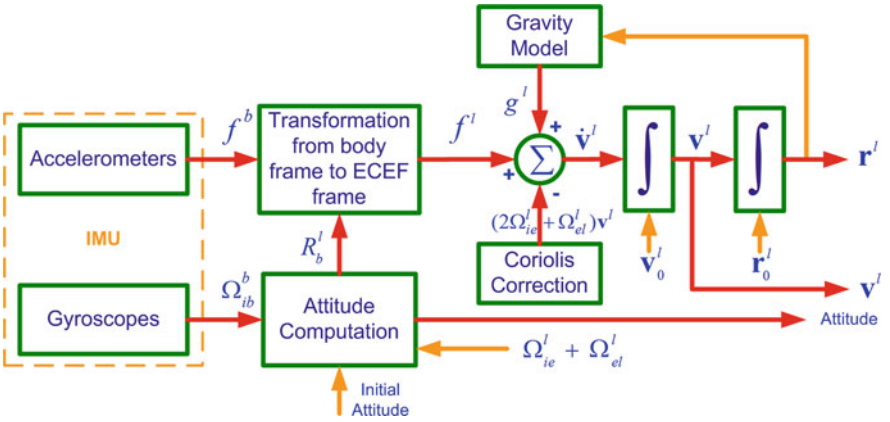


Fig. 5.5 A block diagram depicting the mechanization of an INS in the local-level frame

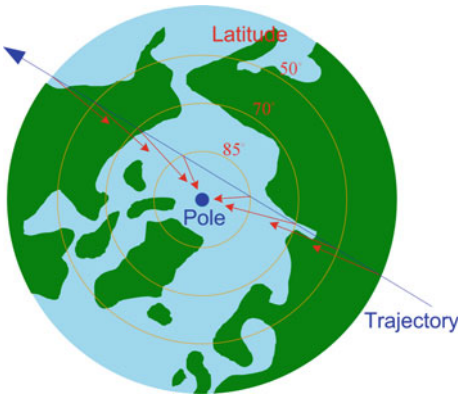


Fig. 5.6 The rotation of the y-axis of the local-level frame (*dark arrows*) for a near-pole crossing trajectory at various latitudes

The rotational rate of the navigation l-frame over the Earth's surface (known as the transport rate) is

$$\boldsymbol{\omega}_{el}^{l} = \begin{bmatrix} \omega_e \\ \omega_n \\ \omega_u \end{bmatrix} = \begin{bmatrix} -\dot{\varphi} \\ \dot{\lambda}\cos\varphi \\ \dot{\lambda}\sin\varphi \end{bmatrix} = \begin{bmatrix} -\dfrac{v_n}{R_M + h} \\ \dfrac{v_e}{R_N + h} \\ \dfrac{v_e \tan\varphi}{R_N + h} \end{bmatrix} \tag{5.66}$$

where ω_e, ω_n and ω_u are the east, north and up angular velocity components as in Fig. 5.7.

It is evident from Eq. (5.66) that the third component of the above vector will introduce numerical instabilities as φ approaches $\pi/2$ and will actually be indeterminate at the pole. This condition is avoided by using the wander azimuth frame

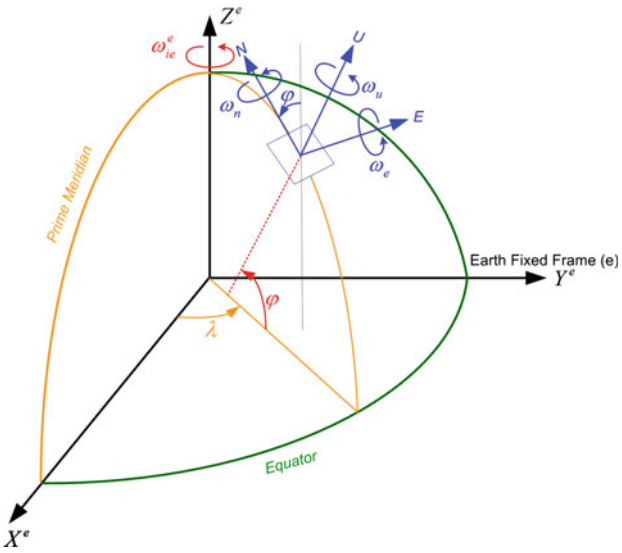


Fig. 5.7 A depiction of the rotational velocity components experienced in the l-frame

in which the third component of Eq. (5.66) is forced to zero and the y-axis of the w-frame will deviate from true north by an angle α, referred to as the *wander azimuth* angle (Bekir 2007). This angle is initialized when initiating the trajectory. In the wander frame, the wander azimuth angle between the true north and the y-axis of the frame varies as the vehicle moves. The output parameters of the w-frame mechanization are eventually transformed into the l-frame.

The angular rate vector of the wander frame with respect to the l-frame can be expressed (Jekeli 2001) as

$$\boldsymbol{\omega}_{lw}^{l} = [0 \quad 0 \quad \dot{\alpha}]^T \tag{5.67}$$

and the angular rate of the wander frame with respect to the e-frame is

$$\boldsymbol{\omega}_{ew}^{l} = \boldsymbol{\omega}_{el}^{l} + \boldsymbol{\omega}_{lw}^{l} \tag{5.68}$$

$$\begin{aligned} \boldsymbol{\omega}_{ew}^{l} &= \begin{bmatrix} -\dot{\varphi} \\ \dot{\lambda}\cos\varphi \\ \dot{\lambda}\sin\varphi \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \dot{\alpha} \end{bmatrix} \\ \boldsymbol{\omega}_{ew}^{l} &= \begin{bmatrix} -\dot{\varphi} \\ \dot{\lambda}\cos\varphi \\ \dot{\lambda}\sin\varphi + \dot{\alpha} \end{bmatrix} \end{aligned} \tag{5.69}$$

and therefore the rotation rate of the w-frame with respect to the e-frame, resolved in the w-frame, is

$$\boldsymbol{\omega}_{ew}^{w} = R_{l}^{w}\boldsymbol{\omega}_{ew}^{l} \tag{5.70}$$

where R_l^w is the rotation matrix from the l-frame to the w-frame. It is simply the transpose of the matrix R_w^l defined in Chap. 2 as

$$R_{w}^{l} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.71}$$

To force the third component ω_u of Eq. (5.66) to be zero we must ensure that

$$\dot{\alpha} = -\dot{\lambda}\sin\varphi \tag{5.72}$$

The mechanization equations are equivalent to those of the l-frame except that all the notations are for the w-frame rather than for the l-frame

$$\begin{bmatrix} \dot{\mathbf{r}}^{w} \\ \dot{\mathbf{v}}^{w} \\ \dot{R}_{b}^{w} \end{bmatrix} = \begin{bmatrix} D^{-1}\mathbf{v}^{w} \\ R_{b}^{w}\mathbf{f}^{b} - (2\Omega_{ie}^{w} + \Omega_{ew}^{w})\mathbf{v}^{w} + \mathbf{g}^{w} \\ R_{b}^{w}(\Omega_{ib}^{b} - \Omega_{iw}^{b}) \end{bmatrix} \tag{5.73}$$

5.4 Parameterization of the Rotation Matrix

The solution of the mechanization equations requires the parameterization of the rotation matrix R_b^l. The three most common methods are Euler angles, direction cosines and the quaternion. Euler angles require only three independent parameters. Direction cosines require nine parameters, six of which are independent. Both of these methods are computationally expensive and therefore inappropriate for real-time computations. Furthermore, Euler angles are prone to singularities. The six independent kinematic equations involved in the derivative of the rotation matrix $\dot{R}_x^y = R_x^y\Omega_{yx}^x$ cannot be solved in closed form, and require numerical integration. The most effective way of parameterizing the transformation matrix is therefore to use the quaternion method.

5.4.1 Solution to Transformation Matrix

The time rate of change of a transformation matrix from the body frame into a computational frame k is

$$\dot{R}_{b}^{k} = R_{b}^{k}\Omega_{kb}^{b} \tag{5.74}$$

For the general case (and for simplicity) we shall delete the superscripts and subscripts

$$\dot{R} = R\Omega \tag{5.75}$$

where R is the transformation matrix from the body frame to the computational frame, and Ω is the skew-symmetric matrix associated with the angular velocities $\omega_x, \omega_y, \omega_z$ of the body frame with respect to the computational frame

$$\Omega = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \tag{5.76}$$

Equation (5.75) requires the solution of nine differential equations in order to obtain the transformation matrix from the angular velocity data. A closed form solution of this equation will now be discussed.

Assuming the angular velocity $\boldsymbol{\omega}$ is constant over the small time interval Δt, then the small incremental angular changes of the rotation of the body frame with respect to the computation frame k is

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \end{bmatrix} = \int \boldsymbol{\omega} dt = \boldsymbol{\omega}\Delta t = \begin{bmatrix} \omega_x \Delta t \\ \omega_y \Delta t \\ \omega_z \Delta t \end{bmatrix} \tag{5.77}$$

and the solution to Eq. (5.75) can be written (see Appendix A) in a recursive form as

$$\begin{aligned} R_{k+1} &= R_k e^{\int \Omega dt} \\ &= R_k e^{\Omega \Delta t} \\ &= R_k \sum_{n=0}^{\infty} \frac{(\Omega \Delta t)^n}{n!} \end{aligned}$$

$$R_{k+1} = R_k \sum_{n=0}^{\infty} \frac{S^n}{n!} \tag{5.78}$$

where $S = \Omega \Delta t$ is the skew-symmetric matrix of vector $\boldsymbol{\theta} = \left(\theta_x, \theta_y, \theta_z\right)^T$ such that

$$\begin{aligned} S &= \begin{bmatrix} 0 & -\omega_z \Delta t & \omega_y \Delta t \\ \omega_z \Delta t & 0 & -\omega_x \Delta t \\ -\omega_y \Delta t & \omega_x \Delta t & 0 \end{bmatrix} \\ S &= \begin{bmatrix} 0 & -\theta_z & \theta_y \\ \theta_z & 0 & -\theta_x \\ -\theta_y & \theta_x & 0 \end{bmatrix} \end{aligned} \tag{5.79}$$

The matrix S has the following characteristics

$$S^3 = -|\boldsymbol{\theta}|^2 S;\ S^4 = -|\boldsymbol{\theta}|^2 S^2; S^5 = |\boldsymbol{\theta}|^4 S;\ S^6 = |\boldsymbol{\theta}|^4 S^2 \tag{5.80}$$

where $|\boldsymbol{\theta}|^2 = \left(\omega_x^2 + \omega_y^2 + \omega_z^2\right)\Delta t^2$.

Equation (5.78) can be expanded as

$$R_{k+1} = R_k\left[I + \frac{S}{1!} + \frac{S^2}{2!} + \frac{S^3}{3!} + \frac{S^4}{4!} + \frac{S^5}{5!} + \ldots\right] \tag{5.81}$$

and in view of the characteristics of Eq. (5.80) this becomes

$$R_{k+1} = R_k\left[I + S + \frac{S^2}{2!} - \frac{|\boldsymbol{\theta}|^2}{3!}S - \frac{|\boldsymbol{\theta}|^2}{4!}S^2 + \frac{|\boldsymbol{\theta}|^4}{5!}S + \frac{|\boldsymbol{\theta}|^4}{6!}S^2 - \ldots\right] \tag{5.82}$$

which can be rearranged as follows

$$R_{k+1} = R_k\left[I + \left(\frac{1}{2!} - \frac{1}{4!}|\boldsymbol{\theta}|^2 + \frac{1}{6!}|\boldsymbol{\theta}|^4 - \ldots\right)S^2 + \left(1 - \frac{1}{3!}|\boldsymbol{\theta}|^2 + \frac{1}{5!}|\boldsymbol{\theta}|^4 - \ldots\right)S\right] \tag{5.83}$$

We know from the definitions of the sine and cosine series that

$$\begin{aligned} \sin\theta &= \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \ldots \\ \frac{\sin\theta}{\theta} &= 1 - \frac{\theta^2}{3!} + \frac{\theta^4}{5!} - \frac{\theta^6}{7!} + \ldots \end{aligned} \tag{5.84}$$

and

$$\begin{aligned} \cos\theta &= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \ldots \\ 1 - \cos\theta &= \frac{\theta^2}{2!} - \frac{\theta^4}{4!} + \frac{\theta^6}{6!} + \ldots \\ \frac{1 - \cos\theta}{\theta^2} &= \frac{1}{2!} - \frac{\theta^2}{4!} + \frac{\theta^4}{6!} + \ldots \end{aligned} \tag{5.85}$$

Equations (5.84) and (5.85) enable Eq. (5.83) to be rewritten as

$$R_{k+1} = R_k\left(I + \frac{\sin|\boldsymbol{\theta}|}{|\boldsymbol{\theta}|}S + \frac{1 - \cos|\boldsymbol{\theta}|}{|\boldsymbol{\theta}|^2}S^2\right) \tag{5.86}$$

and by letting $s = \frac{\sin|\boldsymbol{\theta}|}{|\boldsymbol{\theta}|}$ and $c = \frac{1-\cos|\boldsymbol{\theta}|}{|\boldsymbol{\theta}|^2}$ we get

$$R_{k+1} = R_k\left(I + sS + cS^2\right) \tag{5.87}$$

The above equation allows the direct determination of the transformation matrix in the computational frame from the observables (rotation $\boldsymbol{\omega}$) in the body frame.

However, this equation offers no particular advantage over the quaternion approach (discussed next). It may be noted that the first-order approximation of the Eq. (5.87) for a small value of angle $\boldsymbol{\theta}$ is

$$R_{k+1} = R_k(I + S) \tag{5.88}$$

5.4.2 Quaternions

Solving the mechanization equations requires the parameterization of the rotation matrix R_b^l. The most popular method is the quaternion approach (Kuipers 1999). Euler's theorem states that the rotation of a rigid body (represented by the body frame) with respect to a reference frame (in this case the computational frame) can be expressed in terms of a rotation angle θ about a fixed axis and the direction cosines of the rotation axis that define the rotation direction. Figure 5.8 represents a quaternion where θ is the magnitude of the rotation and α, β and γ define the orientation of the unit vector '$\mathbf{n}$' that lies along the axis of rotation.

A quaternion is a four-parameter representation of a transformation matrix that is defined (Rogers 2007) as follows

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} \frac{\theta_x}{\theta} \sin\frac{\theta}{2} \\ \frac{\theta_y}{\theta} \sin\frac{\theta}{2} \\ \frac{\theta_z}{\theta} \sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{bmatrix} \tag{5.89}$$

where $\theta = \sqrt{\theta_x^2 + \theta_y^2 + \theta_z^2}$ is the rotation angle, and $\frac{\theta_x}{\theta}, \frac{\theta_y}{\theta}$ and $\frac{\theta_z}{\theta}$ are direction cosines of the rotation axis with respect to the computational frame.

The components of a quaternion are related by the constraint

$$q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1 \tag{5.90}$$

This indicates that three independent quaternion components are sufficient to describe the rotation of a rigid body.

The quaternion parameters are functions of time, and the associated differential equation is

$$\dot{\mathbf{q}} = \frac{1}{2}\bar{\Omega}(\omega)\mathbf{q} \tag{5.91}$$

where $\bar{\Omega}(\omega)$ is the skew-symmetric matrix of the following form

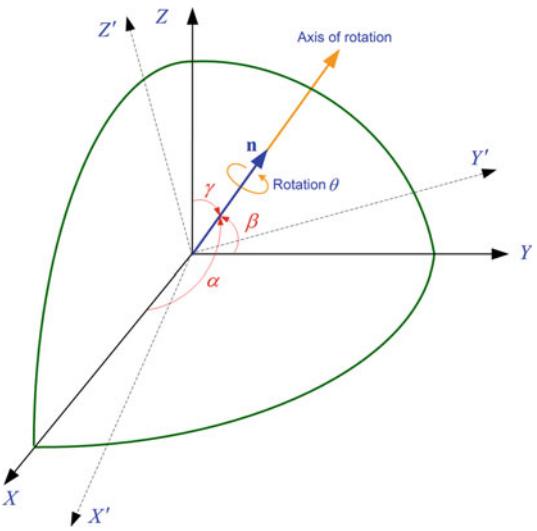


Fig. 5.8 Spatial representation of a quaternion in relation to the reference frame XYZ

$$\bar{\Omega}(\omega) = \begin{bmatrix} 0 & \omega_z & -\omega_y & \omega_x \\ -\omega_z & 0 & \omega_x & \omega_y \\ \omega_y & -\omega_x & 0 & \omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix} = \begin{bmatrix} -\Omega_{3\times 3} & \vdots & \boldsymbol{\omega}_{3\times 1} \\ \cdots & \cdots & \cdots \\ -\boldsymbol{\omega}_{1\times 3}^T & \vdots & 0 \end{bmatrix} \tag{5.92}$$

in which $\boldsymbol{\omega} = \left(\omega_x, \omega_y, \omega_z\right)^T$ is the angular velocity of the body rotation and Ω is the skew-symmetric form of $\boldsymbol{\omega}$.

5.4.2.1 Relationship Between the Transformation Matrix and Quaternion Parameters

Once the quaternion parameters are known at a certain time, the rotation matrix R can be obtained using the following direct relationship

$$R = \begin{bmatrix} R(1,1) & R(1,2) & R(1,3) \\ R(2,1) & R(2,2) & R(2,3) \\ R(3,1) & R(3,2) & R(3,3) \end{bmatrix}$$

$$= \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 - q_3q_4) & 2(q_1q_3 + q_2q_4) \\ 2(q_1q_2 + q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 - q_1q_4) \\ 2(q_1q_3 - q_2q_4) & 2(q_2q_3 + q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \tag{5.93}$$

After determining the initial rotation matrix from the attitude angles measured during the alignment process, the initial values of the quaternion are calculated as

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}_{t_0} = \begin{bmatrix} 0.25\{R(3,2) - R(2,3)\}/q_4 \\ 0.25\{R(1,3) - R(3,1)\}/q_4 \\ 0.25\{R(2,1) - R(1,2)\}/q_4 \\ 0.5\sqrt{1 + R(1,1) + R(2,2) + R(3,3)} \end{bmatrix} \tag{5.94}$$

5.4.3 Solutions of the Quaternion Equation

5.4.3.1 Discrete Closed form (Analytical) Solution

During a short interval of time Δt, the angular velocity of the rotation $\boldsymbol{\omega}$ can be presumed constant, and the closed form of the discrete solution to the quaternion Eq. (5.91) is

$$\begin{aligned} \mathbf{q}_{k+1} &= e^{\int \frac{1}{2}\bar{\Omega}(\omega)dt}\mathbf{q}_k \\ \mathbf{q}_{k+1} &= e^{\frac{1}{2}\bar{\Omega}(\omega)\Delta t}\mathbf{q}_k \end{aligned}$$

$$\mathbf{q}_{k+1} = \left(\sum_{n=0}^{\infty} \frac{1}{2^n n!}\bar{\Omega}^n(\omega)\Delta t^n\right)\mathbf{q}_k \tag{5.95}$$

By the definition

$$\begin{aligned} \bar{S} &= \bar{\Omega}(\omega)\Delta t \\ \bar{S} &= \begin{bmatrix} 0 & \omega_z & -\omega_y & \omega_x \\ -\omega_z & 0 & \omega_x & \omega_y \\ \omega_y & -\omega_x & 0 & \omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix}\Delta t \\ \bar{S} &= \begin{bmatrix} 0 & \theta_z & -\theta_y & \theta_x \\ -\theta_z & 0 & \theta_x & \theta_y \\ \theta_y & -\theta_x & 0 & \theta_z \\ -\theta_x & -\theta_y & -\theta_z & 0 \end{bmatrix} \end{aligned} \tag{5.96}$$

it may be noted that the matrix $\bar{S}$ has the following properties

$$\begin{aligned} \bar{S}^2 &= -\theta^2 I \\ \bar{S}^3 &= -\theta^2 \bar{S}; \quad \bar{S}^4 = -\theta^2 \bar{S}^2 \\ \bar{S}^5 &= \theta^4 \bar{S}; \quad \bar{S}^6 = \theta^4 \bar{S}^2 \end{aligned} \tag{5.97}$$

and therefore Eq. (5.95) becomes

$$\mathbf{q}_{k+1} = \left(\sum_{n=0}^{\infty} \frac{1}{2^n} \bar{S}^n \right) \mathbf{q}_k \tag{5.98}$$

As discussed earlier for the rotation matrix R, and from the above mentioned properties, a similar approach can be adopted to solve Eq. (5.98) as

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \frac{1}{2} \left[2 \left(\cos \frac{\theta}{2} - 1 \right) I + \frac{2}{\theta} \sin \frac{\theta}{2} \bar{S}(\omega) \right] \mathbf{q}_k \tag{5.99}$$

where $\theta = \sqrt{\theta_x^2 + \theta_y^2 + \theta_z^2}$.

5.4.3.2 Numerical Integration Methods

If the rotation rate is slow, then standard numerical integration algorithms can also be used to solve the differential Eqs. (5.75) and (5.91).

Euler's Method

For a first-order differential equation

$$\frac{dy}{dx} = f[x, y(x)],\ y(x_0) = y_0 \tag{5.100}$$

the solution using Euler's method is

$$\begin{aligned} \frac{y_{k+1} - y_k}{\Delta x} &= f(x_k, y_k) \\ y_{k+1} &= y_k + f(x_k, y_k) \Delta x \end{aligned} \tag{5.101}$$

Applying this concept to the quaternion equation $\dot{\mathbf{q}} = \frac{1}{2} \bar{\Omega}(\omega) \mathbf{q}$ gives

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \frac{1}{2} \bar{\Omega}(\omega_k) \mathbf{q}_k \Delta t \tag{5.102}$$

Since truncating the series expansion of analytical solution (5.95) yields the same result as Eq. (5.102), this represents the first-order approximation to the analytical solution given by Eq. (5.99).

5.4.4 Advantages of Quaternion

The quaternion method offers advantages (Rogers 2007) for the parameterization of the rotation matrix R_b^l, namely:

a. Only four equations are solved numerically instead of six differential equations when the rotation matrix R_b^l is manipulated directly by using directions cosines.
b. The quaternion solution avoids the singularity problem that can arise when Euler angles are used.
c. The quaternion computation is relatively simple to perform.

The disadvantages of the quaternion method include the presence of non-linear terms in the result and the need for renormalizations in the computational cycles. Table 5.1 lists the pros and cons of parameterizing the transformation matrix using the three methods mentioned.

5.5 Step by Step Computation of Navigation Parameters in the l-Frame

Owing to the advantages offered by mechanization in the l-frame (as discussed in Sect. 5.3.3) many applications prefer to implement mechanization in this frame of reference. Before we delve into the mathematics, let us review the steps of the mechanization process in the l-frame.

a. Obtain rotation rates $(\omega_x, \omega_y, \omega_z)$ from the gyroscopes and accelerations (f_x, f_y, f_z) from the accelerometers. These measurements are in relation to the inertial frame resolved into the body frame, and they constitute the IMU outputs.
b. Calculate the attitude angles of pitch, roll and azimuth (p, r, A) in terms of the rotation rates $(\omega_x, \omega_y, \omega_z)$. This involves the computation of R_b^l.
c. Use R_b^l computed by the previous step to transform the specific forces in the body frame to the navigation frame, yielding the accelerations in the local-level frame (f_e, f_n, f_u).
d. Since accelerometers also measure gravity and Coriolis forces, we must compensate for these effects.
e. Calculate the east, north and up velocities (V_e, V_n, V_u) by integrating the transformed specific forces (f_e, f_n, f_u).
f. Calculate the geodetic coordinates of the position (φ, λ, h) by integrating the velocities.

The mechanization equations in the l-frame are reproduced here for reference and to assist understanding of the ensuing material

$$\begin{bmatrix} \dot{\mathbf{r}}^l \\ \dot{\mathbf{v}}^l \\ \dot{R}_b^l \end{bmatrix} = \begin{bmatrix} D^{-1}\mathbf{v}^l \\ R_b^l \mathbf{f}^b - (2\Omega_{ie}^l + \Omega_{el}^l)\mathbf{v}^l + \mathbf{g}^l \\ R_b^l(\Omega_{ib}^b - \Omega_{il}^b \end{bmatrix}$$

and mechanization in the l-frame is summarized the Figure 5.9.

Table 5.1 A summary of the characteristics of various methods for the parameterization of the transformation matrix

Methods	Advantages	Disadvantages
Euler Angles	Only three differential equations are needed (three independent parameters)	Non-linear differential equations
	Direct initialization from roll, pitch and yaw angles.	Singularity occurs as the angles approach $\pm 90^{\circ}$
		Order of rotation is important
Direction Cosines	Linear differential equations	Six independent differential equations
	No singularities	Computationally complex
	Direction computation of the transformation matrix	Euler angles are not directly available
Quaternion	Only four differential equations	Euler angles are not directly available
	No singularities	Transformation matrix is not directly available
	Simple computation	Initial conditions using Euler angles are required
	Only three differential equations are needed (three independent parameters)	
	Direct initialization from roll, pitch and yaw angles	

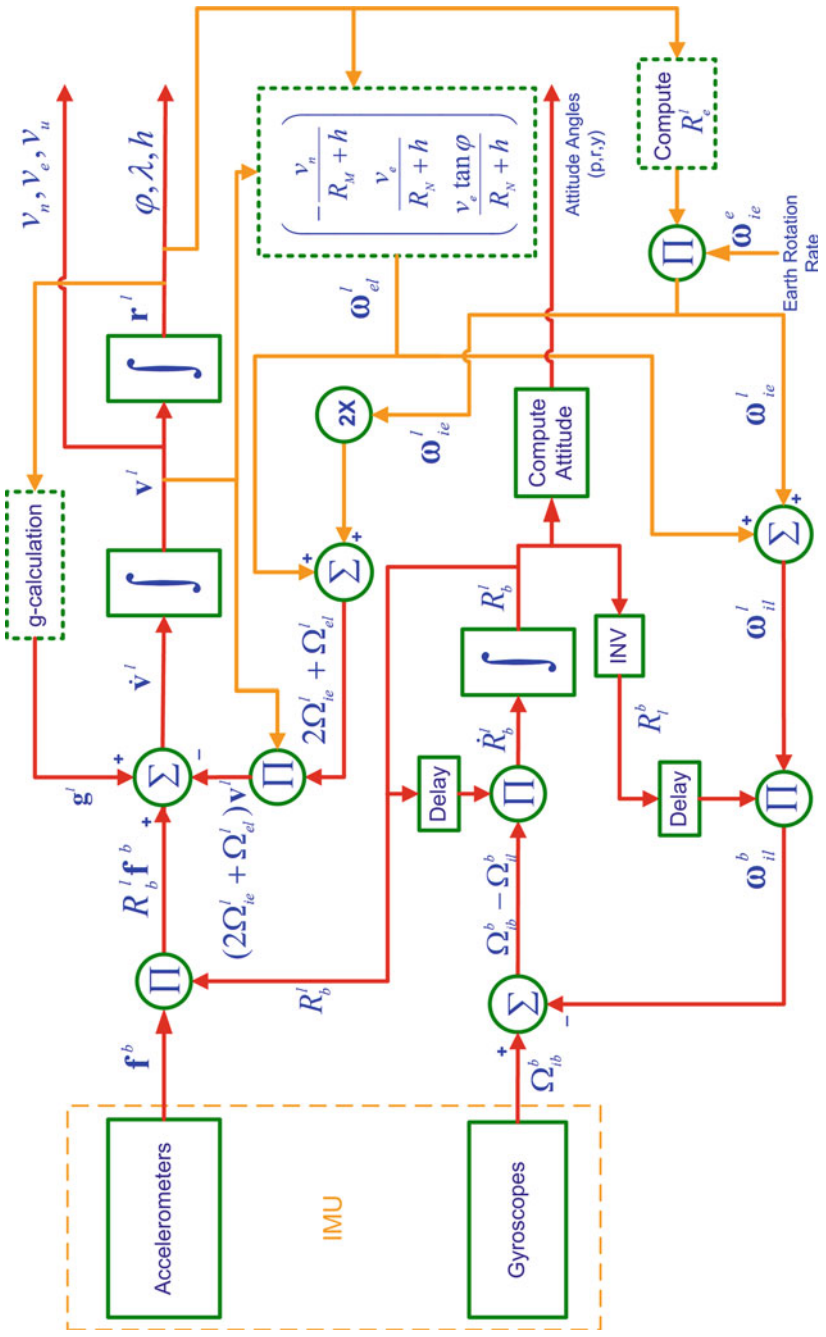


Fig. 5.9 A detailed diagram of mechanization of an INS in the local-level frame

It should be noted that the rotation rate of the l-frame due to the Earth's rotation rate $\boldsymbol{\omega}_{ie}^{l}$ and movement on the curved surface of the Earth $\boldsymbol{\omega}_{el}^{l}$ must be compensated from the measured angular rate of the body $\boldsymbol{\omega}_{ib}^{b}$ prior to integration. Similarly, the Coriolis acceleration due to the Earth's rotation $\boldsymbol{\Omega}_{ie}^{l}$ and movement of the l-frame over Earth's curvature $\boldsymbol{\Omega}_{el}^{l}$ must be subtracted from the measured specific force $\mathbf{f}^{b}$.

The mechanization algorithm provides the position, velocity and attitude components of the moving platform in the following format

- The position in geodetic (curvilinear) coordinates (φ, λ, h)
- The velocities along the east, north and up directions (v_e, v_n, v_u)
- The attitude angles as roll, pitch and yaw.

Mechanization in the l-frame is more intuitive for navigation on or near the Earth's surface because the position of the moving platform is provided in familiar map coordinates (latitude, longitude and altitude) and its attitude is given as angles in the familiar roll, pitch and yaw scheme. Also, the gravity model for the l-frame is simpler. However, additional computations are required to remove the effect of changes in the orientation of the l-frame which are not caused by the movement of the body itself.

5.5.1 Raw Measurement Data

The output of inertial sensors can sometimes (and especially for low cost sensors) be the angular rates and specific forces rather than incremental values. Because we require incremental values for our algorithms the angular rates and specific forces must be changed to their incremental counterparts as follows

$$\begin{aligned} \Delta\tilde{v}^{b} &= \tilde{f}^{b}\Delta t \\ \Delta\tilde{\theta}_{ib}^{b} &= \tilde{\omega}_{ib}^{b}\Delta t \end{aligned} \tag{5.103}$$

where

$\tilde{f}^{b}$ is the specific force (i.e. the output of the accelerometer) (in m/sec^2)

$\tilde{\omega}_{ib}^{b}$ is the rotation rate of the body frame with respect to the inertial frame, resolved in the body frame (i.e. the output of the gyroscope) (in radian/sec)

$\Delta\tilde{v}^{b}$ is the change in specific force during the time interval t (i.e. the velocity) (in m/s)

$\Delta\tilde{\theta}_{ib}^{b}$ is the change in angular rate during the time interval t (i.e. an angle) (in radians)

Δt is the sampling interval (i.e. the reciprocal of the sampling frequency) (in sec)

5.5.2 Correction of the Measurement Data

Although inertial sensors are calibrated in the factory they are usually recalibrated in the laboratory, as explained in Sect. 4.13. As a result of this calibration the biases and scale factors of the sensors are computed and subsequently compensated for in the raw measurements in order to obtain corrected measurements using the following relationship

$$\begin{aligned} \Delta\theta_{ib}^{b} &= \frac{\Delta\tilde{\theta}_{ib}^{b} - b_{gyro}\Delta t}{1 + s_{gyro}} \\ \Delta v^{b} &= \frac{\Delta\tilde{v}^{b} - b_{acc}\Delta t}{1 + s_{acc}} \end{aligned} \tag{5.104}$$

where

b_{gyro} is the drift of the gyroscope (in radians/sec)
s_{gyro} is the gyroscope scale factor (in ppm)
b_{acc} is the bias of the accelerometer (in m/sec^2)
s_{acc} is the accelerometer scale factor (in ppm)
$\Delta\theta_{ib}^{b}$ is the corrected incremental gyroscope output (in radians/sec)
Δv^{b} is the corrected incremental accelerometer output (in m/s)

5.5.3 Calculation and Updating of Rotation Matrix

In l-frame mechanization the updated rotation matrix R_b^l transforms the sensor outputs from the body frame to the l-frame, which requires a determination of the angular increment $\boldsymbol{\theta}_{lb}^{b}$ of the body with respect to the l-frame.

5.5.3.1 Computation of Angular Increment of Body Rotation

The angular rate of the body with respect to the l-frame $\boldsymbol{\omega}_{lb}^{b}$ is

$$\boldsymbol{\omega}_{lb}^{b} = \boldsymbol{\omega}_{ib}^{b} - \boldsymbol{\omega}_{il}^{b} \tag{5.105}$$

where

$\boldsymbol{\omega}_{ib}^{b}$ is the rotation rate of the body frame with respect to the inertial frame, resolved in the body frame (i.e. the output of the gyroscope) (in radian/sec)
$\boldsymbol{\omega}_{il}^{b}$ is the rotation rate of the l-frame with respect to the inertial frame, resolved in the body frame

Furthermore, $\boldsymbol{\omega}_{il}^{b}$ can be computed as

$$\boldsymbol{\omega}_{il}^{b} = R_{l}^{b}\boldsymbol{\omega}_{il}^{l} \tag{5.106}$$

where

$\boldsymbol{\omega}_{il}^{l}$ is the rotation rate of the l-frame with respect to the inertial frame, resolved in the l-frame

R_{l}^{b} is the rotation matrix from the l-frame to the body frame

The rotation rate of the l-frame $\boldsymbol{\omega}_{il}^{l}$ can further be written as

$$\boldsymbol{\omega}_{il}^{l} = \boldsymbol{\omega}_{ie}^{l} + \boldsymbol{\omega}_{el}^{l} \tag{5.107}$$

where $\boldsymbol{\omega}_{ie}^{l}$ is obtained as

$$\boldsymbol{\omega}_{ie}^{l} = R_{e}^{l}\boldsymbol{\omega}_{ie}^{e} \tag{5.108}$$

Substituting for R_{e}^{l} and ω_{ie}^{e} gives

$$\boldsymbol{\omega}_{ie}^{l} = \begin{bmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\sin\varphi\cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi \\ \cos\varphi\cos\lambda & \cos\varphi\sin\lambda & \sin\varphi \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \omega_{e} \end{bmatrix} \tag{5.109}$$

$$\boldsymbol{\omega}_{ie}^{l} = \begin{bmatrix} 0 \\ \omega_{e}\cos\varphi \\ \omega_{e}\sin\varphi \end{bmatrix} \tag{5.110}$$

and the transport rate $\boldsymbol{\omega}_{el}^{l}$ is

$$\boldsymbol{\omega}_{el}^{l} = \begin{bmatrix} -\dot{\varphi} \\ \dot{\lambda}\cos\varphi \\ \dot{\lambda}\sin\varphi \end{bmatrix} = \begin{bmatrix} -\dfrac{v_{n}}{M+h} \\ \dfrac{v_{e}}{N+h} \\ \dfrac{v_{e}\tan\varphi}{N+h} \end{bmatrix} \tag{5.111}$$

After finding the values of the appropriate terms through the above procedure and substituting them into Eq. (5.105), the equation is integrated for interval Δt to obtain the angular increment of the body rotation with respect to the l-frame

$$\boldsymbol{\theta}_{lb}^{b} = \boldsymbol{\theta}_{ib}^{b} - \boldsymbol{\theta}_{il}^{b} \tag{5.112}$$

which can be expanded into component form as

$$\begin{bmatrix} \boldsymbol{\theta}_{lb,x}^{b} \\ \boldsymbol{\theta}_{lb,y}^{b} \\ \boldsymbol{\theta}_{lb,z}^{b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\theta}_{ib,x}^{b} \\ \boldsymbol{\theta}_{ib,y}^{b} \\ \boldsymbol{\theta}_{ib,z}^{b} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\theta}_{il,x}^{b} \\ \boldsymbol{\theta}_{il,y}^{b} \\ \boldsymbol{\theta}_{il,z}^{b} \end{bmatrix} \tag{5.113}$$

5.5.3.2 Updating the Quaternion

The quaternion can be updated by any of the methods described in Sect. 5.4.3, but we will illustrate the analytical method here.

The closed form solution of the quaternion equation is

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \frac{1}{2}\left[2\left(\cos\frac{\boldsymbol{\theta}}{2} - 1\right)I + \frac{2}{\boldsymbol{\theta}}\sin\frac{\boldsymbol{\theta}}{2}\bar{S}(\omega)\right]\mathbf{q}_k \tag{5.114}$$

and the expanded component form is

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}_{k+1} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}_k + \frac{1}{2}\begin{bmatrix} \breve{c} & \breve{s}\omega^l_{lb,z}\Delta t & -\breve{s}\omega^l_{lb,y}\Delta t & \breve{s}\omega^l_{lb,z}\Delta t \\ -\breve{s}\omega^l_{lb,z}\Delta t & \breve{c} & \breve{s}\omega^l_{lb,x}\Delta t & \breve{s}\omega^l_{lb,y}\Delta t \\ \breve{s}\omega^l_{lb,y}\Delta t & -\breve{s}\omega^l_{lb,x}\Delta t & \breve{c} & \breve{s}\omega^l_{lb,z}\Delta t \\ -\breve{s}\omega^l_{lb,x}\Delta t & -\breve{s}\omega^l_{lb,y}\Delta t & -\breve{s}\omega^l_{lb,z}\Delta t & \breve{c} \end{bmatrix}_k \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}_k \tag{5.115}$$

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}_{k+1} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}_k + \frac{1}{2}\begin{bmatrix} \breve{c} & \breve{s}\boldsymbol{\theta}^b_{lb,z} & -\breve{s}\boldsymbol{\theta}^b_{lb,y} & \breve{s}\boldsymbol{\theta}^b_{lb,z} \\ -\breve{s}\boldsymbol{\theta}^b_{lb,z} & \breve{c} & \breve{s}\boldsymbol{\theta}^b_{lb,x} & \breve{s}\boldsymbol{\theta}^b_{lb,y} \\ \breve{s}\boldsymbol{\theta}^b_{lb,y} & -\breve{s}\boldsymbol{\theta}^b_{lb,x} & \breve{c} & \breve{s}\boldsymbol{\theta}^b_{lb,z} \\ -\breve{s}\boldsymbol{\theta}^b_{lb,x} & -\breve{s}\boldsymbol{\theta}^b_{lb,y} & -\breve{s}\boldsymbol{\theta}^b_{lb,z} & \breve{c} \end{bmatrix}_k \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}_k \tag{5.116}$$

The terms s and c are defined as

$$\breve{s} = \frac{2}{\theta}\sin\frac{\theta}{2} \tag{5.117}$$

$$\breve{c} = 2\left(\cos\frac{\theta}{2} - 1\right) \tag{5.118}$$

The magnitude of the incremental angle of the body rotation about a fixed axis is therefore

$$\boldsymbol{\theta} = \sqrt{\left(\boldsymbol{\theta}^b_{lb,x}\right)^2 + \left(\boldsymbol{\theta}^b_{lb,y}\right)^2 + \left(\boldsymbol{\theta}^b_{lb,z}\right)^2} \tag{5.119}$$

5.5.4 Attitude Computation

Once the quaternion has been computed at epoch $k + 1$, the transformation matrix R^l_b can be obtained using the following relationship

$$R_b^l = \begin{bmatrix} R(1,1) & R(1,2) & R(1,3) \\ R(2,1) & R(2,2) & R(2,3) \\ R(3,1) & R(3,2) & R(3,3) \end{bmatrix}$$
$$= \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 - q_3q_4) & 2(q_1q_3 + q_2q_4) \\ 2(q_1q_2 + q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 - q_1q_4) \\ 2(q_1q_3 - q_2q_4) & 2(q_2q_3 + q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \tag{5.120}$$

where (in accordance with Chap. 2) the actual elements of the matrix R_b^l are

$$R_b^l = \begin{bmatrix} \cos y \cos r - \sin y \sin p \sin r & -\sin y \cos p & \cos y \sin r + \sin y \sin p \cos r \\ \sin y \cos r + \cos y \sin p \sin r & \cos y \cos p & \sin y \sin r - \cos y \sin p \cos r \\ -\cos p \sin r & \sin p & \cos p \cos r \end{bmatrix} \tag{5.121}$$

Once the updated transformation matrix R_b^l is known, the attitude angles can be computed as

$$pitch = \sin^{-1}\left[R_b^l(3,2)\right] \tag{5.122}$$

$$yaw = -\tan^{-1}\left[\frac{R_b^l(1,2)}{R_b^l(2,2)}\right] \tag{5.123}$$

$$roll = -\tan^{-1}\left[\frac{R_b^l(3,1)}{R_b^l(3,3)}\right] \tag{5.124}$$

5.5.5 Velocity Computation

The mechanization equation for the rate of change of velocity in the l-frame is

$$\begin{aligned} \dot{\mathbf{v}}^l &= R_b^l \mathbf{f}^b - (2\Omega_{ie}^l + \Omega_{el}^l)\mathbf{v}^l + \mathbf{g}^l \\ \frac{\Delta \mathbf{v}^l}{\Delta t} &= R_b^l \mathbf{f}^b - (2\Omega_{ie}^l + \Omega_{el}^l)\mathbf{v}^l + \mathbf{g}^l \\ \Delta \mathbf{v}^l &= R_b^l \mathbf{f}^b \Delta t - (2\Omega_{ie}^l + \Omega_{el}^l)\mathbf{v}^l \Delta t + \mathbf{g}^l \Delta t \end{aligned} \tag{5.125}$$

$$\Delta \mathbf{v}^l = \Delta \tilde{\mathbf{v}}^l - (2\Omega_{ie}^l + \Omega_{el}^l)\mathbf{v}^l \Delta t + \mathbf{g}^l \Delta t \tag{5.126}$$

where

$\Delta\tilde{\mathbf{v}}^l$	is the measured velocity increment after transformation to the l-frame
$(2\Omega_{ie}^l + \Omega_{el}^l)\mathbf{v}^l\Delta t$	is the Coriolis correction that compensates for the Earth's rotation and the resulting change of orientation of the l-frame
$\mathbf{g}^l\Delta t$	is the gravity correction

Because the z-axis of the l-frame is defined to be normal to the ellipsoid, the normal gravity vector is zero in the other axes

$$\mathbf{g}^l = [\,0 \quad 0 \quad -g\,]^T \tag{5.127}$$

The gravity component g is calculated (in accordance with Chap. 2) as

$$g = a_1\left(1 + a_2\sin^2\varphi + a_3\sin^4\varphi\right) + \left(a_4 + a_5\sin^2\varphi\right)h + a_6h^2 \tag{5.128}$$

Now the velocity at the current epoch can be calculated by the modified Euler formula

$$\mathbf{v}_{k+1}^l = \mathbf{v}_k^l + .5\left(\Delta\mathbf{v}_k^l + \Delta\mathbf{v}_{k+1}^l\right) \tag{5.129}$$

where $\mathbf{v}^l = [\,v_e \quad v_n \quad v_u\,]^T$.

5.5.6 Position Computation

The position coordinates for the l-frame (i.e. latitude, longitude and altitude) are calculated as follows.

The rate of change of latitude is

$$\dot{\varphi} = \frac{v_n}{R_N + h} \tag{5.130}$$

therefore

$$\varphi_{k+1} = \varphi_k + \frac{1}{2}\frac{\left[v_{n,k} + v_{n,k+1}\right]}{R_N + h}\Delta t \tag{5.131}$$

The rate of change of longitude is

$$\dot{\lambda} = \frac{v_e}{(R_N + h)\cos\varphi} \tag{5.132}$$

therefore

$$\lambda_{k+1} = \lambda_k + \frac{1}{2}\frac{\left(v_{e,k} + v_{e,k+1}\right)}{(R_N + h)\cos\varphi}\Delta t \tag{5.133}$$

The rate of change of altitude is $\dot{h} = v_u$, therefore

$$h_{k+1} = h_k + \frac{1}{2}\left(v_{u,k} + v_{u,k+1}\right)\Delta t \tag{5.134}$$

Appendix A

Solving the equation of the form $\dot{y} = yx$

$$\begin{aligned}
&\dot{y} = yx \\
&\frac{dy}{dt} = yx \\
&\frac{dy}{y} = xdt \\
&\ln(y)|_{y_k}^{y_{k+1}} = \int xdt \\
&\ln(y_{k+1}) - \ln(y_k) = \int xdt \\
&\ln\left(\frac{y_{k+1}}{y_k}\right) = \int xdt \\
&\frac{y_{k+1}}{y_k} = e^{\int xdt} \\
&y_{k+1} = y_k e^{\int xdt}
\end{aligned}$$

References

Bekir E (2007) Introduction to modern navigation systems. World Scientific Singapore, Hackensack, NJ

Farrell JA (2008) Aided Navigation: GPS with High Rate Sensors. McGraw-Hill Professional, USA

Georgy J, Karamat TB, Iqbal U, Noureldin A (2010) Enhanced MEMS-IMU/odometer/GPS integration using mixture particle filter. GPS Solutions 15(3):1–14. doi:10.1007/s10291-010-0186-4

Iqbal U, Karamat TB, Okou AF, Noureldin A (2009) Experimental results on an integrated GPS and multisensor system for land vehicle positioning. International journal of navigation and observation, hindawi publishing corporation, Article ID 765010 2009: 18. doi:10.1155/2009/765010

Jekeli C (2001) Inertial navigation systems with geodetic applications. Walter de Gruyter, Berlin, New York

Kuipers JB (1999) Quaternions and rotation sequences : a primer with applications to orbits, aerospace, and virtual reality. Princeton University Press, Princeton NJ

Noureldin A, Karamat TB, Eberts MD, El-Shafie A (2009) Performance enhancement of MEMS-based INS/GPS integration for low-cost navigation applications. IEEE Trans Veh Technol 58(3):1077–1096

Rogers RM (2007) Applied mathematics in integrated navigation systems, 3rd edn. AIAA

Chapter 6
Modeling INS Errors by Linear State Equations

The accuracy of an INS is affected by various sources. These include errors during the initial alignment procedure, sensor errors, and the limitations of the processing algorithm. To see the effect of these errors on the navigational output parameters (position, velocity and attitude) it is vital to understand their propagation through the navigation equations. Once the nature of the errors is known, one can mitigate them by proper modeling and estimation techniques. This usually requires external aiding sources in order to limit the errors and predict their behavior. Hence error models are required for the analysis and estimation of the error sources associated with any inertial navigation system. The estimator options include Kalman filters (KF) (Karamat et al. 2009; Noureldin et al. 2009), particle filters (PF) (Georgy et al. 2008, 2010, 2011) and artificial intelligence (AI) techniques (Noureldin et al. 2007; Chiang et al. 2006). Traditionally various forms of KF are used for sensor fusion and therefore we will restrict ourselves to this technique.

Mechanization equations describe physical process by a deterministic dynamic system. Navigation parameters can be determined using these state equations by taking kinematic measurements along a nominal trajectory. Since the solution of these equations incorporate errors (both deterministic and stochastic), sensor error models are required for analysis and estimation purposes. The errors of dynamic systems are variable in time, and are therefore described by differential equations. And since these equations are non-linear, we must linearize them prior to applying Kalman filtering.

In general a dynamic non-linear system can be expressed by a set of first-order differential equations as follows

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}, t) \tag{6.1}$$

where $\mathbf{x}$ is a vector of *internal parameters* of the dynamic system that is referred to as the *state vector*. The exact values of $\mathbf{x}$ are usually not known, and only an approximate value can be computed based upon the nominal trajectory and noisy

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_6,

measurements. For inertial systems, the output of mechanization can be thought of as a *nominal* trajectory.

After linearizing Eq. (6.1) about a nominal solution (see Chap. 2 for the detailed derivation of the linearization) and retaining only the first-order terms

$$\delta\dot{\mathbf{x}}(t) = \frac{\partial f[\mathbf{x}(t)]}{\partial \mathbf{x}}\delta\mathbf{x}(t) \tag{6.2}$$

By letting

$$F(t) = \frac{\partial f[\mathbf{x}(t)]}{\partial \mathbf{x}}$$

and substituting F into Eq. (6.2) we obtain linearized error state equation of the non-linear system

$$\delta\dot{\mathbf{x}}(t) = F_x(t)\delta\mathbf{x}(t) \tag{6.3}$$

For the inertial navigation system, the state vector $\delta\mathbf{x}$ consists of the position errors, velocity errors, attitude errors and the errors in the inertial sensors. The matrix F is called the *dynamic matrix* which propagates the errors over time.

Since inertial sensors (accelerometers and gyroscopes) contain both correlated and white noise components we can rewrite Eq. (6.3) as

$$\delta\dot{\mathbf{x}}(t) = F_x(t)\delta\mathbf{x}(t) + G\mathbf{w}(t) \tag{6.4}$$

where

$\mathbf{w}(t)$ are the random forcing functions, assumed to be Gaussian white noise associate with the inertial sensors

G is the noise distribution matrix.

6.1 Local-Level Frame Error State Equations

The navigation equations can be mechanized in the i-frame, e-frame or l-frame as detailed in Chap. 5. That said, mechanization equations are usually sought in the l-frame because of the advantages mentioned in that chapter. The derivation of the error state equations in the l-frame is more complex. However, once understood, the equations for the other frames can easily be derived based on the knowledge of l-frame error state equations. As this will be briefly illustrated later in this chapter, here we will focus on the l-frame error state equations so that the reader will gain sufficient understanding of the procedure to be able to implement the algorithm.

As explained in Chap. 5, the mechanization equations in the l-frame can be expressed in a compact form as

$$\begin{bmatrix} \dot{\mathbf{r}}^l \\ \dot{\mathbf{v}}^l \\ \dot{R}_b^l \end{bmatrix} = \begin{bmatrix} D^{-1}\mathbf{v}^l \\ R_b^l \mathbf{f}^b - (2\Omega_{ie}^l + \Omega_{el}^l)\mathbf{v}^l + \mathbf{g}^l \\ R_b^l(\Omega_{ib}^b - \Omega_{il}^b) \end{bmatrix} \tag{6.5}$$

The error state vector for the mechanization equations in the local-level frame consists of the errors along the curvilinear geodetic coordinates (latitude error $\delta\varphi$, longitude error $\delta\lambda$, and altitude error δh), the errors along the Earth-referenced velocities (east-velocity error δv_e, north-velocity error δv_n, and up-velocity error δv_u) and the error along the three attitude angles (pitch error δp, roll error δr, and azimuth error δA). It also includes the accelerometer biases and gyro drifts. The complete vector of error states is therefore

$$\mathbf{x}_{15\times1}^l = \left[\delta\mathbf{r}_{3\times1}^l, \delta\mathbf{v}_{3\times1}^l, \boldsymbol{\varepsilon}_{3\times1}^l, \delta\boldsymbol{\omega}_{3\times1}, \delta\mathbf{f}_{3\times1}\right]^T \tag{6.6}$$

where

$\delta\mathbf{r}^l = [\delta\varphi, \delta\lambda, \delta h]^T$ is the position error vector
$\delta\mathbf{v}^l = [\delta v_e, \delta v_n, \delta v_u]^T$ is the Earth-referenced velocity error vector
$\boldsymbol{\varepsilon}^l = [\delta p, \delta r, \delta A]^T$ is the attitude error vector
$\delta\boldsymbol{\omega} = \left[\delta\omega_x, \delta\omega_y, \delta\omega_z\right]^T$ is gyroscope error vector (consisting of drifts)
$\delta\mathbf{f} = \left[\delta f_x, \delta f_y, \delta f_z\right]^T$ is accelerometer error vector (consisting of biases).

6.1.1 Position Errors for Local-Level Frame

From Eq. (6.5) the relationship between the time rate of change of coordinate vector $\mathbf{r}^l$ and the Earth-referenced velocity vector $\mathbf{v}^l$ is

$$\dot{\mathbf{r}}^l = D^{-1}\mathbf{v}^l \tag{6.7}$$

where

$$\dot{\mathbf{r}}^l = (\dot{\varphi}, \dot{\lambda}, \dot{h})$$

$$\mathbf{v}^l = (v_e, v_n, v_u)$$

$$D^{-1} = \begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Substituting these terms into Eq. (6.7) gives

$$\begin{bmatrix} \dot{\varphi} \\ \dot{\lambda} \\ \dot{h} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_e \\ v_n \\ v_u \end{bmatrix} \tag{6.8}$$

The error in the coordinate vector is the difference between the true coordinate vector $\mathbf{r}^l$ and the one computed during the INS mechanization $\hat{\mathbf{r}}^l$. The time rate of change of this error $\delta\dot{\mathbf{r}}^l$ is

$$\delta\dot{\mathbf{r}}^l = \dot{\hat{\mathbf{r}}}^l - \dot{\mathbf{r}}^l \tag{6.9}$$

As discussed earlier, by employing the Taylor series expansion to a first-order approximation it is possible to express the time rate of change of the coordinate error $\delta\dot{\mathbf{r}}^l$ as

$$\begin{aligned} \delta\dot{\mathbf{r}}^l &= \frac{\partial}{\partial \mathbf{r}^l}\left(\dot{\mathbf{r}}^l\right)\delta\mathbf{r}^l \\ \delta\dot{\mathbf{r}}^l &= \frac{\partial}{\partial \mathbf{r}^l}\left\{ \begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_e \\ v_n \\ v_u \end{bmatrix} \right\}\delta\mathbf{r}^l \end{aligned} \tag{6.10}$$

$$\begin{aligned} \delta\dot{\mathbf{r}}^l = &\begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} + \\ &\begin{bmatrix} 0 & \frac{-\delta h}{(R_M+h)^2} & 0 \\ \frac{\tan\varphi\delta\varphi}{(R_N+h)\cos\varphi} - \frac{\delta h}{(R_N+h)^2\cos\varphi} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_e \\ v_n \\ v_u \end{bmatrix} \end{aligned} \tag{6.11}$$

The next two steps will manipulate the second term in this equation.

$$\begin{aligned} \delta\dot{\mathbf{r}}^l = &\begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} + \\ &\begin{bmatrix} \frac{-v_n\delta h}{(R_M+h)^2} \\ \left\{\frac{\tan\varphi\delta\varphi}{(R_N+h)\cos\varphi} - \frac{\delta h}{(R_N+h)^2\cos\varphi}\right\}v_e \\ 0 \end{bmatrix} \end{aligned} \tag{6.12}$$

The second term on the right-hand side of this equation can be rewritten as the product of a matrix and a vector of position errors

$$\delta\dot{\mathbf{r}}^l = \begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} + \begin{bmatrix} 0 & 0 & \frac{-v_n}{(R_M+h)^2} \\ \frac{v_e \tan\varphi}{(R_N+h)\cos\varphi} & 0 & \frac{-v_e}{(R_N+h)^2\cos\varphi} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} \tag{6.13}$$

The reader should recall that $\dot{\lambda} = \frac{v_e}{(R_N+h)\cos\varphi}$ and $\dot{\varphi} = \frac{v_n}{R_M+h}$. The above equation and similar ones that will appear later in this chapter may use either $\dot{\lambda}$S and $\dot{\varphi}$ or their right-hand-side equivalents. It should also be noted that the terms which involve the reciprocal of the square of the Earth's radius (R_M or R_N) are very small and can be neglected. These include $\frac{v_n}{(R_M+h)^2}$ and $\frac{v_e}{(R_N+h)^2\cos\varphi}$. Furthermore, the second term of the above equation involves multiplication of $\frac{v_e \tan\varphi}{(R_N+h)\cos\varphi}$ (or $\dot{\lambda}\tan\varphi$) with the latitude error $\delta\varphi$. This will lead to a relatively small error quantity that can also be neglected. Therefore the time rate of change of the coordinate errors can be reduced to

$$\delta\dot{\mathbf{r}}^l = \begin{bmatrix} \delta\dot{\varphi} \\ \delta\dot{\lambda} \\ \delta\dot{h} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} \tag{6.14}$$

which can be written more compactly as

$$\delta\dot{\mathbf{r}}^l = D^{-1}\delta\mathbf{v}^l \tag{6.15}$$

6.1.2 Velocity Errors for Local-Level Frame

From Chap. 5 we know that vehicle accelerations are obtained from the specific force measurements $\tilde{f}^b$ after compensating for both the Coriolis acceleration and the effect of gravity.

$$\dot{\hat{\mathbf{v}}}^l = \hat{R}_b^l\tilde{\mathbf{f}}^b - (2\hat{\Omega}_{ie}^l + \hat{\Omega}_{el}^l)\hat{\mathbf{v}}^l + \hat{\mathbf{g}}^l \tag{6.16}$$

Taking into account the errors in the measurements

$$\dot{\hat{\mathbf{v}}}^l = (I+\Psi)R_b^l(\mathbf{f}^b + \delta\mathbf{f}^b) - \left[2(\Omega_{ie}^l + \delta\Omega_{ie}^l) + \Omega_{el}^l + \delta\Omega_{el}^l)\right](\mathbf{v}^l + \delta\mathbf{v}^l) + (\mathbf{g}^l + \delta\mathbf{g}^l) \tag{6.17}$$

where $\dot{\hat{\mathbf{v}}}^l = \dot{\mathbf{v}}^l + \delta\dot{\mathbf{v}}^l$.

After mathematical manipulation of this equation and discarding the second-order terms, the error on the velocity states becomes

$$\delta\dot{\mathbf{v}}^l = \Psi R_b^l \mathbf{f}^b + R_b^l \delta\mathbf{f}^b - (2\Omega_{ie}^l + \Omega_{el}^l)\delta\mathbf{v}^l - (2\delta\Omega_{ie}^l + \delta\Omega_{el}^l)\mathbf{v}^l + \delta\mathbf{g}^l \tag{6.18}$$

Recalling that $\mathbf{f}^l = R_b^l \mathbf{f}^b$, and applying the skew-symmetric matrix property $A\mathbf{b} = -B\mathbf{a}$, we finally get

$$\delta\dot{\mathbf{v}}^l = -\mathbf{F}^l\boldsymbol{\varepsilon}^l + R_b^l \delta\mathbf{f}^b - 2(\Omega_{ie}^l + \Omega_{el}^l)\delta\mathbf{v}^l + \mathbf{V}^l(2\delta\boldsymbol{\omega}_{ie}^l + \delta\boldsymbol{\omega}_{el}^l) + \delta\mathbf{g}^l \tag{6.19}$$

where

$\mathbf{F}^l$	is the skew-symmetric matrix of the corresponding specific force vector $\mathbf{f}^l$
$\mathbf{V}^l$	is the skew-symmetric matrix of the corresponding velocity vector $\mathbf{v}^l$
$\delta\mathbf{g}^l$	is the error in the normal gravity vector
$\delta\boldsymbol{\omega}_{ie}^l$	is the error in the rotation rate of the Earth, resolved in the l-frame
$\delta\boldsymbol{\omega}_{el}^l$	is vector of the error in the angular velocity vector $\boldsymbol{\omega}_{el}^l$.

Now we will explore Eq. (6.19), which describes the velocity errors, and expand the components in a matrix form to get a sense of what these terms really are. The terms are labeled as follows

$$\delta\dot{\mathbf{v}}^l = \underbrace{-\mathbf{F}^l\boldsymbol{\varepsilon}^l}_{1} + \underbrace{R_b^l \delta\mathbf{f}^b}_{2} - \underbrace{2(\Omega_{ie}^l + \Omega_{el}^l)\delta\mathbf{v}^l}_{3} + \underbrace{\mathbf{V}^l(2\delta\boldsymbol{\omega}_{ie}^l + \delta\boldsymbol{\omega}_{el}^l)}_{4} + \underbrace{\delta\mathbf{g}^l}_{5} \tag{6.20}$$

We will start by expanding the first term of Eq. (6.20)

$$-\mathbf{F}^l\boldsymbol{\varepsilon}^l = \begin{bmatrix} 0 & f_u & -f_n \\ -f_u & 0 & f_e \\ f_n & -f_e & 0 \end{bmatrix} \begin{bmatrix} \delta p \\ \delta r \\ \delta A \end{bmatrix} \tag{6.21}$$

where f_e, f_n and f_u are the body accelerations along the east, north and up directions.

In the second term of Eq. (6.20) the accelerometer biases $\delta\mathbf{f}^b$ are transformed from the body frame to the local-level frame using the R_b^l matrix

$$R_b^l \delta\mathbf{f}^b = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} \delta f_x \\ \delta f_y \\ \delta f_z \end{bmatrix} \tag{6.22}$$

For the third term of Eq. (6.20) we use the following definitions (from Chap. 4)

$$\boldsymbol{\omega}_{el}^l = \begin{bmatrix} -\dot{\varphi} \\ \dot{\lambda}\cos\varphi \\ \dot{\lambda}\sin\varphi \end{bmatrix}, \boldsymbol{\omega}_{ie}^l = \begin{bmatrix} 0 \\ \omega_e\cos\varphi \\ \omega_e\sin\varphi \end{bmatrix}, \delta\mathbf{v}^l = \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} \tag{6.23}$$

to obtain

$$-2(\boldsymbol{\Omega}_{ie}^{l}+\boldsymbol{\Omega}_{el}^{l})\delta\mathbf{v}^{l}=-2\left\{\begin{bmatrix}0&-\omega_e\sin\varphi&\omega_e\cos\varphi\\ \omega_e\sin\varphi&0&0\\ -\omega_e\cos\varphi&0&0\end{bmatrix}+\begin{bmatrix}0&-\dot{\lambda}\sin\varphi&\dot{\lambda}\cos\varphi\\ \dot{\lambda}\sin\varphi&0&\dot{\varphi}\\ -\dot{\lambda}\cos\varphi&-\dot{\varphi}&0\end{bmatrix}\right\}\begin{bmatrix}\delta v_e\\ \delta v_n\\ \delta v_u\end{bmatrix} \tag{6.24}$$

$$-2(\boldsymbol{\Omega}_{ie}^{l}+\boldsymbol{\Omega}_{el}^{l})\delta\mathbf{v}^{l}=\begin{bmatrix}0&(2\omega_e+\dot{\lambda})\sin\varphi&-(2\omega_e+\dot{\lambda})\cos\varphi\\ -(2\omega_e+\dot{\lambda})\sin\varphi&0&-\dot{\varphi}\\ (2\omega_e+\dot{\lambda})\cos\varphi&\dot{\varphi}&0\end{bmatrix}\begin{bmatrix}\delta v_e\\ \delta v_n\\ \delta v_u\end{bmatrix} \tag{6.25}$$

Now consider the fourth term of Eq. (6.20). Given the definitions of $\boldsymbol{\omega}_{ie}^{l}$ and $\boldsymbol{\omega}_{el}^{l}$ in Eq. (6.23) we can obtain $\delta\boldsymbol{\omega}_{ie}^{l}$ and $\delta\boldsymbol{\omega}_{el}^{l}$ as follows

$$\delta\boldsymbol{\omega}_{ie}^{l}=\begin{bmatrix}0\\ -\omega_e\sin\varphi\delta\varphi\\ \omega_e\cos\varphi\delta\varphi\end{bmatrix},\ \delta\boldsymbol{\omega}_{el}^{l}=\begin{bmatrix}-\delta\dot{\varphi}\\ -\dot{\lambda}\sin\varphi\delta\varphi+\cos\varphi\delta\dot{\lambda}\\ \dot{\lambda}\cos\varphi\delta\varphi+\sin\varphi\delta\dot{\lambda}\end{bmatrix} \tag{6.26}$$

Therefore

$$\begin{aligned}(2\delta\boldsymbol{\omega}_{ie}^{l}+\delta\boldsymbol{\omega}_{el}^{l})&=\left\{\begin{bmatrix}0\\ -2\omega_e\sin\varphi\delta\varphi\\ 2\omega_e\cos\varphi\delta\varphi\end{bmatrix}+\begin{bmatrix}-\delta\dot{\varphi}\\ -\dot{\lambda}\sin\varphi\delta\varphi+\cos\varphi\delta\dot{\lambda}\\ \dot{\lambda}\cos\varphi\delta\varphi+\sin\varphi\delta\dot{\lambda}\end{bmatrix}\right\}\\ (2\delta\boldsymbol{\omega}_{ie}^{l}+\delta\boldsymbol{\omega}_{el}^{l})&=\begin{bmatrix}-\delta\dot{\varphi}\\ -2\omega_e\sin\varphi\delta\varphi-\dot{\lambda}\sin\varphi\delta\varphi+\cos\varphi\delta\dot{\lambda}\\ 2\omega_e\cos\varphi\delta\varphi+\dot{\lambda}\cos\varphi\delta\varphi+\sin\varphi\delta\dot{\lambda}\end{bmatrix}\\ (2\delta\boldsymbol{\omega}_{ie}^{l}+\delta\boldsymbol{\omega}_{el}^{l})&=\begin{bmatrix}-\delta\dot{\varphi}\\ -(2\omega_e+\dot{\lambda})\sin\varphi\delta\varphi+\cos\varphi\delta\dot{\lambda}\\ (2\omega_e+\dot{\lambda})\cos\varphi\delta\varphi+\sin\varphi\delta\dot{\lambda}\end{bmatrix}\\ (2\delta\boldsymbol{\omega}_{ie}^{l}+\delta\boldsymbol{\omega}_{el}^{l})&=\begin{bmatrix}0\\ -(2\omega_e+\dot{\lambda})\sin\varphi\delta\varphi\\ (2\omega_e+\dot{\lambda})\cos\varphi\delta\varphi\end{bmatrix}+\begin{bmatrix}-\delta\dot{\varphi}\\ \cos\varphi\delta\dot{\lambda}\\ \sin\varphi\delta\dot{\lambda}\end{bmatrix}\end{aligned} \tag{6.27}$$

$$\begin{aligned}(2\delta\boldsymbol{\omega}_{ie}^{l}+\delta\boldsymbol{\omega}_{el}^{l})=&\begin{bmatrix}0&0&0\\ -(2\omega_e+\dot{\lambda})\sin\varphi&0&0\\ (2\omega_e+\dot{\lambda})\cos\varphi&0&0\end{bmatrix}\begin{bmatrix}\delta\varphi\\ \delta\lambda\\ \delta h\end{bmatrix}\\ &+\begin{bmatrix}-1&0&0\\ 0&\cos\varphi&0\\ 0&\sin\varphi&0\end{bmatrix}\begin{bmatrix}\delta\dot{\varphi}\\ \delta\dot{\lambda}\\ \delta\dot{h}\end{bmatrix}\end{aligned} \tag{6.28}$$

Substituting for $[\delta\dot{\varphi}, \delta\dot{\lambda}, \delta\dot{h}]^T$ from Eq. (6.13) into (6.28) gives

$$(2\delta\boldsymbol{\omega}_{ie}^{l} + \delta\boldsymbol{\omega}_{el}^{l}) = \begin{bmatrix} 0 & 0 & 0 \\ -(2\omega_e + \dot{\lambda})\sin\varphi & 0 & 0 \\ (2\omega_e + \dot{\lambda})\cos\varphi & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} + \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & 0 \\ 0 & \sin\varphi & 0 \end{bmatrix} \left\{ \begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} + \begin{bmatrix} 0 & 0 & \frac{-\dot{\varphi}}{(R_M+h)} \\ \dot{\lambda}\tan\varphi & 0 & \frac{-\dot{\lambda}}{(R_N+h)} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} \right\} \tag{6.29}$$

$$(2\delta\boldsymbol{\omega}_{ie}^{l} + \delta\boldsymbol{\omega}_{el}^{l}) = \begin{bmatrix} 0 & 0 & 0 \\ -(2\omega_e + \dot{\lambda})\sin\varphi & 0 & 0 \\ (2\omega_e + \dot{\lambda})\cos\varphi & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{(R_M+h)} & 0 \\ \frac{1}{R_N+h} & 0 & 0 \\ \frac{\tan\varphi}{(R_N+h)} & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} + \begin{bmatrix} 0 & 0 & \frac{\dot{\varphi}}{(R_M+h)} \\ \dot{\lambda}\sin\varphi & 0 & \frac{-\dot{\lambda}\cos\varphi}{(R_N+h)} \\ \dot{\lambda}\sin\varphi\tan\varphi & 0 & \frac{-\dot{\lambda}\sin\varphi}{(R_N+h)} \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} \tag{6.30}$$

Adding the first and third terms on the right-hand side

$$(2\delta\boldsymbol{\omega}_{ie}^{l} + \delta\boldsymbol{\omega}_{el}^{l}) = \begin{bmatrix} 0 & 0 & \frac{\dot{\varphi}}{(R_M+h)} \\ -2\omega_e\sin\varphi & 0 & \frac{-\dot{\lambda}\cos\varphi}{(R_N+h)} \\ \left(\begin{array}{c}(2\omega_e + \dot{\lambda})\cos\varphi \\ +\dot{\lambda}\sin\varphi\tan\varphi\end{array}\right) & 0 & \frac{-\dot{\lambda}\sin\varphi}{(R_N+h)} \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} + \begin{bmatrix} 0 & -\frac{1}{(R_M+h)} & 0 \\ \frac{1}{R_N+h} & 0 & 0 \\ \frac{\tan\varphi}{(R_N+h)} & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} \tag{6.31}$$

$$(2\delta\boldsymbol{\omega}_{ie}^{l}+\delta\boldsymbol{\omega}_{el}^{l})=\begin{bmatrix}0 & 0 & \frac{\dot{\varphi}}{(R_M+h)}\\ -2\omega_e\sin\varphi & 0 & -\frac{\dot{\lambda}\cos\varphi}{(R_N+h)}\\ 2\omega_e\cos\varphi+\dot{\lambda}(\cos\varphi+\frac{\sin^2\varphi}{\cos\varphi}) & 0 & -\frac{\dot{\lambda}\sin\varphi}{(R_N+h)}\end{bmatrix}\begin{bmatrix}\delta\varphi\\ \delta\lambda\\ \delta h\end{bmatrix}+\begin{bmatrix}0 & \frac{-1}{(R_M+h)} & 0\\ \frac{1}{(R_N+h)} & 0 & 0\\ \frac{\tan\varphi}{(R_N+h)} & 0 & 0\end{bmatrix}\begin{bmatrix}\delta v_e\\ \delta v_n\\ \delta v_u\end{bmatrix} \tag{6.32}$$

Using the definitions $\dot{\varphi}=\frac{v_n}{(R_M+h)}$ and $\dot{\lambda}=\frac{v_e}{(R_N+h)\cos\varphi}$ gives

$$(2\delta\boldsymbol{\omega}_{ie}^{l}+\delta\boldsymbol{\omega}_{el}^{l})=\begin{bmatrix}0 & 0 & \frac{v_n}{(R_M+h)^2}\\ -2\omega_e\sin\varphi & 0 & -\frac{v_e}{(R_N+h)^2}\\ 2\omega_e\cos\varphi+\frac{\dot{\lambda}}{\cos\varphi} & 0 & -\frac{v_e\tan\varphi}{(R_N+h)^2}\end{bmatrix}\begin{bmatrix}\delta\varphi\\ \delta\lambda\\ \delta h\end{bmatrix}+\begin{bmatrix}0 & \frac{-1}{(R_M+h)} & 0\\ \frac{1}{(R_N+h)} & 0 & 0\\ \frac{\tan\varphi}{(R_N+h)} & 0 & 0\end{bmatrix}\begin{bmatrix}\delta v_e\\ \delta v_n\\ \delta v_u\end{bmatrix} \tag{6.33}$$

Ignoring the terms which are divided by $(R_M+h)^2$ or $(R_N+h)^2$ (as explained earlier) yields

$$(2\delta\boldsymbol{\omega}_{ie}^{l}+\delta\boldsymbol{\omega}_{el}^{l})=\begin{bmatrix}0 & 0 & 0\\ -2\omega_e\sin\varphi & 0 & 0\\ 2\omega_e\cos\varphi+\frac{\dot{\lambda}}{\cos\varphi} & 0 & 0\end{bmatrix}\begin{bmatrix}\delta\varphi\\ \delta\lambda\\ \delta h\end{bmatrix}+\begin{bmatrix}0 & \frac{-1}{(R_M+h)} & 0\\ \frac{1}{(R_N+h)} & 0 & 0\\ \frac{\tan\varphi}{(R_N+h)} & 0 & 0\end{bmatrix}\begin{bmatrix}\delta v_e\\ \delta v_n\\ \delta v_u\end{bmatrix} \tag{6.34}$$

The skew-symmetric matrix of the velocities is given as

$$\mathbf{V}^{l}=\begin{bmatrix}0 & -v_u & v_n\\ v_u & 0 & -v_e\\ -v_u & v_e & 0\end{bmatrix} \tag{6.35}$$

Multiplying the above matrix with Eq. (6.34) yields the fourth term of the right-hand side of Eq. (6.20)

$$\mathbf{V}^l(2\delta\boldsymbol{\omega}_{ie}^l + \delta\boldsymbol{\omega}_{el}^l) = \begin{bmatrix} 2\omega_e(v_u\sin\varphi + v_n\cos\varphi) + \frac{v_n\dot{\lambda}}{\cos\varphi} & 0 & 0 \\ -2\omega_e v_e\cos\varphi - \frac{v_e\dot{\lambda}}{\cos\varphi} & 0 & 0 \\ -2\omega_e v_e\sin\varphi & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} + \begin{bmatrix} \frac{-v_u}{(R_N+h)} + \frac{v_n\tan\varphi}{(R_N+h)} & 0 & 0 \\ \frac{-v_e\tan\varphi}{(R_N+h)} & \frac{-v_u}{(R_M+h)} & 0 \\ \frac{v_e}{(R_N+h)} & \frac{v_n}{(R_M+h)} & 0 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} \tag{6.36}$$

The final term on the right-hand side of Eq. (6.20) is the error in normal gravity due primarily to the error in the altitude

$$\delta\mathbf{g}^l = \begin{bmatrix} 0 \\ 0 \\ \frac{2g}{R}\delta h \end{bmatrix} \tag{6.37}$$

where g represents the normal component of gravity and R is the mean radius of the Earth.

By combining Eqs. (6.21, 6.22, 6.36 and 6.37) we get the complete definition of the time rate of change of the velocity errors

$$\delta\dot{\mathbf{v}}^l = \underbrace{\begin{bmatrix} 0 & f_u & -f_n \\ -f_u & 0 & f_e \\ f_n & -f_e & 0 \end{bmatrix} \begin{bmatrix} p \\ r \\ A \end{bmatrix}}_{1} + \underbrace{\begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} \delta f_x \\ \delta f_y \\ \delta f_z \end{bmatrix}}_{2} + \begin{bmatrix} 2\omega_e(v_u\sin\varphi + v_n\cos\varphi) + \frac{v_n\dot{\lambda}}{\cos\varphi} & 0 & 0 \\ -2\omega_e v_e\cos\varphi - \frac{v_e\dot{\lambda}}{\cos\varphi} & 0 & 0 \\ -2\omega_e v_e\sin\varphi & 0 & \frac{2g}{R} \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} + \begin{bmatrix} \frac{-v_u}{(R_N+h)} + \frac{v_n\tan\varphi}{(R_N+h)} & (2\omega_e+\dot{\lambda})\sin\varphi & -(2\omega_e+\dot{\lambda})\cos\varphi \\ -\dot{\lambda}\sin\varphi - (2\omega_e+\dot{\lambda})\sin\varphi & \frac{-v_u}{(R_M+h)} & -\dot{\varphi} \\ \dot{\lambda}\cos\varphi + (2\omega_e+\dot{\lambda})\cos\varphi & \frac{2v_n}{(R_M+h)} & 0 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} \tag{6.38}$$

Again the reader should remember that $\dot{\varphi} = \frac{v_n}{(R_M+h)}$ and $\dot{\lambda} = \frac{v_e}{(R_N+h)\cos\varphi}$, and hence

$$\delta\dot{\mathbf{v}}^l = \underbrace{\begin{bmatrix} 0 & f_u & -f_n \\ -f_u & 0 & f_e \\ f_n & -f_e & 0 \end{bmatrix}\begin{bmatrix} \delta p \\ \delta r \\ \delta A \end{bmatrix}}_{1} + \underbrace{\begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}\begin{bmatrix} \delta f_x \\ \delta f_y \\ \delta f_z \end{bmatrix}}_{2} +$$

$$\underbrace{\begin{bmatrix} 2\omega_e(v_u\sin\varphi + v_n\cos\varphi) + \frac{v_n v_e}{(R_N+h)\cos^2\varphi} & 0 & 0 \\ -2\omega_e v_e\cos\varphi - \frac{v_e^2}{(R_N+h)\cos^2\varphi} & 0 & 0 \\ -2\omega_e v_e\sin\varphi & 0 & 0 \end{bmatrix}\begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix}}_{3} +$$

$$\underbrace{\begin{bmatrix} \frac{-v_u}{(R_N+h)} + \frac{v_n\tan\varphi}{(R_N+h)} & \left(2\omega_e + \frac{v_e}{(R_N+h)\cos\varphi}\right)\sin\varphi & -\left(2\omega_e + \frac{v_e}{(R_N+h)\cos\varphi}\right)\cos\varphi \\ -\left(2\omega_e + \frac{v_e}{(R_N+h)\cos\varphi}\right)\sin\varphi & \frac{-\dot{h}}{(R_M+h)} & \frac{-v_n}{(R_M+h)} \\ \left(2\omega_e + \frac{v_e}{(R_N+h)\cos\varphi}\right)\cos\varphi & \frac{2v_n}{(R_M+h)} & 0 \end{bmatrix}\begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix}}_{4} \tag{6.39}$$

The third and the fourth terms on the right-hand side of Eq. (6.39) contain expressions where velocities are divided by the Earth's radius or multiplied by the Earth's rotation rate. Some of them are also multiplied by the latitude or longitude errors (in radians). As these terms are relatively minor they can safely be ignored for most applications and Eq. (6.39) can be approximated as

$$\begin{bmatrix} \delta\dot{v}_e \\ \delta\dot{v}_n \\ \delta\dot{v}_u \end{bmatrix} = \begin{bmatrix} 0 & f_u & -f_n \\ -f_u & 0 & f_e \\ f_n & -f_e & 0 \end{bmatrix}\begin{bmatrix} \delta p \\ \delta r \\ \delta A \end{bmatrix} + \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}\begin{bmatrix} \delta f_x \\ \delta f_y \\ \delta f_z \end{bmatrix} \tag{6.40}$$

In general, f_e and f_n are very small acceleration components (close to zero) in comparison to f_u, which is close to the acceleration due to gravity (9.8 m/sec^2). Hence there can be strong coupling between δv_e and δr, as well as between δv_n and δp. On the other hand, there will be weak coupling between δA and either δv_e or δv_n.

6.1.3 Attitude Errors for Local-Level Frame

The errors in the transformation matrix between the body and the computational frames are caused by angular velocity errors in measuring the rotation between the two frames. The Euler angles between the body frame and the l-frame are obtained from the transformation matrix R_b^l, which is gained (Schwarz and Wei 1999) by solving the following differential equation

$$\dot{R}_b^l = R_b^l\Omega_{lb}^b \tag{6.41}$$

Taking into account the measurement and computational errors, the calculated rate of change of the transformation matrix $\dot{R}_b^l$ is

$$\dot{\hat{R}}_b^l = \hat{R}_b^l \hat{\Omega}_{lb}^b \tag{6.42}$$

The computed transformation matrix $\hat{R}_b^l$ can be written as

$$\hat{R}_b^l = R_b^l + \delta R_b^l \tag{6.43}$$

By letting $\delta R_b^l = \Psi^l R_b^l$ we obtain

$$\hat{R}_b^l = \left(I + \Psi^l\right) R_b^l \tag{6.44}$$

where Ψ^l is the skew-symmetric matrix of the attitude errors arising from the transformation errors

$$\Psi^l = \begin{bmatrix} 0 & -\delta A & \delta r \\ \delta A & 0 & -\delta p \\ -\delta r & \delta p & 0 \end{bmatrix} \tag{6.45}$$

and the matrix $\left(I + \Psi^l\right)$ is basically an orthogonal transformation that contains a small rotation angle (expressed by Ψ^l) between the actual and the approximate computational reference frames.

To get the rate of change of the transformation matrix $\hat{R}_b^l$ in terms of the rate of change of attitude errors ε^l, Eq. (6.44) is differentiated to yield

$$\begin{aligned} \frac{d}{dt}\left(\hat{R}_b^l\right) &= \frac{d}{dt}\left(R_b^l + \Psi^l R_b^l\right) \\ \dot{\hat{R}}_b^l &= \dot{R}_b^l + \dot{\Psi}^l R_b^l + \Psi^l \dot{R}_b^l \end{aligned} \tag{6.46}$$

and since $\dot{R}_b^l = R_b^l \Omega_{lb}^b$

$$\dot{\hat{R}}_b^l = R_b^l \Omega_{lb}^b + \dot{\Psi}^l R_b^l + \Psi^l R_b^l \Omega_{lb}^b \tag{6.47}$$

Now we substitute $\hat{R}_b^l$ from Eq. (6.44) into Eq. (6.42) to obtain

$$\dot{\hat{R}}_b^l = \left(I + \Psi^l\right) R_b^l \hat{\Omega}_{lb}^b \tag{6.48}$$

and since $\hat{\Omega}_{lb}^b = \Omega_{lb}^b + \delta\Omega_{lb}^b$ this gives

$$\dot{\hat{R}}_b^l = \left(I + \Psi^l\right) R_b^l \left(\Omega_{lb}^b + \delta\Omega_{lb}^b\right) \tag{6.49}$$

$$\dot{\hat{R}}_b^l = \left(R_b^l + \Psi^l R_b^l\right)\left(\Omega_{lb}^b + \delta\Omega_{lb}^b\right) \tag{6.50}$$

$$\dot{\hat{R}}_b^l = R_b^l \Omega_{lb}^b + R_b^l \delta\Omega_{lb}^b + \Psi^l R_b^l \Omega_{lb}^b + \Psi^l R_b^l \delta\Omega_{lb}^b \tag{6.51}$$

where $\delta\Omega_{lb}^{b}$ represent the linear angular velocity errors.

Combining Eqs. (6.47 and 6.51) gives

$$R_b^l\Omega_{lb}^b + \dot{\Psi}^l R_b^l + \Psi^l R_b^l\Omega_{lb}^b = R_b^l\Omega_{lb}^b + R_b^l\delta\Omega_{lb}^b + \Psi^l R_b^l\Omega_{lb}^b + \Psi^l R_b^l\delta\Omega_{lb}^b \tag{6.52}$$

in which the following terms can be canceled from both sides

$$\begin{aligned} &\cancel{R_b^l\Omega_{lb}^b} + \dot{\Psi}^l R_b^l + \cancel{\Psi^l R_b^l\Omega_{lb}^b} = \cancel{R_b^l\Omega_{lb}^b} + R_b^l\delta\Omega_{lb}^b + \cancel{\Psi^l R_b^l\Omega_{lb}^b} + \Psi^l R_b^l\delta\Omega_{lb}^b \\ &\dot{\Psi}^l R_b^l = R_b^l\delta\Omega_{lb}^b + +\Psi^l R_b^l\delta\Omega_{lb}^b \end{aligned} \tag{6.53}$$

Substituting $\delta R_b^l = \Psi^l R_b^l$ gives

$$\dot{\Psi}^l R_b^l = R_b^l\delta\Omega_{lb}^l + \delta R_b^l\delta\Omega_{lb}^l \tag{6.54}$$

and by neglecting the second-order effect $\delta R_b^l\delta\Omega_{lb}^b$ we get

$$\begin{aligned} \dot{\Psi}^l R_b^l &= R_b^l\delta\Omega_{lb}^l \\ \dot{\Psi}^l &= R_b^l\delta\Omega_{lb}^b R_l^b \end{aligned} \tag{6.55}$$

which can be expressed in vector from as

$$\dot{\boldsymbol{\varepsilon}}^l = R_b^l\delta\boldsymbol{\omega}_{lb}^b \tag{6.56}$$

This shows how the time derivative of the attitude errors $\boldsymbol{\varepsilon} = [\delta p, \delta r, \delta A]^T$ is represented by angular velocity errors $\delta\boldsymbol{\omega}_{lb}^b$. The angular velocity of the body frame with respect to the l-frame, $\hat{\boldsymbol{\omega}}_{lb}^b$, is obtained by subtracting the angular velocity of the l-frame with respect to inertial frame $\hat{\boldsymbol{\omega}}_{il}^b$ from the rotational rate of the body frame as measured by the gyroscope $\tilde{\boldsymbol{\omega}}_{ib}^b$, therefore

$$\begin{aligned} \hat{\boldsymbol{\omega}}_{lb}^b &= \tilde{\boldsymbol{\omega}}_{ib}^b - \hat{\boldsymbol{\omega}}_{il}^b \\ \hat{\boldsymbol{\omega}}_{lb}^b &= \tilde{\boldsymbol{\omega}}_{ib}^b - \hat{R}_l^b\hat{\boldsymbol{\omega}}_{il}^l \end{aligned} \tag{6.57}$$

and by linearizing this and neglecting the second-order terms the angular velocity error $\delta\boldsymbol{\omega}_{lb}^b$ is

$$\begin{aligned} \delta\boldsymbol{\omega}_{lb}^b &= \delta\boldsymbol{\omega}_{ib}^b - \delta\left(R_l^b\boldsymbol{\omega}_{il}^l\right) \\ \delta\boldsymbol{\omega}_{lb}^b &= \delta\boldsymbol{\omega}_{ib}^b - \delta R_l^b\boldsymbol{\omega}_{il}^l - R_l^b\delta\boldsymbol{\omega}_{il}^l \\ \delta\boldsymbol{\omega}_{lb}^b &= \delta\boldsymbol{\omega}_{ib}^b - \left(\delta R_b^l\right)^T\boldsymbol{\omega}_{il}^l - R_l^b\delta\boldsymbol{\omega}_{il}^l \\ \delta\boldsymbol{\omega}_{lb}^b &= \delta\boldsymbol{\omega}_{ib}^b - \left(\Psi^l R_b^l\right)^T\boldsymbol{\omega}_{il}^l - R_l^b\delta\boldsymbol{\omega}_{il}^l \\ \delta\boldsymbol{\omega}_{lb}^b &= \delta\boldsymbol{\omega}_{ib}^b - R_l^b\left(\Psi^l\right)^T\boldsymbol{\omega}_{il}^l - R_l^b\delta\boldsymbol{\omega}_{il}^l \end{aligned} \tag{6.58}$$

By noticing that $\left(\Psi^l\right)^T = -\Psi^l$ and rearranging

$$\delta\boldsymbol{\omega}_{lb}^{b} = R_l^b \Psi^l \boldsymbol{\omega}_{il}^{l} - R_l^b \delta\boldsymbol{\omega}_{il}^{l} + \delta\boldsymbol{\omega}_{ib}^{b} \tag{6.59}$$

From the skew-symmetric matrix property $A\mathbf{b} = -B\mathbf{a}$, we can say that

$$\Psi^l \boldsymbol{\omega}_{il}^{l} = -\Omega_{il}^{l} \boldsymbol{\varepsilon}^l \tag{6.60}$$

where Ω_{il}^{l} is the skew-symmetric matrix corresponding to the vector $\boldsymbol{\omega}_{il}^{l}$, and $\boldsymbol{\varepsilon}^l$ is the attitude error vector corresponding to the skew-symmetric matrix Ψ^l.

Substituting into Eq. (6.59) gives

$$\delta\boldsymbol{\omega}_{lb}^{b} = -R_l^b \Omega_{il}^{l} \boldsymbol{\varepsilon}^l - R_l^b \delta\boldsymbol{\omega}_{il}^{l} + \delta\boldsymbol{\omega}_{ib}^{b} \tag{6.61}$$

from which it is apparent that the error in the computed angular velocity $\delta\boldsymbol{\omega}_{lb}^{b}$ has contributions from the coordinate transformation error $\boldsymbol{\varepsilon}^l$ between the b-frame and the l-frame, the error in the angular velocities $\boldsymbol{\omega}_{il}^{l}$, and the measurement errors $\delta\boldsymbol{\omega}_{ib}^{b}$.

Substituting $\delta\boldsymbol{\omega}_{lb}^{b}$ from Eq. (6.61) into Eq. (6.56) gives

$$\dot{\boldsymbol{\varepsilon}}^l = R_b^l \left(-R_l^b \Omega_{il}^{l} \boldsymbol{\varepsilon}^l - R_l^b \delta\boldsymbol{\omega}_{il}^{l} + \delta\boldsymbol{\omega}_{ib}^{b}\right) \tag{6.62}$$

and the set of differential equations for the attitude errors in the local-level frame

$$\dot{\boldsymbol{\varepsilon}}^l = -\Omega_{il}^{l} \boldsymbol{\varepsilon}^l - \delta\boldsymbol{\omega}_{il}^{l} + R_b^l \delta\boldsymbol{\omega}_{ib}^{b} \tag{6.63}$$

where the vector $\boldsymbol{\varepsilon}^l$contains the attitude errors δp, δr, δA, the vector $\delta\boldsymbol{\omega}_{il}^{l}$ represents the errors caused by navigation parameters errors ($\delta\mathbf{r}, \delta\mathbf{v}$, etc.), the vector $\delta\boldsymbol{\omega}_{ib}^{b}$ depicts the errors in the measurement of body rotational rates, and the term Ω_{il}^{l} represent the skew-symmetric matrix containing the angular velocities for the rotation of the l-frame with respect to the i-frame.

Now we will examine Eq. (6.63) in order to achieve a sense of what these terms really are.

We can write $\delta\boldsymbol{\omega}_{il}^{l}$ as a combination of following two terms

$$\delta\boldsymbol{\omega}_{il}^{l} = \delta\boldsymbol{\omega}_{ie}^{l} + \delta\boldsymbol{\omega}_{el}^{l} \tag{6.64}$$

As shown between Eqs. (6.26 and 6.33), we can express this in terms of position and velocity errors

$$\delta\boldsymbol{\omega}_{il}^{l} = \begin{bmatrix} 0 & 0 & \frac{\dot{\phi}}{(R_M+h)} \\ -\omega_e \sin\phi & 0 & -\frac{\dot{\lambda}\cos\phi}{(R_N+h)} \\ \omega_e \cos\phi + \frac{\dot{\lambda}}{\cos\phi} & 0 & -\frac{\dot{\lambda}\sin\phi}{(R_N+h)} \end{bmatrix} \begin{bmatrix} \delta\phi \\ \delta\lambda \\ \delta h \end{bmatrix} + \begin{bmatrix} 0 & \frac{-1}{(R_M+h)} & 0 \\ \frac{1}{(R_N+h)} & 0 & 0 \\ \frac{\tan\phi}{(R_N+h)} & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} \tag{6.65}$$

Similarly, Ω_{il}^{l} can be written as

$$\Omega_{il}^{l} = \Omega_{ie}^{l} + \Omega_{el}^{l} \tag{6.66}$$

$$\Omega_{il}^{l} = \begin{bmatrix} 0 & -\dot{\lambda}\sin\phi & \dot{\lambda}\cos\phi \\ \dot{\lambda}\sin\phi & 0 & \dot{\phi} \\ -\dot{\lambda}\cos\phi & -\dot{\phi} & 0 \end{bmatrix} + \begin{bmatrix} 0 & -\omega_e\sin\phi & \omega_e\cos\phi \\ \omega_e\sin\phi & 0 & 0 \\ -\omega_e\cos\phi & 0 & 0 \end{bmatrix} \tag{6.67}$$

$$\Omega_{il}^{l} = \begin{bmatrix} 0 & -\left(\omega_e + \dot{\lambda}_e\right)\sin\phi & \left(\omega_e + \dot{\lambda}_e\right)\cos\phi \\ \left(\omega_e + \dot{\lambda}_e\right)\sin\phi & 0 & \dot{\phi} \\ -\left(\omega_e + \dot{\lambda}_e\right)\cos\phi & -\dot{\phi} & 0 \end{bmatrix} \tag{6.68}$$

Substituting Eqs. (6.65 and 6.68) in Eq. (6.63) and rearranging the terms gives

$$\dot{\boldsymbol{\varepsilon}}^{l} = \underbrace{\begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{-1}{(R_N+h)} & 0 & 0 \\ \frac{-\tan\phi}{(R_N+h)} & 0 & 0 \end{bmatrix}\begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix}}_{1} + \underbrace{\begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}\begin{bmatrix} \delta\omega_x \\ \delta\omega_y \\ \delta\omega_z \end{bmatrix}}_{2} + \underbrace{\begin{bmatrix} 0 & \left(\omega_e + \dot{\lambda}_e\right)\sin\phi & -\left(\omega_e + \dot{\lambda}_e\right)\cos\phi \\ -\left(\omega_e + \dot{\lambda}_e\right)\sin\phi & 0 & -\dot{\phi} \\ \left(\omega_e + \dot{\lambda}_e\right)\cos\phi & \dot{\phi} & 0 \end{bmatrix}\begin{bmatrix} \delta p \\ \delta r \\ \delta A \end{bmatrix}}_{3} + \underbrace{\begin{bmatrix} 0 & 0 & \frac{-\dot{\phi}}{(R_M+h)} \\ \omega_e\sin\phi & 0 & -\frac{\dot{\lambda}\cos\phi}{(R_N+h)} \\ -\omega_e\cos\phi + \frac{\dot{\lambda}}{\cos\phi} & 0 & -\frac{\dot{\lambda}\sin\phi}{(R_N+h)} \end{bmatrix}\begin{bmatrix} \delta\phi \\ \delta\lambda \\ \delta h \end{bmatrix}}_{4} \tag{6.69}$$

The components of the third and the fourth terms are either divided by the Earth's radius (or its square) or are multiplied by the rotation of the Earth and as such can safely be ignored for most navigational applications. This simplifies the time rate of change of the attitude errors to

$$\dot{\boldsymbol{\varepsilon}}^{l} = \begin{bmatrix} \delta\dot{p} \\ \delta\dot{r} \\ \delta\dot{A} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{(R_M+h)} & 0 \\ \frac{-1}{(R_N+h)} & 0 & 0 \\ \frac{-\tan\phi}{(R_N+h)} & 0 & 0 \end{bmatrix}\begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} + \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}\begin{bmatrix} \delta\omega_x \\ \delta\omega_y \\ \delta\omega_z \end{bmatrix} \tag{6.70}$$

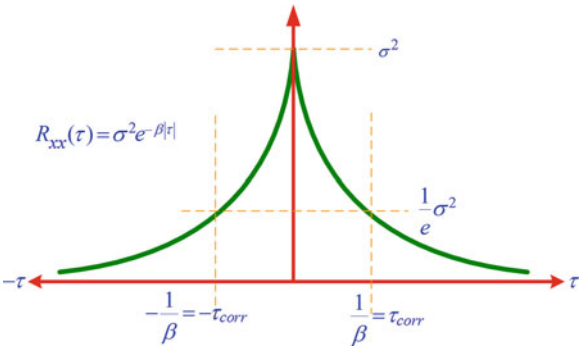


Fig. 6.1 The autocorrelation sequence of a random variable described by the first-order Gauss-Markov process

in which it is evident that δr is coupled with δv_e, and δp is coupled with δv_n. These implicit interrelationships are known as the Schuler effect.

6.1.4 Inertial Sensor Error States

The set of linear differential Eqs. (6.15, 6.16 and 6.63) are non-homogeneous due to the effects of sensor errors which are mainly gyroscope drifts and accelerometer biases that possess deterministic as well as non-deterministic parts. The deterministic part is computed during calibration in the laboratory and compensated in the measurements. The non-deterministic part of the sensor errors is random and modeled by stochastic models. These errors are usually correlated in time, and the methods commonly used to model them include the random walk process, the first-order Gauss-Markov (GM) process, and the autoregressive (AR) process (Noureldin et al. 2009).

The first-order GM process is commonly used to model the stochastic sensor errors, and its general form is

$$\dot{x} = -\beta x + \sqrt{2\beta\sigma^2}w \tag{6.71}$$

where

x	is the random process
β	is the reciprocal of the correlation time of the process
W	is zero-mean uncorrelated Gaussian noise vector of unit variance
σ^2	is the variance of the white noise associated with the random process.

The first-order GM model is a decaying exponential autocorrelation sequence of x, as shown in Fig. 6.1.

6.1.4.1 Accelerometer Bias Errors

For accelerometer stochastic biases, the time rate of change of the accelerometer bias errors can be expressed as

$$\delta\dot{\mathbf{f}}^b = \begin{bmatrix} \delta\dot{f}_x \\ \delta\dot{f}_y \\ \delta\dot{f}_z \end{bmatrix} = \begin{bmatrix} -\beta_{fx} & 0 & 0 \\ 0 & -\beta_{fy} & 0 \\ 0 & 0 & -\beta_{fz} \end{bmatrix} \begin{bmatrix} \delta f_x \\ \delta f_y \\ \delta f_z \end{bmatrix} + \begin{bmatrix} \sqrt{2\beta_{fx}\sigma_{fx}^2} \\ \sqrt{2\beta_{fy}\sigma_{fy}^2} \\ \sqrt{2\beta_{fz}\sigma_{fz}^2} \end{bmatrix} w(t) \qquad (6.72)$$

where

$\beta_{fx}, \beta_{fy}, \beta_{fz}$	are the reciprocals of the correlation times associated with the autocorrelation sequence of δf_x, δf_y and δf_z
$\sigma_{fx}^2, \sigma_{fy}^2, \sigma_{fz}^2$	are the variances associated with the accelerometer errors
$w(t)$	is white Gaussian noise with variance equal to one.

6.1.4.2 Gyroscope Drift Errors

For gyroscope stochastic biases, the time rate of change of the bias errors can be written as

$$\delta\dot{\boldsymbol{\omega}} = \begin{bmatrix} \delta\dot{\omega}_x \\ \delta\dot{\omega}_y \\ \delta\dot{\omega}_z \end{bmatrix} = \begin{bmatrix} -\beta_{\omega x} & 0 & 0 \\ 0 & -\beta_{\omega y} & 0 \\ 0 & 0 & -\beta_{\omega z} \end{bmatrix} \begin{bmatrix} \delta\omega_x \\ \delta\omega_y \\ \delta\omega_z \end{bmatrix} + \begin{bmatrix} \sqrt{2\beta_{\omega x}\sigma_{\omega x}^2} \\ \sqrt{2\beta_{\omega y}\sigma_{\omega y}^2} \\ \sqrt{2\beta_{\omega z}\sigma_{\omega z}^2} \end{bmatrix} w(t) \quad (6.73)$$

where

$\beta_{\omega x}, \beta_{\omega y}, \beta_{\omega z}$	are the reciprocals of the correlation times associated with the autocorrelation sequence of $\delta\omega_x$, $\delta\omega_y$ and $\delta\omega_z$
$\sigma_{\omega x}^2, \sigma_{\omega y}^2, \sigma_{\omega z}^2$	are the variances associated with the gyroscope errors
$w(t)$	is white Gaussian noise with variance equal to one.

For both the accelerometer and the gyroscope error models, the white noise term $w(t)$ is of unity variance because the variance has been included inside the process. The above equations for the stochastic drift and bias are in the b-frame and, being independent of any reference frame, apply equally to all computational frames.

6.1.5 Summary of Local-Level Frame Error State Equations

From the position, velocity and attitude error state Eqs. (6.14, 6.40 and 6.70), and the sensor error Eqs. (6.72 and 6.73), the state equations for the INS error states in the l-frame can be summarized in a compact form as

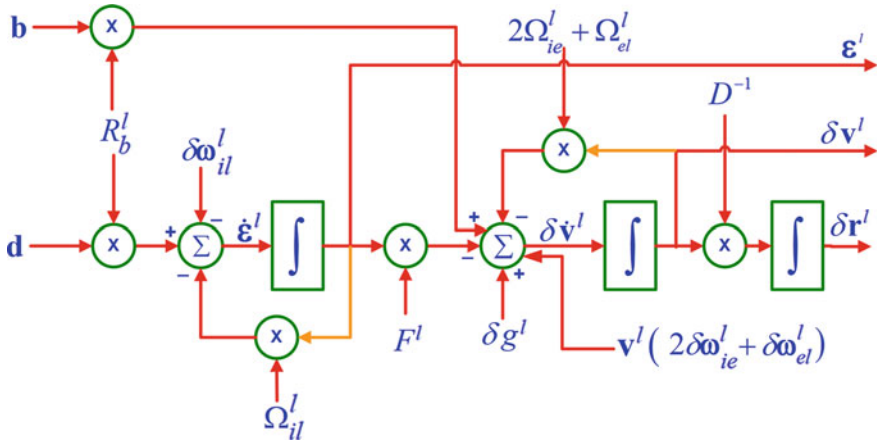


Fig. 6.2 Flow diagram of the error state equation in the l-frame

$$\dot{\mathbf{x}}^l(t) = \begin{bmatrix} \delta\dot{\mathbf{r}}^l \\ \delta\dot{\mathbf{v}}^l \\ \dot{\boldsymbol{\varepsilon}}^l \\ \delta\dot{\boldsymbol{\omega}}_{ib}^b \\ \delta\dot{\mathbf{f}}^b \end{bmatrix} = \begin{bmatrix} D^{-1}\delta\mathbf{v}^l \\ -\mathbf{F}^l\varepsilon^l - 2(\boldsymbol{\Omega}_{el}^l + \boldsymbol{\Omega}_{el}^l)\delta\mathbf{v}^l - \mathbf{V}^l(2\delta\omega_{ie}^l + \delta\omega_{ie}^l) + \delta\mathbf{g}^l + R_b^l\delta\mathbf{f}^b \\ -\boldsymbol{\Omega}_{il}^l\varepsilon^l - \delta\omega_{il}^l + R_b^l\delta\omega_{ib}^b \\ -\beta_\omega\delta\omega_{ib}^b + \sqrt{2\beta_\omega\sigma_\omega^2}\mathbf{w} \\ -\beta_f\delta\mathbf{f}^b + \sqrt{2\beta_f\sigma_f^2}\mathbf{w} \end{bmatrix} \tag{6.74}$$

where

$$\beta_\omega = \begin{bmatrix} -\beta_{\omega x} & 0 & 0 \\ 0 & -\beta_{\omega y} & 0 \\ 0 & 0 & -\beta_{\omega z} \end{bmatrix}$$

$$\beta_f = \begin{bmatrix} -\beta_{fx} & 0 & 0 \\ 0 & -\beta_{fy} & 0 \\ 0 & 0 & -\beta_{fz} \end{bmatrix}$$

The conceptual flow diagram of the above equations is shown in Fig. 6.2.

The above set of equations are the first-order differentiation of the error states and can be represented in a compact form as

$$\dot{\mathbf{x}}^l = F^l\mathbf{x}^l + G^l\mathbf{w} \tag{6.75}$$

Now we expand Eq. (6.74) as Eq. (6.76) to present the individual terms to see the overall picture and understand which terms are coupled together.

$$
\begin{bmatrix}
\delta\dot{\varphi}\\ \delta\dot{\lambda}\\ \delta\dot{h}\\ \delta\dot{V}_e\\ \delta\dot{V}_n\\ \delta\dot{V}_u\\ \delta\dot{p}\\ \delta\dot{r}\\ \delta\dot{y}\\ \delta\dot{\omega}_x\\ \delta\dot{\omega}_y\\ \delta\dot{\omega}_z\\ \delta\dot{f}_x\\ \delta\dot{f}_y\\ \delta\dot{f}_z
\end{bmatrix}
=
\begin{pmatrix}
0 & 0 & 0 & 0 & \frac{1}{R_M+h} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & f_u & -f_n & 0 & 0 & 0 & R_b^l & R_b^l & R_b^l\\
0 & 0 & 0 & 0 & 0 & 0 & -f_u & 0 & f_e & 0 & 0 & 0 & R_b^l & R_b^l & R_b^l\\
0 & 0 & 0 & 0 & 0 & 0 & f_n & -f_e & 0 & 0 & 0 & 0 & R_b^l & R_b^l & R_b^l\\
0 & 0 & 0 & 0 & \frac{1}{R_M+h} & 0 & 0 & 0 & 0 & R_b^l & R_b^l & R_b^l & 0 & 0 & 0\\
0 & 0 & 0 & -\frac{1}{R_N+h} & 0 & 0 & 0 & 0 & 0 & R_b^l & R_b^l & R_b^l & 0 & 0 & 0\\
0 & 0 & 0 & -\frac{\tan\varphi}{R_N+h} & 0 & 0 & 0 & 0 & 0 & R_b^l & R_b^l & R_b^l & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta_{\omega x} & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{\omega y} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{\omega z} & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{fx} & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{fy} & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{fz}
\end{pmatrix}
\begin{bmatrix}
\delta\varphi\\ \delta\lambda\\ \delta h\\ \delta V_e\\ \delta V_n\\ \delta V_u\\ \delta p\\ \delta r\\ \delta y\\ \delta\omega_x\\ \delta\omega_y\\ \delta\omega_z\\ \delta f_x\\ \delta f_y\\ \delta f_z
\end{bmatrix}
+ Gw
\tag{6.76}
$$

The distribution vector of Eq. (6.76) consists of the variances and is

$$
\mathbf{G} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2\beta_{\omega x}\sigma_{\omega x}^2} & \sqrt{2\beta_{\omega y}\sigma_{\omega y}^2} & \sqrt{2\beta_{\omega z}\sigma_{\omega z}^2}\\
 & & & & & & & & & \sqrt{2\beta_{fx}\sigma_{fx}^2} & \sqrt{2\beta_{fy}\sigma_{fy}^2} & \sqrt{2\beta_{fz}\sigma_{fz}^2}
\end{bmatrix}
\tag{6.77}
$$

6.2 Schuler Effect

The Schuler effect is related to the coupling in the horizontal plane between the velocity errors δv_e and δv_n, and between the pitch and roll errors δp andδr. As will be explained later, this coupling bounds the errors in both the horizontal and vertical velocities and the pitch and roll angles.

6.2.1 Error Model Along the East Channel

From the modeling of the attitude errors in Eq. (6.70) the rate of change of the error in the roll angle can be written as

$$
\delta\dot{r} = \frac{-1}{R_N + h}\delta v_e + \delta\omega_n \tag{6.78}
$$

And similarly from Eq. (6.40) the rate of change of the error in the east velocity can be written as

$$\delta\dot{v}_e = f_u\delta r - \underbrace{f_n\delta A}_{\text{very small}} + \delta f_e \tag{6.79}$$

For terrestrial navigation the acceleration component f_u is relatively strong (a value close to that of gravity) and it is usually much larger than f_n. This results in a strong coupling between δv_e and δr. And because the term $f_n\delta A$ can be neglected Eq. (6.79) reduces to

$$\delta\dot{v}_e = g\delta r + \delta f_e \tag{6.80}$$

where f_u is replaced by g for simplicity.

If we differentiate this and substitute the expression from the rate of the change of the roll error in Eq. (6.78) the following non-homogenous linear second-order differential equation can be obtained

$$\delta\ddot{v}_e + \frac{g}{R_N + h}\delta v_e = g\delta\omega_n \tag{6.81}$$

the solution of which yields velocity errors δv_e oscillating over time with a very small frequency equal to 1/5000 Hz called the Schuler frequency f_s with a time interval of 84.4 min (King 1998). Consequently, the velocity error becomes bounded over time. The Schuler frequency is

$$f_s = \frac{1}{2\pi}\sqrt{\frac{g}{R_N + h}} \cong \sqrt{\frac{g}{R}} \tag{6.82}$$

where g represents gravity and R is the mean radius of the Earth.

Similarly, if we differentiate both sides of Eq. (6.78) and substitute $\delta\dot{v}_e$ from Eq. (6.80) we get the following non-homogenous linear second-order differential equation

$$\delta\ddot{r} + \frac{g}{R_N + h}\delta r = \frac{-1}{R_N + h}\delta f_e \tag{6.83}$$

the solution of which yields roll errors δr oscillating over time at the Schuler frequency. As a result, the attitude error becomes bounded over time.

The Fig. 6.3 illustrates the strong coupling between the east velocity error δv_e and roll error δr which implies that if the east velocity of the INS is updated from some external source (like GPS) to enable accurate estimation of δv_e, this will result in an accurate estimation of δr. In other words, the velocity update makes δv_e an observable component of the estimator, while the strong coupling of δv_e with δr, will make the roll error also observable.

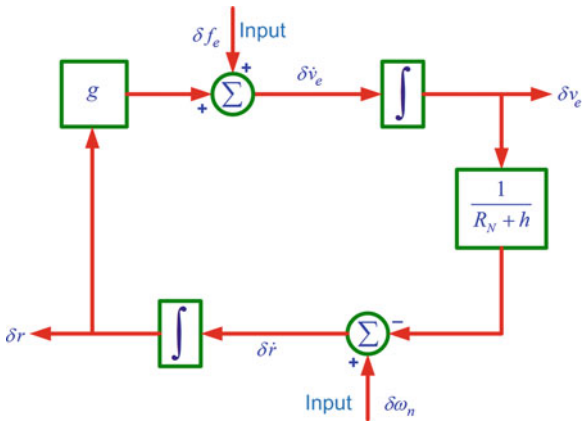


Fig. 6.3 Illustration of the schuler loop-east channel

6.2.2 Error Model Along the North Channel

The Schuler effect manifests itself in the north channel as well. From the modeling of the attitude errors in Eq. (6.70) the error in the time rate of change of the pitch angle is

$$\delta \dot{p} = \frac{1}{R_M + h} \delta v_n + \delta \omega_e \tag{6.84}$$

From Eq. (6.40) the rate of change of the error in the north velocity is

$$\delta \dot{v}_n = -f_u \delta p + f_e \delta A + \delta f_n \tag{6.85}$$

For similar reasons to those mentioned for the east channel, the error in the rate of change of the north velocity can be simplified as

$$\delta \dot{v}_n = -g \delta p + \delta f_n \tag{6.86}$$

By differentiating this and substituting the expression for the rate of change of the pitch error in Eq. (6.84) we obtain the following non-homogenous linear second-order differential equation

$$\delta \ddot{v}_n + \frac{g}{R_M + h} \delta v_n = -g \delta \omega_e \tag{6.87}$$

the solution of which yields velocity errors δv_n oscillating over time at the Schuler frequency. As a result, the north velocity error becomes bounded over time.

If we now differentiate both sides of Eq. (6.84) and substitute $\delta \dot{v}_n$ from Eq. (6.86) then we obtain the following non-homogenous linear second-order differential equation

$$\delta \ddot{p} + \frac{g}{R_M + h} \delta p = \frac{1}{R_M + h} \delta f_n \tag{6.88}$$

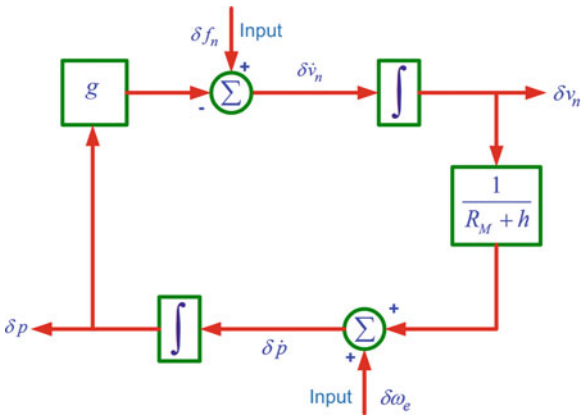


Fig. 6.4 Illustration of the schuler loop-north channel

the solution of which yields roll errors δr oscillating over time with Schuler frequency f_s. As a result, the roll attitude error becomes bounded over time.

The strong coupling between the north velocity error δv_n and pitch error δp is shown in Fig. 6.4. This behaves in a similar manner to the east channel. When the north velocity of the INS is updated from an external source, δv_n is estimated accurately and because of the strong mutual coupling δp will also be estimated accurately. Therefore, the velocity update not only makes δv_n an observable to the estimator but also δp due to the strong coupling between them.

6.2.3 *Understanding the Error Behavior of the Inertial System*

The error analysis of the inertial system in the previous section indicates that the velocity errors δv_e and δv_n are composed of two components

The Schuler part: which gives the strong coupling between δv_e and δr, and between δv_n and δp.
The non-stationary part: which depends upon the body acceleration components along the horizontal plane and relates the velocity errors to the azimuth error.

These components can be rewritten as follows

$$\begin{array}{ccc} \delta \dot{v}_{e,Sch} = f_u \delta r + \delta f^e & \& & \delta \dot{v}_{e,nst} = -f_n \delta A \\ \underbrace{\delta \dot{v}_{n,Sch} = -f_u \delta p + \delta f^n}_{\text{Schuler part}} & \& & \underbrace{\delta \dot{v}_{n,nst} = f_e \delta A}_{\text{Non-stationary part}} \end{array} \tag{6.89}$$

Owing to the Schuler effect, the Schuler parts of the velocity errors δv_e and δv_n (related to δr and δp respectively) are bounded in time. Since there is no strong coupling with the azimuth error the second part, which is related to δA, continues to change with time depending on the azimuth drift and so is known as the non-stationary part. In fact, with velocity updates any optimal estimation tool can

estimate and compensate only for the Schuler part of the velocity errors. On the other hand, if not updated the non-stationary parts of both δv_e and δv_n can jeopardize the long term accuracy of the position components along the north and the east directions.

If we twice integrate the time rate of change of the non-stationary parts of the velocity errors along the east and the north directions as follows

$$\delta \dot{v}_{e,nst} = -f_n \delta A \rightarrow \iint \rightarrow \delta P_{e,nst}(t_k) = \delta P_{e,nst}(t_{k-1}) - v_n \delta A \, \Delta t \tag{6.90}$$

and

$$\delta \dot{v}_{n,nst} = f_e \delta A \rightarrow \iint \rightarrow \delta P_{n,nst}(t_k) = \delta P_{n,nst}(t_{k-1}) + v_e \delta A \, \Delta t \tag{6.91}$$

it is evident that the non-stationary parts of δp_e and δp_n drift with time in a rate of $v_n \delta A$ and $v_e \delta A$ respectively. Hence the azimuth error δA has an important role in determining the long term position accuracy. Since it is modulated by the velocity components v_n and v_e, an azimuth error is particularly significant for a vehicle traveling at high velocity.

References

Chiang KW, Noureldin A, El-Sheimy N (2006) The utilization of artificial neural networks for multi-sensor system integration in navigation and positioning instruments. Paper presented at the IEEE transactions on instrumentation and measurement

Georgy J, Iqbal U, Bayoumi M, Noureldin A (2008) Reduced inertial sensor system (RISS)/GPS integration using particle filtering for land vehicles. In: Proceedings of the 21th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2008), avannah, Georgia, USA, ION GNSS 2008, pp 30–37

Georgy J, Karamat T, Iqbal U, Noureldin A (2011) Enhanced MEMS-IMU/Odometer/GPS integration using mixture particle filter. GPS Solutions 15(3):239–252

Georgy J, Karamat TB, Iqbal U, Noureldin A (2010) Enhanced MEMS-IMU/odometer/GPS integration using mixture particle filter. GPS solutions:1–14. doi:10.1007/s10291-010-0186-4

Karamat TB, Georgy J, Iqbal U, Noureldin AA (2009) Tightly-coupled reduced multi-sensor system for urban navigation. In: Proceedings of the 22nd international technical meeting of the satellite division of the institute of navigation (ION GNSS 2009), Savannah, GA, USA, 22–25 Sep 2009

King AD (1998) Inertial navigation—forty years of evolution. GEC Rev 13(3):140–149

Noureldin A, El-Shafie A, El-Sheimy N (2007) Adaptive neuro-fuzzy module for INS/GPS integration utilizing position and velocity update with real-time cross validation. IET Radar Sonar Navig 1(5):388–396

Noureldin A, Karamat TB, Eberts MD, El-Shafie A (2009) Performance enhancement of MEMS-Based INS/GPS integration for low-cost navigation applications. IEEE Trans Veh Technol 58(3):1077–1096

Schwarz KP, Wei M (1999) ENGO 623: Partial lecture on INS/GPS integration for geodetic applications, University of Calgary, department of geomatics engineering

Chapter 7
Kalman Filter

As stated in the previous chapter the accuracy of an INS is affected by the errors in the inertial sensors, initialization and computational algorithms. The situation is worse for the low cost MEMS sensors where the INS output can drift rapidly and render them essentially unusable as standalone sensors for navigation applications owing to severe stochastic errors. The main errors are accelerometer biases and gyro drifts. In order to improve the accuracy of an INS, k regular time intervals must be made to estimate the stochastic errors in the inertial sensors as a basis for compensation. As noted in the case of velocities and attitude angles along the horizontal channel, in the absence of ongoing optimal estimation of the velocity errors and their removal they will result in significant inaccuracies in the values of v_n and/or v_e over the long term. These velocity values modulate the azimuth error and jeopardize the positioning accuracy in the long term. Hence it is vital to establish accurate computation of the velocity components in order to cancel the pitch and the roll errors and thereby limit the growth rate of the position components as per the Schuler effect discussed in the previous chapter. Typically, the following types of measurement updates are used

1. Position or coordinate update known as "CUPT".
2. Velocity or zero velocity update known as "ZUPT".
3. Attitude update.

Various methods are available to update an INS with external measurements, including Kalman filtering (KF), particle filtering (PF) and artificial intelligence (AI). Traditionally, Kalman filtering is used when integrating an INS with aiding systems such as GPS for the estimation and compensation of the inertial sensors errors. In fact, the very first practical use of KF was in the Apollo program in the 1960 s for navigating in deep space.

In general, a Kalman filter is an algorithm for optimally estimating the error states of a system from measurements contaminated by noise. This is a sequential recursive algorithm that provides an optimal least mean variance estimation of the

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_7,

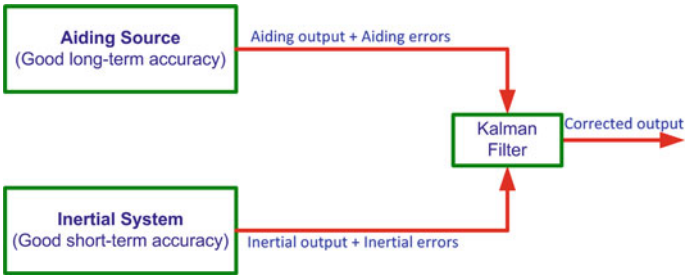


Fig. 7.1 Typical use of KF in a navigation application

error states. In addition to its benefits as an optimal estimator, KF provides real-time statistical data related to the estimation accuracy of the error states, which is very useful for quantitative error analysis. The beauty of KF is that it utilizes all of the available measurements, regardless of their precision, to estimate the current value of the state of the system by appropriately weighting these measurements.

KF uses the following information (Maybeck 1979) to accomplish its task

1. Information about the model of the system and its measurements.
2. Statistical knowledge of the systems noise, the measurement errors, and the uncertainty in the system model.
3. Available knowledge about the initial conditions of the system states.

For inertial navigation applications Kalman filtering is used in complementary configuration in which redundant measurements of the same signal with different noise characteristics are combined as a means of minimizing the error (Brown and Hwang 1997). An INS provides good high frequency information, but its errors grow increasingly over time owing to the implicit mathematical integration in the mechanization algorithm, which causes the bias errors of both the accelerometers and gyroscopes to accumulate at the output. The bias errors of the inertial sensors usually appear at the low frequency part of the sensor output and they are known as long term errors. On the other hand many other navigational systems (such as GPS) provide good low frequency characteristics and are prone to high frequency noise. Hence KF is employed to benefit from accurate low frequency data from an external source and therefore limit the long term errors of an INS. There are many sources that can provide reliable external aid for an INS, including radar, GPS, the speed from an air data system, wheel sensors, laser ranging, and a stored image data base. For the purpose of the book we will consider GPS as the external aiding source because of its accuracy, global availability and low cost. Figure 7.1 depicts a typical use of KF in navigation, where data from external sensors with good long term accuracy is combined with INS data with good short term accuracy in order to provide the best overall estimate of position, velocity and attitude.

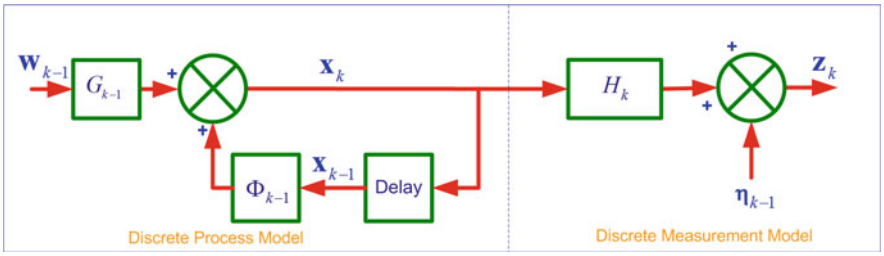


Fig. 7.2 The discrete-time process and measurement model

KF methodologies are therefore used to implement

1. Continuous position and velocity updates from sensors and systems other than INS (e.g. GPS).
2. Zero velocity updates (ZUPT) by halting the vehicle at regular intervals.
3. Coordinate updates (CUPT) at certain control stations whose coordinates are well known.

7.1 Discrete-Time KF

The application of KF requires that both the system and the measurement models of the underlying process be linear. A discrete-time linear system can be described as

$$\mathbf{x}_k = \Phi_{k,k-1}\,\mathbf{x}_{k-1} + G_{k-1}\mathbf{w}_{k-1} \tag{7.1}$$

where

$\mathbf{x}_k$	is the state vector
$\Phi_{k,k-1}$	is the state transition matrix (STM)
G_{k-1}	is the noise distribution matrix
$\mathbf{w}_{k-1}$	is the process noise vector
k	is the measurement epoch.

Since a noise source can affect more than one component of the state vector of a dynamic system, we introduced a noise distribution vector G which takes into account the coupling of common noise disturbances into various components of the state dynamics.

The discrete-time linear measurement equation of the system is

$$\mathbf{z}_k = H_k\mathbf{x}_k + \boldsymbol{\eta}_k \tag{7.2}$$

where

$\mathbf{z}_k$	is the measurement vector of the system output
H_k	is the observation or design
$\boldsymbol{\eta}_k$	is the measurement noise.

Figure 7.2 shows the discrete-time system corresponding to Eqs. (7.1) and (7.2).

The state transition matrix (STM) Φ represents the known dynamic behavior of the system (in this case the INS error model) which relates the state vector from epoch $k-1$ to k. Given the dynamic coefficient matrix F of a continuous time system the STM is

$$\Phi = \exp(F\Delta t) \tag{7.3}$$

To linearize this for use by KF we take the first two terms of the Taylor series expansion of the equation as follows

$$\Phi = (I + F\Delta t) \tag{7.4}$$

where I is identity matrix and Δt is sampling interval.

7.1.1 KF Assumptions

Kalman filtering relies on the following assumptions (Maybeck 1979; Minkler and Minkler 1993).

1. The system (both the process and the measurements) can be described by linear models.
2. The system noise $\mathbf{w}_k$ and the measurement noise $\boldsymbol{\eta}_k$ are uncorrelated zero-mean white noise processes with known auto covariance functions, hence

$$E[\mathbf{w}_k] = 0, \quad E[\boldsymbol{\eta}_k] = 0 \quad \forall k \tag{7.5}$$

$$E[\mathbf{w}_k \boldsymbol{\eta}_j^T] = 0 \quad \forall k, j \tag{7.6}$$

$$E[\mathbf{w}_k \mathbf{w}_j^T] = \begin{cases} Q_k, & k = j \\ 0 & k \neq j \end{cases} \tag{7.7}$$

$$E[\boldsymbol{\eta}_k \boldsymbol{\eta}_j^T] = \begin{cases} R_k, & k = j \\ 0 & k \neq j \end{cases} \tag{7.8}$$

   where Q_k and R_k are known positive definitive matrices. In INS/GPS integration, Q_k represents the covariance matrix of the system noise associated with the INS errors, and R_k represents the covariance matrix of the measurement noise associated with the GPS position and velocity updates.
3. The initial system state vector $\mathbf{x}_0$ is a random vector uncorrelated to both the process and measurement noises, hence

$$E[\mathbf{x}_0 \mathbf{w}_k^T] = 0, \quad E[\mathbf{x}_0 \, \boldsymbol{\eta}_k^T] = 0 \quad \forall k \tag{7.9}$$

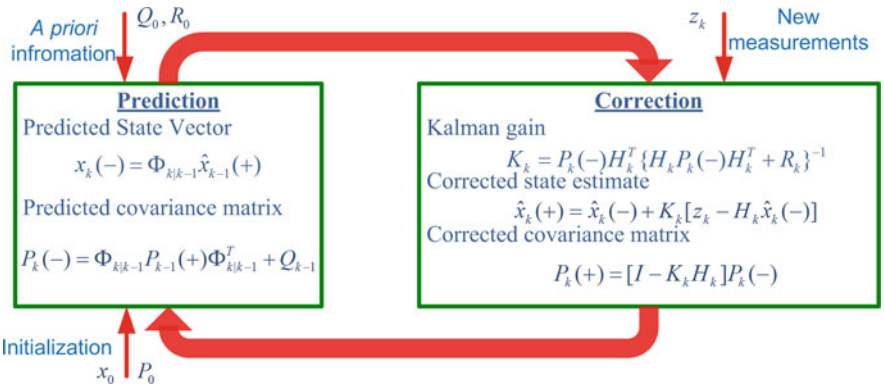


Fig. 7.3 The KF recursive process of 'prediction' and 'correction'

4. The mean value of the initial state $\bar{\mathbf{x}}_0$ and its covariance matrix P_0 are known, and can be expressed as

$$\bar{\mathbf{x}}_0 = E[\mathbf{x}_0] \tag{7.10}$$

$$P_0 == E[(\mathbf{x}_0 - \bar{\mathbf{x}}_0)(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T] \tag{7.11}$$

Under these assumptions KF is the optimal filter of any conceivable form.

7.2 KF Procedure

KF is a recursive algorithm which estimates the states of a system by operating as a feedback loop. Based on the known system model the filter estimates the system state at an epoch and then receives the feedback through measurements which are themselves contaminated by noise (Welch and Bishop 2001).

As shown in Fig. 7.3 the operation of a KF has two phases: (1) prediction or time update, and (2) correction or measurement update. In the prediction phase the system model is applied to propagate both the current state of the system and its covariance estimates from epoch $k-1$ to k. Then in the correction phase the measurements are used to update the previous estimates, thereby improving on the last estimate. Table 7.1 explains the variables in the KF literature, most of which will be further explained in the ensuing discussion.

Table 7.1 Important terms encountered in the KF literature

Φ: State transition matrix of a discrete-time linear dynamic system	H: Measurement sensitivity matrix or observation matrix which defines the linear relationship between the state of the dynamic systems and measurements that can be made
$\mathbf{x}$: State vector of a linear dynamic system	
$\hat{\mathbf{x}}(-)$: Predicted or *a priori* value of the estimated state vector of a linear dynamic system	$P(-)$: Predicted or *a priori* matrix of the estimation covariance of state estimation uncertainty in matrix form
$\hat{\mathbf{x}}(+)$: Corrected or *a posteriori* value of the estimated state vector of a linear dynamic system	$P(+)$: Corrected or *a posteriori* matrix of estimation covariance of state estimation uncertainty in matrix form
$\mathbf{z}$: Measurement vector or observation vector	$\mathbf{w}$: Process noise
K: Kalman gain matrix	$\boldsymbol{\eta}$: Measurement noise

7.2.1 Time Update or Prediction

The estimate of the system state $\mathbf{x}$ at time k given only the information up to time $k-1$, is called prediction $\hat{\mathbf{x}}_k(-)$. It is also the *a priori* estimate because it applies 'prior' to a measurement. Since the system noise is zero-mean, the best prediction of the state at time k is

$$\hat{\mathbf{x}}_k(-) = \Phi_{k|k-1}\mathbf{x}_{k-1}(+) \tag{7.12}$$

where $\hat{\mathbf{x}}_{k-1}(+)$ is the best estimate of state during the last epoch and is called the *a posteriori* estimate (described later). This is solely based on the process model, which is represented by the STM $\Phi_{k|k-1}$. KF also propagates the uncertainty about its estimate from epoch $k-1$ to k. This is called error covariance, and is the expected value of the variance of the error in the states at time k given all the information up to time $k-1$. It is represented by the covariance matrix $P_k(-)$, also known as the *a priori* covariance matrix

$$P_k(-) = \Phi_{k|k-1}P_{k-1}(+)\Phi_{k|k-1}^T + G_{k-1}Q_{k-1}G_{k-1}^T \tag{7.13}$$

where $P_{k-1}(+)$ represents the best estimate of the covariance in the last epoch and is based upon the measurement at epoch $k-1$. This is the *a posteriori* estimate of covariance (described later). It should be noted that the *a priori* covariance matrix $P_k(-)$ depends on both the process noise and the *a posteriori* covariance $P_{k-1}(+)$.

7.2.2 Measurement Update or Correction

After the KF has predicted its estimate of the state, this is corrected whenever a measurement from an external source becomes available. Firstly, based upon the

measurement covariance R_k, a weighting factor called the Kalman gain K, is computed such that it minimizes the mean squared error of the estimate

$$K_k = P_k(-)H_k^T[H_k P_k(-)H_k^T + R_k]^{-1} \tag{7.14}$$

It is the Kalman gain which makes KF stand out in the category of optimal estimation algorithms. As is evident from the equation, K depends upon both the *a priori* covariance $P_k(-)$ and the measurement noise covariance R_k. If the measurements are noisy (when R_k increases) or the process noise is lower (when $P_k(-)$ reduces), then K becomes relatively smaller. When there is more noise in the process (when $P_k(-)$increases) or the measurements are less noisy (when R reduces), then K becomes relatively larger. As we will see, when K is large it assigns more weight to the measurements and when it is small it shows greater faith in the prediction. In the context of INS/GPS integration, K takes relatively larger values when the GPS is more accurate and less noisy. In such a case the measurement covariance matrix becomes relatively small. If Eq. (7.14) is carefully examined, it is evident that small values of R lead to relatively larger values of K.

When a new measurement $\mathbf{z}_k$ is obtained at time t_k, it is compared with the predicted measurement $H_k\hat{\mathbf{x}}_k(-)$ based upon the *a priori* state estimate. Their difference is weighted by K and the prediction of the state vector is updated to generate the best estimate. The estimate of the state at time t_k is therefore

$$\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + K_k[\mathbf{z}_k - H_k\hat{\mathbf{x}}_k(-)] \tag{7.15}$$

where $H_k\hat{\mathbf{x}}_k(-)$ is the predicted observation called $\hat{\mathbf{z}}_k$, and $\mathbf{z}_k - H_k\hat{\mathbf{x}}_k(-)$ is the innovation sequence, a vector of the difference between the actual observation $\mathbf{z}_k$ and the predicted observation $\hat{\mathbf{z}}_k$. The innovation sequence is

$$\mathbf{v}_k = z_k - \hat{z}_k = z_k - H_k\hat{x}_k(-) \tag{7.16}$$

and represents the *amount* of useful information gathered from new measurements whereas K *weights* the useful information for the next update. It can be seen that when K is large (because there is less noise in the measurements or because the process noise is greater) the new information based on the measurements is given more weight. When K is smaller (because either the measurements are noisy or there is less the process noise) the innovation is small and the new information is given little weight.

Based on the value of K, the Kalman filter also updates the uncertainty of its new prediction $\hat{\mathbf{x}}_k(+)$, which is called the *a posteriori* covariance

$$P_k(+) = [I - K_k H_k]P_k(-) \tag{7.17}$$

Equation (7.17) is strongly contested by P. D. Joseph (Bucy and Joseph 2005) to be a *bad* simplification. He argues that even the smallest error in computing K in Eq. (7.14) could result in horrific errors when using Eq. (7.17). This was a real problem in the 1960s and caused serious issues in KF design. Instead he advocates the use of an expanded form of the equation, known as the *Joseph* form

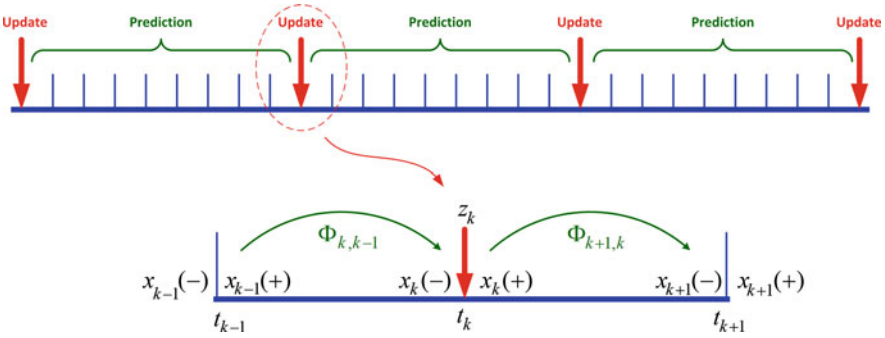


Fig. 7.4 A depiction of typical prediction and update rates

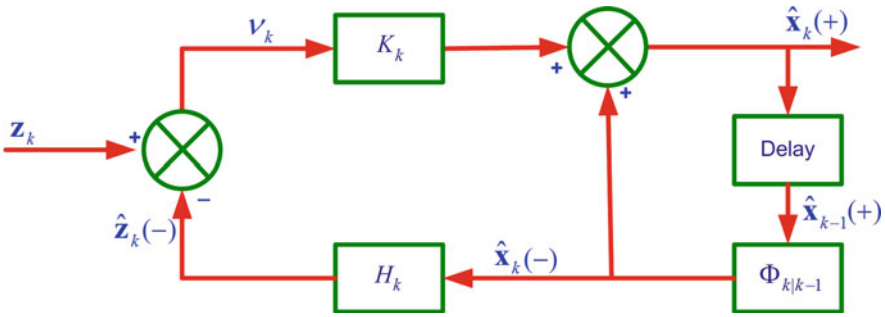


Fig. 7.5 A block diagram showing the information flow in a KF

$$P_k(+) = [I - K_k H_k] P_k(-) [I - K_k H_k]^T + K_k R_k K_k^T \tag{7.18}$$

which is numerically stable and yields correct answers even when the computation of K has an error (e.g. owing to rounding off). It is noteworthy that this form of the a posteriori error covariance $P_k(+)$ helps to avoid divergence by virtue of the assurance of positive semi-definiteness of $P_k(+)$.

In most INS applications the KF update procedure is implemented at a lower rate than the predictions. For example, in a typical application of integrating GPS and INS through KF the prediction maybe carried out at 100 Hz whereas the update may occur at 1 Hz. Figure 7.4 shows the process for typical prediction and update rates.

The specific equations for the time and measurement updates are presented in Tables 7.2 and 7.3 respectively. Figure 7.6 shows the flow diagram of the KF algorithm.

A system level block diagram of the discrete-time KF is shown in Fig. 7.5.

Table 7.2 KF time update (predictor) equations

System dynamic model:	$\mathbf{x}_k = \Phi_{k\|k-1}\mathbf{x}_{k-1} + G_{k-1}\mathbf{w}_{k-1}$ $\mathbf{w}_k \sim N(0, Q_k)$	(7.19)
Predicted state vector:	$\hat{\mathbf{x}}_k(-) = \Phi_{k\|k-1}\mathbf{x}_{k-1}(+)$	(7.20)
Predicted covariance matrix:	$P_k(-) = \Phi_{k\|k-1}P_{k-1}(+)\Phi_{k\|k-1}^T + G_{k-1}Q_{k-1}G_{k-1}^T$	(7.21)

Table 7.3 Discrete-time KF measurement update (corrector) equations

Measurement model:	$z_k = H_k x_k + \boldsymbol{\eta}_k$ $\boldsymbol{\eta}_k \sim N(0, R_k)$	(7.22)
Kalman gain matrix:	$K_k = P_k(-)H_k^T[H_k P_k(-)H_k^T + R_k]^{-1}$	(7.23)
Corrected state estimate:	$\hat{x}_k(+) = \hat{x}_k(-) + K_k[z_k - H_k\hat{x}_k(-)]$	(7.24)
Corrected covariance matrix:	$P_k(+) = [I - K_k H_k]P_k(-)$	(7.25)
Joseph form:	$P_k = (I - K_k H_k)P_k(-)(I - K_k H_k)^T + K_k R_k K_k^T$	(7.26)

7.3 KF Algorithm Steps

A KF essentially comprises five equations which operate in a sequential manner. The flow of the algorithm is depicted in Fig. 7.6, and the steps in the process are explained below.

1. Firstly, the filter is initialized. This requires providing the filter with the initial estimate for its states $\hat{\mathbf{x}}_0$ and the uncertainty in the initial estimate P_0. The estimate of P_0 is based upon knowledge of the approximate accuracy of the initial state estimates and is usually set to relatively high value. We also need to provide the filter with the initial estimates of the system noise covariance matrix Q and measurement noise covariance matrix R. These are estimated on the basis of prior experience with the system and are tuned to get the best estimates of the states.
2. In the first part of the prediction step, the STM Φ is computed and then, using this matrix, the initial state is propagated from the epoch $k - 1$ to k, which is denoted by $\hat{\mathbf{x}}_k(-)$.
3. In the second part of the prediction step the covariance of the predicted state P_k is calculated. This is based on the STM, the previous value of the state covariance P_{k-1}, the last value of the process noise covariance Q_{k-1}, and the noise distribution matrix G_{k-1}. It may be noted that if the process noise is high, P_k will increase and result in a lower confidence in the predicted state $\hat{\mathbf{x}}_k(-)$.
4. In the first step of the update stage the Kalman gain K_k is computed. This depends upon the *a priori* error covariance $P_k(-)$, the process noise covariance R_k, and the design matrix H_k. When $P_k(-)$ is higher the gain is higher, and when R_k is higher the gain is lower, and vice versa of course.
5. In the second part of the prediction stage the estimated (or the *a priori*) state $\hat{\mathbf{x}}_k(-)$ is corrected whenever a measurement is received. This is based on the

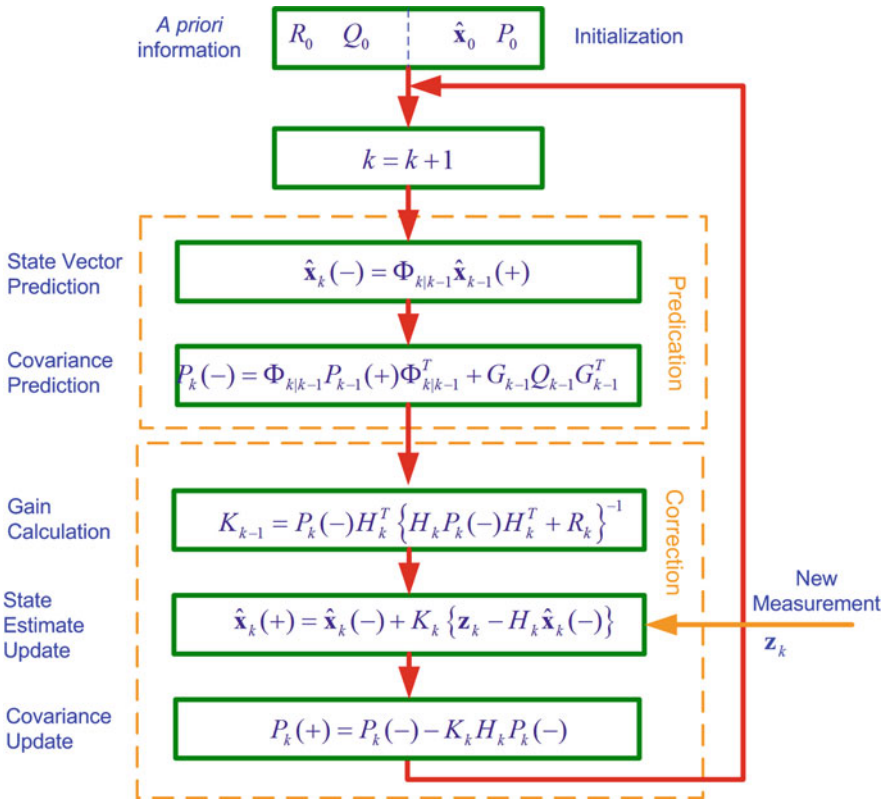


Fig. 7.6 KF algorithm flow diagram

difference of the predicted measurement $H_k\hat{\mathbf{x}}_k(+)$ and the actual measurement $\mathbf{z}_k$. This difference contains the new information that forms the basis for the correction. When K is higher, this difference is weighted more heavily and added to the *a priori* estimate in order to update this to the *a posteriori* estimate $\hat{\mathbf{x}}_k(+)$. But when K is lower, the new information obtained from the measurement is given less weight and the a priori estimate is considered to be relatively accurate.

6. After correcting the state estimate, the KF goes a step further and also updates the *a priori* error covariance $P_k(-)$ to the *a posteriori* error covariance $P_k(+)$ to indicate the level of trust in the corrected estimate $\hat{\mathbf{x}}_k(+)$, which is proportional to gain K_k and $P_k(-)$.
7. Now the KF is ready to go through another loop based on the *a posteriori* estimates that will constitute the *a priori* estimates for the new epoch, as shown in Fig. 7.6.

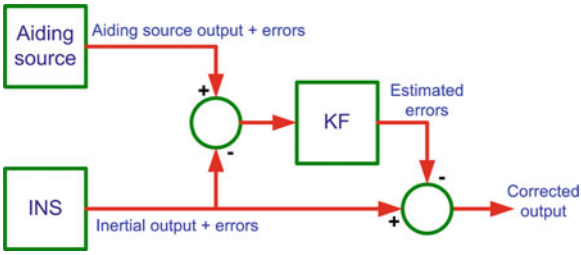


Fig. 7.7 A block diagram of a linearized KF (LKF) which corresponds to an open-loop configuration

7.4 Non-Linear Kalman Filtering

As stated earlier, KF is optimal for linear systems only; however, linear systems are idealized systems which, strictly speaking, do not exist. All systems are non-linear, and even Ohm's law is only an approximation over a limited range and its linearity breaks down above a certain voltage threshold (Simon 2006). However, because many systems are very close to linear, these estimation approaches give satisfactory results. Techniques of linear estimation theory can be applied to non-linear systems by first *linearizing* them. Although KF assumes a linear system model, it has been successfully applied in many situations in which the dynamics were non-linear. In fact the first practical use of KF was for a non-linear system used in Apollo spacecraft. The idea is to linearize the non-linear model, then apply conventional Kalman filtering. When a system is linearized its states represent the deviations from the reference trajectory, and the deviations estimated by KF are subtracted from the reference. Therefore, the non-linear filters operate on the *error states* rather than the *total states* or *whole states*. The error states are formed by taking the difference between the INS states and the aiding source states (such as from GPS). The methods available for linearization are linearized KF (LKF) and extended KF (EKF).

7.4.1 Linearized KF

When Kalman filtering is applied to a system which has been linearized around a nominal trajectory (Maybeck 1982a) this is known as linearized KF (LKF). The nominal trajectory is usually known in advance. For example, the route of a ship or a passenger aircraft is planned in advance and gives the nominal trajectory. Similarly for a satellite orbit, an ellipse or even a circle can be used as a nominal trajectory. For INS/GPS integration the output of the INS is considered to be the nominal trajectory. The linearized KF corresponds to the open-loop configuration in which the filter estimated errors are subtracted from the INS output but they are not fed back to the inertial system (Fig. 7.7).

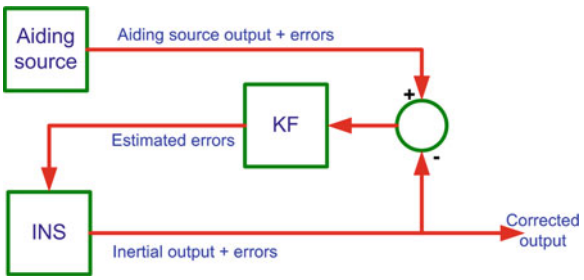


Fig. 7.8 A block diagram of an extended KF (EKF) which corresponds to a closed-loop configuration

7.4.2 Extended KF

Sometimes it is not possible to know the nominal trajectory in advance; therefore the current best estimate of the actual trajectory is used as the nominal trajectory (Grewal and Andrews 2001). When KF is applied to the trajectory linearized using the previous best estimate rather than using a predefined nominal trajectory, it is called Extended KF (EKF). The KF is initialized with the best known estimate and then the estimated value is used as the nominal trajectory for the next epoch, and so on. EKF therefore corresponds to the closed-loop filter configuration in which the estimated errors are fed back to the inertial system to correct its output (Fig. 7.8). This configuration is preferred (especially for low cost and MEMS sensors) so that the INS errors remain small and the linearity assumption required for the KF technique is upheld.

For an in-depth exposition of KF, LKF and EKF the reader is referred to the excellent texts of (Maybeck 1979; Gelb (editor) 1974; Minkler and Minkler 1993; Brown and Hwang 1997; Grewal and Andrews 2008; Simon 2006; Maybeck 1982a, b).

7.5 KF Divergence Control

The process of simplifying the actual system for mathematical representation and linearization gives rise to inadequacies in the system model which can result in the divergence of the KF. This problem was discovered at the very beginning of KF theory, and proposed divergence control techniques include adding fictitious noise to the KF process model, the Schmidt epsilon technique, finite memory filtering and fading memory filtering.

7.5.1 Addition of Fictitious Noise to the KF Process Model

As pointed out in Chap. 5, many simplifications and assumptions must be made in order to make the mathematical model of a real system tractable. Furthermore,

most systems are non-linear and they have to be linearized before we can use KF. This linearization process imposes further approximations on the system model. These inadequacies result in poor KF performance and may lead to divergence. The deterioration in the KF performance can be mitigated by introducing artificial noise into the system by increasing the process noise covariance Q. A similar effect can be achieved by decreasing the measurement noise covariance R. So in reality the Q and R parameters are initialized based on prior experience and are then tuned in order to optimize the performance of the system. For further details, the reader is referred to (Maybeck 1982a; Gelb (editor) 1974).

7.5.2 Schmidt Epsilon Technique

One way to control divergence in KF is to define a threshold for the covariance matrix of the estimation error so that it does not go below a certain value. While in operation the covariance matrix $P_k(-)$ keeps getting smaller, reducing the filter gain K. As is evident from Eq. (7.14), when $P_k(-)$ becomes too small K approaches zero, new measurements are not given sufficient weight, and the filter rejects new measurements and relies on its estimates alone. This is called *Kalman filter incest* (Biezad 1999) and it can be cured by preventing the computed error covariance from falling below a minimum threshold. In covariance *thresholding*, a lower bound is imposed upon the computed error covariance matrix $P_k(-)$ simply by adding to it a fixed quantity denoted by epsilon ε (Schmidt 1968). This value is determined empirically by the designer and can be either an educated guess or a result of trial and error.

The modified Kalman gain equation can therefore be rewritten as

$$K_k = \{P_k(-) + \varepsilon I\} H_k^T \left\{H_k P_k(-) H_k^T + R_k\right\}^{-1} \tag{7.27}$$

and all the other KF equations remain the same.

7.5.3 Finite Memory Filtering

In KF, the latest estimate $\hat{\mathbf{x}}_k(+)$ is based on all the measurements from the start of operations to epoch k. The filter must account for all the previous measurements in order to furnish the best estimate. But sometimes the older data is no longer valid and instead of being of use to the filter it may cause problems. Therefore, in fading memory filters the information from only the most recent measurement data is used and the older information is discarded. For example, the KF could be programmed to use only the last 100 samples of the measurement. The details of this algorithm can be found in (Maybeck 1982a; Gelb (editor) 1974).

7.5.4 Fading Memory Filtering

Sometimes a system model is appropriate for a limited length of time, after which it no longer represents the system adequately. In this case we would rather discard the older measurements gradually, which can be achieved simply by progressively increasing the measurements noise covariance R for that data in order to limit its influence on the latest estimate (Fagin 1964). In mathematical terms (Minkler and Minkler 1993) this is

$$R_k = R\alpha^{-2(k+1)} \tag{7.28}$$

where α is a positive constant. Over time k becomes greater and R_k gradually decreases, thereby giving greater weight to the most recent data.

A similar effect can be achieved (Schlee et al. 1967; Sorenson and Sacks 1971; Simon 2006) by increasing the computed error covariance $P_k(-)$ by a factor of α^2 as follows

$$P_k(-) = \alpha^2 \Phi_{k-1} P_{k-1}(+) \Phi_{k-1}^T + Q_{k-1} \tag{7.29}$$

where the term α can be chosen to be ≥ 1.

In both cases, all the other KF equations remain the same.

7.6 Explanatory Examples

We shall now give some simple examples to help the reader to understand how KF operates.

7.6.1 A Simple Navigation Example

Consider a body moving in a straight line with constant acceleration. The time rate of change of the velocity error is equal to the accelerometer bias error. In addition, the time rate of change of the position error δp will be equal to the velocity error. This can be expressed mathematically as

$$\begin{aligned} \delta\dot{p} &= \delta v \\ \delta\dot{v} &= \delta f \end{aligned} \tag{7.30}$$

Moreover, the accelerometer errors are modelled using the first-order Gauss Markov process

$$\delta\dot{f} = -\alpha\delta f + \sqrt{2\alpha\sigma^2}w(t) \tag{7.31}$$

The system equation in its matrix form is

$$\underbrace{\begin{bmatrix}\delta\dot{p}\\ \delta\dot{v}\\ \delta\dot{f}\end{bmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & -\alpha\end{bmatrix}}_{F}\underbrace{\begin{bmatrix}\delta p\\ \delta v\\ \delta f\end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix}0\\ 0\\ \sqrt{2\alpha\sigma^2}\end{bmatrix}}_{G}\underbrace{w(t)}_{w} \tag{7.32}$$

Let us now assume that GPS position updates are continuously available. The measurement (update) equation is obtained by comparing the output of the aiding source (GPS) to the INS output. The observation $\boldsymbol{z}$ supplied to the Kalman filter is therefore

$$\boldsymbol{z} = p_{GPS} - (p_{INS} + \delta\eta_p) = \delta p + \delta\eta_p \tag{7.33}$$

where $\delta\eta_p$ is the GPS measurement noise. Hence $\boldsymbol{z}$ is related to the error state vector $\boldsymbol{x}$ as follows

$$\boldsymbol{z} = \underbrace{\begin{bmatrix}1 & 0 & 0\end{bmatrix}}_{H}\underbrace{\begin{bmatrix}\delta p\\ \delta v\\ \delta f\end{bmatrix}}_{\mathrm{x}} + \underbrace{\delta\eta_p}_{\eta} \tag{7.34}$$

If we have velocity measurement updates instead of position updates, then

$$\boldsymbol{z} = v_{GPS} - (v_{INS} + \delta\eta_v) = \delta v + \delta\eta_v \tag{7.35}$$

where $\delta\eta_v$ is the GPS measurement noise. In this case z is related to the error state vector $\boldsymbol{x}$ as follows

$$\boldsymbol{z} = \underbrace{\begin{bmatrix}0 & 1 & 0\end{bmatrix}}_{H}\underbrace{\begin{bmatrix}\delta p\\ \delta v\\ \delta f\end{bmatrix}}_{\mathrm{x}} + \underbrace{\delta\eta_v}_{\eta} \tag{7.36}$$

If we have both velocity and position updates, then

$$\boldsymbol{z} = \underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}}_{H}\underbrace{\begin{bmatrix}\delta p\\ \delta v\\ \delta f\end{bmatrix}}_{\mathrm{x}} + \underbrace{\begin{bmatrix}\delta\eta_p\\ \delta\eta_v\end{bmatrix}}_{\boldsymbol{\eta}} \tag{7.37}$$

7.6.2 Zero Velocity Update

In applications where there are no continuous update measurements available or when the measurements are temporary denied, the surveying accuracy may be put

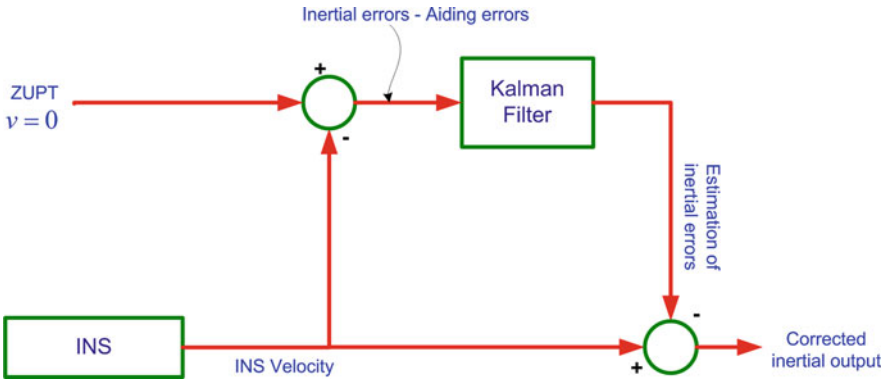


Fig. 7.9 The operation of the KF during the ZUPT

at risk by a potentially unlimited increase in the velocity errors and their direct contributions to both the attitude and position errors. ZUPTs provide a dramatic improvement in the accuracy of inertial surveying as compared to the standalone systems.

The procedure to carry out the ZUPT is as follows

1. Stop the vehicle every 2–4 min for 30–60 s.
2. Obtain between 30 and 40 velocity measurements while halted.
3. Use these measurements as KF updates or for velocity curve fitting.

The advantage of ZUPT is that because no additional equipment is needed, it is simple and inexpensive to perform. Also, there are no synchronization problems between the predictions and measurements. However, its major limitation is that the vehicle must be readily halted, which is not always possible (e.g. for airborne and sea vehicles).

7.6.2.1 Improving Surveying Accuracy by ZUPT

An optimal estimation procedure like KF should be used to restore the behavior of the INS velocity errors between ZUPT stations. At such a station the differences between the INS output velocities and those of the ZUPT along the three mutually orthogonal direction are fed to the Kalman filter (Fig. 7.9).

The system equation is independent of the type of update measurement, and is given as

$$\dot{\mathbf{x}} = F\mathbf{x} + G\mathbf{w} \tag{7.38}$$

where $\boldsymbol{x}$ is the state vector consisting of fifteen states as follows

$$\mathbf{x} = \Big[\underbrace{\delta\phi, \delta\lambda, \delta h}_{\text{Position errors}}, \underbrace{\delta v_e, \delta v_n, \delta v_u}_{\text{Velocity errors}}, \underbrace{\delta p, \delta r, \delta A}_{\text{Attitude errors}}, \underbrace{\delta\omega_x, \delta\omega_y, \delta\omega_z}_{\text{Gyro errors}}, \underbrace{\delta f_x, \delta f_y, \delta f_z}_{\text{Accelerometer errors}}\Big]^T \tag{7.39}$$

The design matrix H of the measurement update equation depends upon the type of measurement available, and is given as

$$\mathbf{z} = H\mathbf{x} + \delta\boldsymbol{\eta}_p \tag{7.40}$$

With ZUPT measurements the design matrix will be

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{7.41}$$

and therefore the measurement update equation is

$$\underbrace{\begin{bmatrix} v_{e,0} - v_{e,INS} \\ v_{n,0} - v_{n,INS} \\ v_{u,0} - v_{u,INS} \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{H} \mathbf{x} + \underbrace{\begin{bmatrix} \delta\eta_{ve} \\ \delta\eta_{vn} \\ \delta\eta_{vu} \end{bmatrix}}_{\boldsymbol{\eta}} \tag{7.42}$$

where

$v_{e,INS}, v_{n,INS}, v_{u,INS}$	are the INS output velocities
$v_{e,0}, v_{n,0}, v_{u,0}$	are the zero update velocities, which can be modeled either as zeros or as very small values corresponding to any vibrations which might exist in the system
H	is the design matrix
$\mathbf{x}$	is the INS error state vector given in Eq. (7.39)
$\delta\eta_{ve}, \delta\eta_{vn}, \delta\eta_{vu}$	are the uncertainties around the zero velocity

Figure 7.10 illustrate the trend of the velocity and the position in the periods during and between the ZUPTs. It is evident that the ZUPTs prevent the error in the position from growing unhindered. It can be appreciated that in the absence of ZUPTs the position error would grow very rapidly and make the position estimate totally unreliable.

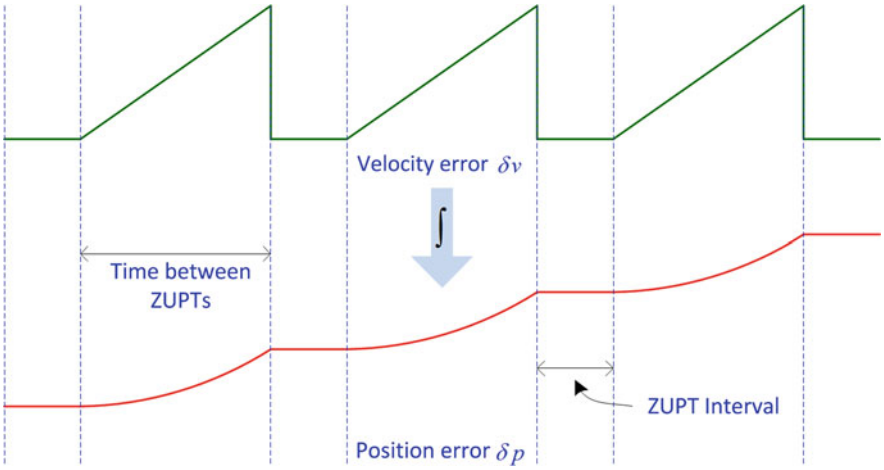


Fig. 7.10 The velocity and the position between and during ZUPTs

The performance of the ZUPT is affected by two factors

1. The duration between two neighboring ZUPTs, during which the velocity errors grow with time. The choice of the time duration between ZUPTs is related to the accuracy and the performance of the IMU employed.
2. The ZUPT time interval should be chosen to enable the KF algorithm to converge and restore the INS errors.

The significance of ZUPT can be assessed by the following factors

1. It limits the growth of velocity errors and can even reset them to zero.
2. It facilitates estimation of the accelerometer bias errors.
3. It facilitates estimation of the misalignment of the three attitude angles.

In general the ZUPT procedure is effective in limiting the long term growth of surveying errors.

7.6.3 Coordinate Update

In some applications the coordinate update can be provided either continuously or at CUPT stations whose coordinates are accurately predefined. If the navigation system were left working as a standalone INS without updates, we would expect significant long term error growth, even when using highly accurate navigational grade inertial sensors. With CUPT, the motion of the body is interrupted to permit the output of the INS to be compared with the coordinate measurements supplied by the aiding source.

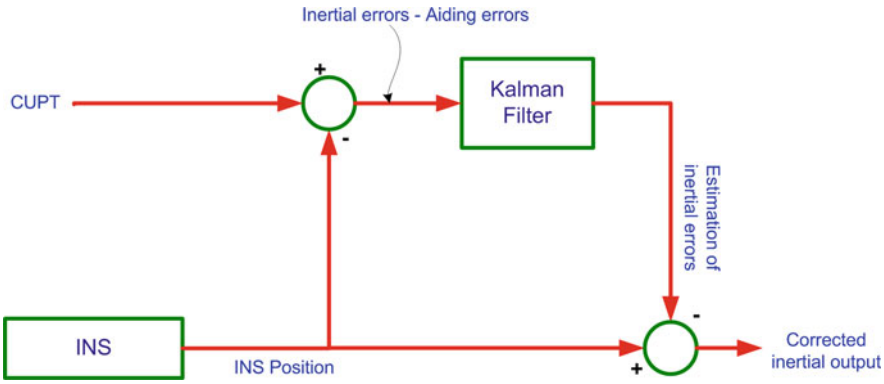


Fig. 7.11 The operation of the KF during CUPT

7.6.3.1 Improving Surveying Accuracy by CUPT

This procedure is similar to ZUPT except that instead of the *velocities* the update is for the *position coordinates*. An optimal estimation procedure like KF is used to restore the behavior of the INS position errors between CUPT stations. At each station the differences between the INS position outputs and of the CUPT are fed to the Kalman filter (Fig. 7.11).

As in the case of ZUPT, the system equation is independent of the type of update measurement and is given as

$$\dot{\mathbf{x}} = F\mathbf{x} + G\mathbf{w} \tag{7.43}$$

where $\boldsymbol{x}$ is the state vector consisting of 15 states are follows

$$\mathbf{x} = \Big[\underbrace{\delta\phi, \delta\lambda, \delta h}_{\text{Position errors}}, \underbrace{\delta v_e, \delta v_n, \delta v_u}_{\text{Velocity errors}}, \underbrace{\delta p, \delta r, \delta A}_{\text{Attitude errors}}, \underbrace{\delta\omega_x, \delta\omega_y, \delta\omega_z}_{\text{Gyro errors}}, \underbrace{\delta f_x, \delta f_y, \delta f_z}_{\text{Accelerometer errors}}\Big]^T \tag{7.44}$$

The design matrix H of the measurement update equation depends upon the type of measurement available, and is given as

$$\mathbf{z} = H\mathbf{x} + \delta\boldsymbol{\eta}_p \tag{7.45}$$

With CUPT measurements the design matrix will be

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{7.46}$$

and therefore the measurement update equation is

$$\underbrace{\begin{bmatrix} \varphi_{GPS} - \varphi_{INS} \\ \lambda_{GPS} - \lambda_{INS} \\ h_{GPS} - h_{INS} \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{H} \mathbf{x} + \underbrace{\begin{bmatrix} \delta\eta_{\varphi} \\ \delta\eta_{\lambda} \\ \delta\eta_{h} \end{bmatrix}}_{\boldsymbol{\eta}} \tag{7.47}$$

where

$\varphi_{INS}, \lambda_{INS}, h_{INS}$	are the INS output position coordinates
$\varphi_{GPS}, \lambda_{GPS}, h_{GPS}$	are the GPS output position coordinates
H	is the design matrix
$\mathbf{x}$	is the INS error state vector given in Eq. (7.44)
$\delta\eta_{\varphi}, \delta\eta_{\lambda}, \delta\eta_{h}$	are the uncertainties position measurements.

References

Biezad DJ (1999) Integrated navigation and guidance systems. The American Institute of Aeronautics and Astronautics, Reston. ISBN:1563472910978156347291097816347291

Brown RG, Hwang PYC (1997) Introduction to random signals and applied kalman filtering : with MATLAB exercises and solutions, 3rd edn. Wiley, New York

Bucy RS, Joseph PD (2005) Filtering for stochastic processes with applications to guidance. American Mathematical Society

Fagin SL (1964) Recursive linear regression theory and error analysis of optimal systems. IEEE Int Conv Rec 2:216–240

Gelb (editor) A (1974) Applied optimal estimation. M.I.T. Press, Cambridge, Mass

Grewal MS, Andrews AP (2001) Kalman filtering: theory and practice using MATLAB, 2nd edn. Wiley, New York

Grewal MS, Andrews AP (2008) Kalman filtering : theory and practice using MATLAB®. Wiley, Oxford

Maybeck PS (1979) Stochastic models, estimation, and control, vol 1. Academic, New York

Maybeck PS (1982a) Stochastic models, estimation, and control, vol 2. New York, Academic

Maybeck PS (1982b) Stochastic models, estimation, and control, vol 3. New York, Academic

Minkler G, Minkler J (1993) Theory and application of kalman filtering. Magellan Book Co., Palm Bay, FL

Schlee FH, Standish CJ, Toda NF (1967) Divergence in the kalman filter. AIAA J 5(6): 1114–1120

Schmidt SF (1968) Case study of kalman filtering in the C-5 aircraft navigation system. Joint Autom, Control Conference, Ann Arbor, MI

Simon D (2006) Optimal state estimation : Kalman, H infinity, and nonlinear approaches. Wiley, Hoboken

Sorenson HW, Sacks JE (1971) Recursive fading memory filtering. Inf Sci 3(2):101–119

Welch G, Bishop G (2001) SIGGRAPH course 8: an introduction to the Kalman filter. University of North Carolina at Chapel Hill, Chapel Hill

Chapter 8
INS/GPS Integration

There are contrasting pros and cons to INS and GPS. An INS is a self-contained autonomous navigation system that provides a bandwidth exceeding 200 Hz. It has good short term accuracy and provides attitude information in addition to position and velocity. But long term errors grow without bound as the inertial sensor errors accumulate due to intrinsic integration in the navigation algorithm. In contrast to an INS, GPS has good long term accuracy with errors limited to a few meters and user hardware costing as little as $100. But it has poor short term accuracy and a lower output data rate. A standard GPS receiver usually does not provide attitude information, but it can with extra hardware and software. GPS needs a direct line of sight to at least four satellites, which is not always possible because the signals from satellites suffer from obstruction by tall buildings, trees, tunnels, degradation through the atmosphere and multipath interference.

Capitalizing on the complementary characteristics of these two systems, their synergistic integration overcomes their individual drawbacks and provides a more accurate and robust navigation solution than either could achieve on its own. The integrated navigation solution is a continuous high data rate system that provides a full navigation solution (position, velocity and attitude) with improved accuracy in both the short and long term. Optimal estimation techniques, predominantly based on Kalman filtering, are employed to optimally fuse the GPS and INS positioning and navigation information to yield a reliable navigation solution. GPS prevents the inertial solution from drifting and INS provides continuity in the navigational solution, attitude information, and bridges GPS signal outages. A typical INS/GPS integration is depicted in Fig. 8.1.

The estimator compares the outputs of the INS and GPS and estimates errors in inertial position, velocity and attitudes, plus some other parameters. Traditionally

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_8,

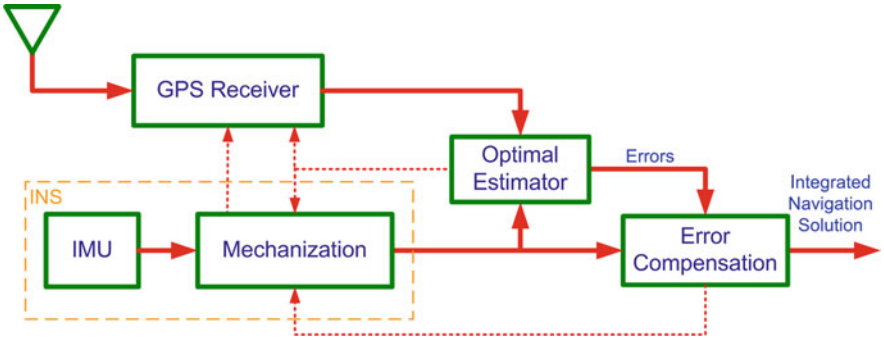


Fig. 8.1 An overview of a typical INS/GPS integration

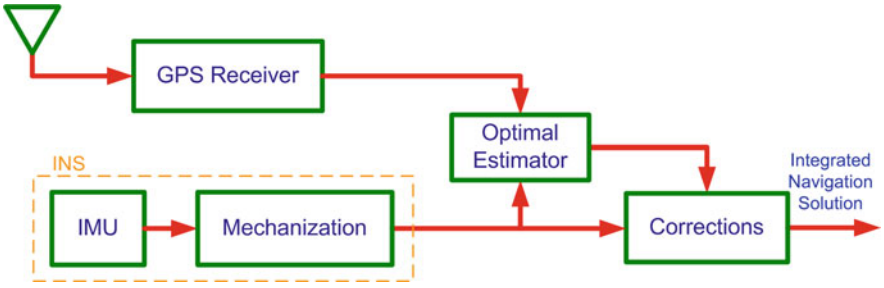


Fig. 8.2 An open-loop implementation of INS/GPS integration

the estimator is a KF or a variant such as LKF and EKF, but other filters (e.g. a particle filter) or AI techniques are available. In Fig. 8.1 the inertial output is corrected using the estimated errors to produce the integrated navigation solution. Dotted lines in the figure depict the optional paths, the presence of which depends upon the specific type of integrations scheme, as will be discussed later.

8.1 Error Feedback Schemes

Two types of error feedback mechanisms are used in INS/GPS integration based on error-state KF: open-loop and closed-loop. As pointed out in Chap. 7, these architectures correspond to LKF and EKF respectively.

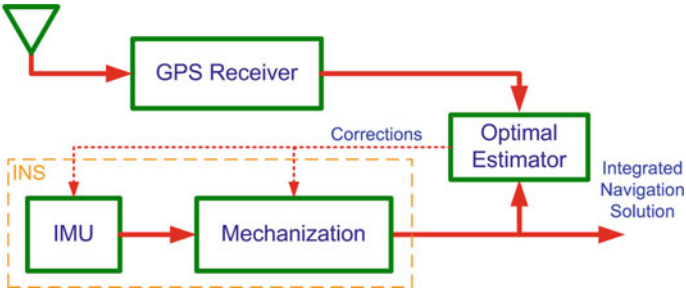


Fig. 8.3 A closed-loop implementation of INS/GPS integration

8.1.1 Open-Loop INS/GPS Architecture

In the open-loop configuration the correction of position, velocity and attitude are performed external to the INS, where the estimated errors are subtracted from the INS solution at each iteration. The errors, or the corrected state, are not fed back to the INS. The advantage of this configuration is that, in addition to the integrated navigation solution, the raw INS solution can support integrity monitoring and continuing service in the event of a problem with the Kalman filter (Minkler and Minkler 1993; Groves 2008). However, due to unhindered INS drift, the errors in the INS grow larger with time, making the linearity assumption invalid. Hence the open-loop configuration is more prone to KF performance issues (Maybeck 1979). A schematic of the open-loop INS/GPS integration is shown in Fig. 8.2.

8.1.2 Closed-Loop INS/GPS Architecture

In the closed-loop configuration the error estimates from KF are fed back in order to correct the INS (Minkler and Minkler 1993). The output of the INS forms the integrated solution. KF position, velocity and attitude estimates are reset to zero after the error estimates are fed back. It may be observed from Fig. 8.3 that in closed-loop integration, KF estimated accelerometer and gyroscope errors are also fed back to correct the IMU measurements before they are used in mechanization. These errors are applied on every iteration of mechanization, with feedback from KF periodically updating the accelerometer and gyroscope errors (Groves 2008; Maybeck 1979).

8.2 Types of Integration

Capitalizing on the complementary qualities of GPS and INS, various INS/GPS integration architectures have been proposed (Wagner and Wieneke 2003; Gebre-Egziabher 2007) to achieve the maximum advantage depending on the

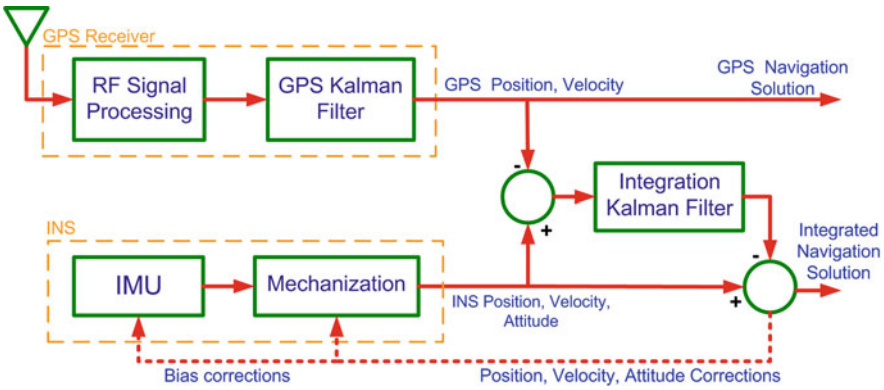


Fig. 8.4 A block diagram of a loosely coupled INS/GPS integration

application and the requirements of simplicity versus robustness; specifically loosely coupled, tightly coupled and ultra-tightly (or deeply) coupled.

8.2.1 Loosely Coupled INS/GPS Integration

In this integration, the GPS and INS operate independently and provide separate navigation solutions. To improve the solution the position and/or velocity from GPS is fed to some optimal estimator, usually a Kalman filter. The INS solution is also supplied to the filter which takes the difference between the two and, based upon the error models, estimates the INS errors. The INS solution is corrected for these errors to produce the integrated navigation solution in the form of position, velocity and attitude. The distinguishing feature of this configuration is a separate filter for the GPS. This integration is an example of cascaded integration because the two filters (the GPS and the integration filter) are used in succession. It is also called a *decentralized* approach because there is a separate filter used for GPS. A block diagram of the loosely coupled INS/GPS integration is shown in Fig. 8.4.

Loosely coupled integration is simple to implement and is robust. It provides three separate navigation solutions for open-loop (raw INS, raw GPS, and the integrated solution) and two for close-loop (no raw INS). The main problem is its inability to provide GPS-aiding when the effective number of satellites falls below the minimum. And because the outputs of the GPS KF are time correlated, the KF assumption of uncorrelated measurement noise is jeopardized, thereby impairing the system performance.

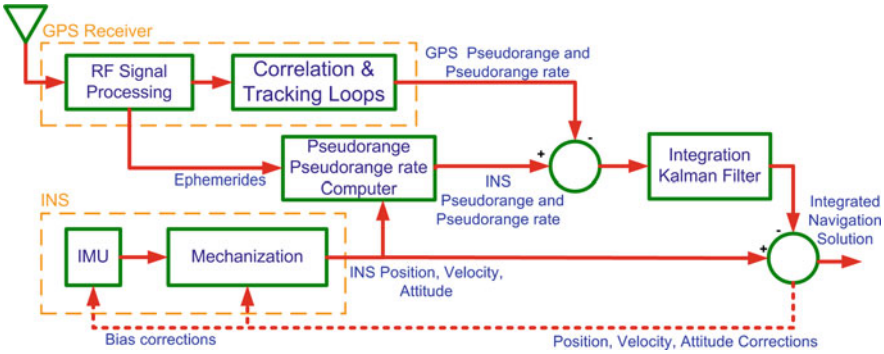


Fig. 8.5 A block diagram of a tightly coupled integration of INS/GPS

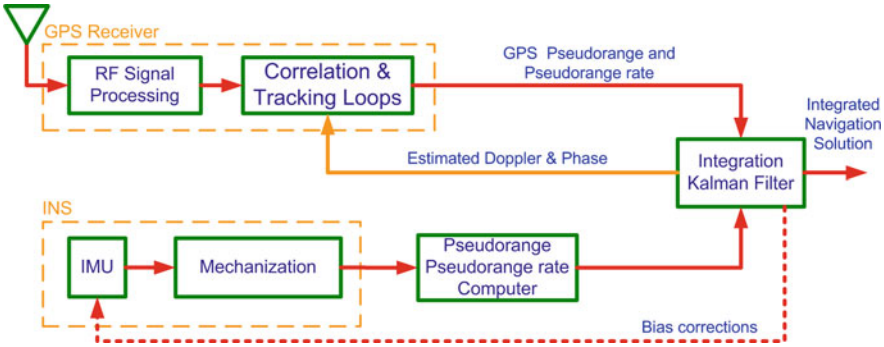


Fig. 8.6 A block diagram of an ultra-tight integration of GPS/INS

8.2.2 Tightly Coupled INS/GPS Integration

This architecture is also called *centralized* integration owing to the use of a single common master filter. The difference between the pseudo-range and pseudo-range rate measurements from the GPS and corresponding values predicted by the INS are fed to the KF to estimate the errors in the INS. The output from the INS is then corrected for these errors to obtain the integrated navigation solution. It is possible to use either pseudo-range or pseudo-range rate measurements, but often both are used because they are complementary (Groves 2008). This architecture is shown in Fig. 8.5. The tightly coupled integration eliminates the problem of correlated measurements that arises due to cascaded Kalman filtering in the loosely coupled approach. Furthermore, this integration can provide a GPS update even if fewer than four satellites are visible. This is useful because a typical real-life trajectory includes urban environments as well as thick forest canopies and steep hills where the number of visible satellites can drop to fewer than four. On the other hand the tightly coupled

approach is more complex to implement, as the algorithm involves processing raw GPS data. The other limitation is that there is no standalone GPS solution. Using the same hardware, a tightly coupled INS/GPS integration almost always performs better than a loosely coupled integration in term of accuracy and robustness.

8.2.3 Ultra-Tight INS/GPS Integration

This type of integration, also called deep integration, is shown in Fig. 8.6. There are two major differences (Gebre-Egziabher 2007) between this architecture and those outlined above. Firstly, there is a basic change in the architecture of the GPS receiver to provide a different implementation of the tracking loops. Secondly, the information from the INS is used as an integral part of the GPS receiver; hence the INS and GPS are no longer independent navigators. This scheme usually requires access to the internal GPS hardware and is complex to implement. The advantages of this scheme are that it lowers the tracking bandwidth and is more resistant to jamming. This type of integration can work under a lower signal to noise ratio and provides a solution in scenarios in which the number of GPS satellites falls below four. As seen from Fig. 8.6, the defining characteristic of this type of integration is the feedback path from the estimator which provides the estimated Doppler and carrier phase data to the GPS structure to enable its phase-lock loops (PLL) and delay-lock loops (DLL) to provide better accuracy, robustness and overall system improvement (Alban et al. 2003).

8.3 Dynamic Error Model of INS Equations

An integrated GPS/INS requires an optimal estimator to perform the data fusion and estimation. Various kinds of estimators can be used depending on the needs of the system. As KF has traditionally been used for the integration of navigational sensors we will look at this in greater detail. Owing to the advantages explained in Chap. 5, the local-level frame is commonly chosen for the implementation of this integration for navigation applications. Chapter 6 explained that INS errors arise due to initial condition errors, inertial sensor errors and computational errors. As a result, error models are required for analysis and estimation of the various error sources associated with an INS. These error models will be summarized here prior to moving on to discuss the full KF model for both loosely coupled and tightly coupled INS/GPS integration schemes.

As shown in Chap. 6, the error state vector for the local-level frame mechanization equations consists of the errors along the curvilinear geodetic coordinates (latitude error $\delta\varphi$, longitude error $\delta\lambda$ and altitude error δh), the errors along the Earth-referenced velocities (east-velocity error δv_e, north-velocity error δv_n and up-velocity error δv_u) and the error in the three attitude angles (pitch error δp, roll

error δr and azimuth error δA). It also includes the accelerometer biases and gyro drifts. The complete error states vector is therefore

$$\mathbf{x}^l_{15\times 1} = \left[\delta\mathbf{r}^l_{3\times 1},\ \delta\mathbf{v}^l_{3\times 1},\ \boldsymbol{\varepsilon}^l_{3\times 1},\ \delta\boldsymbol{\omega}_{3\times 1},\ \delta\mathbf{f}_{3\times 1}\right]^T \tag{8.1}$$

where

$\delta\mathbf{r}^l = [\delta\varphi,\ \delta\lambda,\ \delta h]^T$	is the position error vector
$\delta\mathbf{v}^l = [\delta v_e,\ \delta v_n,\ \delta v_u]^T$	is the Earth-referenced velocity error vector
$\boldsymbol{\varepsilon}^l = [\delta p,\ \delta r,\ \delta A]^T$	is the attitude error vector
$\delta\boldsymbol{\omega} = \left[\delta\omega_x, \delta\omega_y, \delta\omega_z\right]^T$	is the gyroscope error vector (consisting of drifts)
$\delta\mathbf{f} = \left[\delta f_x,\ \delta f_y,\ \delta f_z\right]^T$	is accelerometer error vector (consisting of biases)

The linearized error models for these errors are summarized below.

Coordinate errorsin the l-frame

$$\underbrace{\begin{bmatrix}\delta\dot{\varphi}\\ \delta\dot{\lambda}\\ \delta\dot{h}\end{bmatrix}}_{\delta\dot{\mathbf{r}}^l} = \underbrace{\begin{bmatrix}0 & \frac{1}{(R_M+h)} & 0\\ \frac{1}{(R_N+h)\cos\varphi} & 0 & 0\\ 0 & 0 & 1\end{bmatrix}}_{F_r}\underbrace{\begin{bmatrix}\delta v_e\\ \delta v_n\\ \delta v_u\end{bmatrix}}_{\delta\mathbf{v}^l} \tag{8.2}$$

Velocity errorsin the l-frame

$$\underbrace{\begin{bmatrix}\delta\dot{v}_e\\ \delta\dot{v}_n\\ \delta\dot{v}_u\end{bmatrix}}_{\delta\dot{\mathbf{v}}} = \underbrace{\begin{bmatrix}0 & f_u & -f_n\\ -f_u & 0 & f_e\\ f_n & -f_e & 0\end{bmatrix}}_{F_v}\underbrace{\begin{bmatrix}\delta p\\ \delta r\\ \delta A\end{bmatrix}}_{\boldsymbol{\varepsilon}} + \underbrace{\begin{bmatrix}R_{11} & R_{12} & R_{13}\\ R_{21} & R_{22} & R_{23}\\ R_{31} & R_{32} & R_{33}\end{bmatrix}}_{R^l_b}\underbrace{\begin{bmatrix}\delta f_x\\ \delta f_y\\ \delta f_z\end{bmatrix}}_{\delta\mathbf{f}} \tag{8.3}$$

where f_e, f_n and f_u are the body accelerations transformed into the local-level frame.

Attitude errorsin the l-frame

$$\dot{\boldsymbol{\varepsilon}}^l = \underbrace{\begin{bmatrix} \delta\dot{p} \\ \delta\dot{r} \\ \delta\dot{A} \end{bmatrix}}_{\dot{\varepsilon}^l} = \underbrace{\begin{bmatrix} 0 & \dfrac{1}{(R_M + h)} & 0 \\ \dfrac{-1}{(R_N + h)} & 0 & 0 \\ \dfrac{-\tan\varphi}{(R_N + h)} & 0 & 0 \end{bmatrix}}_{F_\varepsilon} \underbrace{\begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix}}_{\delta \mathbf{v}^l} + \underbrace{\begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}}_{R_b^l} \underbrace{\begin{bmatrix} \delta\omega_x \\ \delta\omega_y \\ \delta\omega_z \end{bmatrix}}_{\delta\boldsymbol{\omega}} \tag{8.4}$$

Accelerometer bias errors

For the stochastic biases of the accelerometers, the time rate of change of the bias errors can be expressed as

$$\underbrace{\begin{bmatrix} \delta\dot{f}_x \\ \delta\dot{f}_y \\ \delta\dot{f}_z \end{bmatrix}}_{\delta\dot{\mathbf{f}}^b} = \underbrace{\begin{bmatrix} -\beta_{fx} & 0 & 0 \\ 0 & -\beta_{fy} & 0 \\ 0 & 0 & -\beta_{fz} \end{bmatrix}}_{F_f} \underbrace{\begin{bmatrix} \delta f_x \\ \delta f_y \\ \delta f_z \end{bmatrix}}_{\delta\mathbf{f}^b} + \underbrace{\begin{bmatrix} \sqrt{2\beta_{fx}\sigma_{fx}^2} \\ \sqrt{2\beta_{fy}\sigma_{fy}^2} \\ \sqrt{2\beta_{fz}\sigma_{fz}^2} \end{bmatrix}}_{\boldsymbol{\sigma}_f} w(t) \tag{8.5}$$

where

$\beta_{fx},\ \beta_{fy},\ \beta_{fz}$	are the reciprocals of the correlation times associated with the autocorrelation sequence of δf_x, δf_y and δf_z
$\sigma_{fx}^2,\ \sigma_{fy}^2,\ \sigma_{fz}^2$	are the variances associated with the accelerometer errors
$w(t)$	is the unit-variance white Gaussian noise.

Gyroscope drift errors

Similarly for the stochastic biases of the gyroscopes, the time rate of change of the bias errors can be written as

$$\delta\dot{\boldsymbol{\omega}} = \underbrace{\begin{bmatrix} \delta\dot{\omega}_x \\ \delta\dot{\omega}_y \\ \delta\dot{\omega}_z \end{bmatrix}}_{\delta\dot{\boldsymbol{\omega}}} = \underbrace{\begin{bmatrix} -\beta_{\omega x} & 0 & 0 \\ 0 & -\beta_{\omega y} & 0 \\ 0 & 0 & -\beta_{\omega z} \end{bmatrix}}_{F_\omega} \underbrace{\begin{bmatrix} \delta\omega_x \\ \delta\omega_y \\ \delta\omega_z \end{bmatrix}}_{\delta\boldsymbol{\omega}} + \underbrace{\begin{bmatrix} \sqrt{2\beta_{\omega x}\sigma_{\omega x}^2} \\ \sqrt{2\beta_{\omega y}\sigma_{\omega y}^2} \\ \sqrt{2\beta_{\omega z}\sigma_{\omega z}^2} \end{bmatrix}}_{\sigma_\omega} w(t) \tag{8.6}$$

where

$\beta_{\omega x},\ \beta_{\omega y},\ \beta_{\omega z}$	are the reciprocals of the correlation times associated with the autocorrelation sequence of $\delta\omega_x$, $\delta\omega_y$ and $\delta\omega_z$
$\sigma_{\omega x}^2,\ \sigma_{\omega y}^2,\ \sigma_{\omega z}^2$	are the variances associated with the gyroscope errors
$w(t)$	is the unit-variance white Gaussian noise.

8.4 Models for Loosely Coupled INS/GPS Integration

8.4.1 System Model

The system model of continuous KF for loosely coupled integration is given by

$$\delta\dot{\mathbf{x}} = F\delta\mathbf{x} + Gw \tag{8.7}$$

The state vector includes error components of position, velocity and attitude as well as accelerometer biases and gyroscope drifts

$$\delta\mathbf{x}^l_{15\times1} = \left[\delta\mathrm{r}^l_{3\times1},\ \delta\mathbf{v}^l_{3\times1},\ \boldsymbol{\varepsilon}^l_{3\times1},\ \delta\boldsymbol{\omega}_{3\times1},\ \delta\mathrm{f}_{3\times1}\right]^T \tag{8.8}$$

where

$\delta\mathbf{r}^l = [\delta\varphi,\ \delta\lambda,\ \delta h]^T$	is the position error vector
$\delta\mathbf{v}^l = [\delta v_e,\ \delta v_n,\ \delta v_u]^T$	is the Earth-referenced velocity error vector
$\boldsymbol{\varepsilon}^l = [\delta p, \delta r, \delta A]^T$	is the attitude error vector
$\delta\boldsymbol{\omega} = \left[\delta\omega_x, \delta\omega_y, \delta\omega_z\right]^T$	is the gyroscope error vector (consisting of drifts)
$\delta\mathbf{f} = \left[\delta f_x, \delta f_y, \delta f_z\right]^T$	is accelerometer error vector (consisting of biases)

w is the unit-variance white Gaussian noise.

The term G in Eq. (8.7) is the noise distribution vector, which includes the variances associated with the state vector

$$G = \left[\boldsymbol{\sigma}_{r,1\times3},\ \boldsymbol{\sigma}_{v,1\times3},\ \boldsymbol{\sigma}_{\varepsilon,1\times3},\ \boldsymbol{\sigma}_{\omega,1\times3},\ \boldsymbol{\sigma}_{f,1\times3}\right] \tag{8.9}$$

The term F is the dynamic coefficient matrix which, as previously explained, contains the INS error models for the position, velocity, attitude and the inertial sensors. According to the expressions for the errors in Eq. (8.2) through Eq. (8.6), we can write F matrix as

$$F = \begin{bmatrix} 0_{3\times3} & F_r & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & F_v & 0_{3\times3} & R^l_b \\ 0_{3\times3} & F_\varepsilon & 0_{3\times3} & R^l_b & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & F_\omega & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & F_f \end{bmatrix} \tag{8.10}$$

where various sub-matrices were defined in Sect. 8.3.

The system model for the loosely coupled INS/GPS integration can therefore be written as

$$\begin{bmatrix} \delta\dot{\mathbf{r}}^l_{3\times1} \\ \delta\dot{\mathbf{v}}^l_{3\times1} \\ \dot{\boldsymbol{\varepsilon}}^l_{3\times1} \\ \delta\dot{\boldsymbol{\omega}}_{3\times1} \\ \delta\dot{\mathbf{f}}_{3\times1} \end{bmatrix} = \begin{bmatrix} 0_{3\times3} & F_r & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & F_v & 0_{3\times3} & R^l_b \\ 0_{3\times3} & F_\varepsilon & 0_{3\times3} & R^l_b & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & F_\omega & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & F_f \end{bmatrix} \begin{bmatrix} \delta\mathbf{r}^l_{3\times1} \\ \delta\mathbf{v}^l_{3\times1} \\ \boldsymbol{\varepsilon}^l_{3\times1} \\ \delta\boldsymbol{\omega}_{3\times1} \\ \delta\mathbf{f}_{3\times1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\sigma}_{r,1\times3} \\ \boldsymbol{\sigma}_{v,1\times3} \\ \boldsymbol{\sigma}_{\varepsilon,1\times3} \\ \boldsymbol{\sigma}_{\omega,1\times3} \\ \boldsymbol{\sigma}_{f,1\times3} \end{bmatrix} w \tag{8.11}$$

The discrete form of Eq. (8.7) can be written as

$$\delta\mathbf{x}_k = (I + F\Delta t)\delta\mathbf{x}_{k-1} + G\Delta t w_{k-1} \tag{8.12}$$

and Eq. (8.11) can be written in the following discrete form

$$\begin{bmatrix} \delta\mathbf{r}^l_{3\times1} \\ \delta\mathbf{v}^l_{3\times1} \\ \boldsymbol{\varepsilon}^l_{3\times1} \\ \delta\boldsymbol{\omega}_{3\times1} \\ \delta\mathbf{f}_{3\times1} \end{bmatrix} = \begin{bmatrix} I_{3\times3} & F_r\Delta t & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & I_{3\times3} & F_v\Delta t & 0_{3\times3} & R^l_b\Delta t \\ 0_{3\times3} & F_\varepsilon\Delta t & I_{3\times3} & R^l_b\Delta t & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} + F_\omega\Delta t & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} + F_f\Delta t \end{bmatrix} \begin{bmatrix} \delta\mathbf{r}^l_{3\times1} \\ \delta\mathbf{v}^l_{3\times1} \\ \boldsymbol{\varepsilon}^l_{3\times1} \\ \delta\boldsymbol{\omega}_{3\times1} \\ \delta\mathrm{f}_{3\times1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\sigma}_{r,1\times3} \\ \boldsymbol{\sigma}_{v,1\times3} \\ \boldsymbol{\sigma}_{\varepsilon,1\times3} \\ \boldsymbol{\sigma}_{\omega,1\times3} \\ \boldsymbol{\sigma}_{f,1\times3} \end{bmatrix} \Delta t w_{k-1} \tag{8.13}$$

Equation (8.14) is the expanded system model of Eq. 8.11, showing all the system states and how they are coupled through the dynamic matrix.

$$\begin{bmatrix} \delta\dot{\varphi} \\ \delta\dot{\lambda} \\ \delta\dot{h} \\ \delta\dot{V}_e \\ \delta\dot{V}_n \\ \delta\dot{V}_u \\ \delta\dot{p} \\ \delta\dot{r} \\ \delta\dot{A} \\ \delta\dot{\omega}_x \\ \delta\dot{\omega}_y \\ \delta\dot{\omega}_z \\ \delta\dot{f}_x \\ \delta\dot{f}_y \\ \delta\dot{f}_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{R_M+h} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{(R_N+h)\cos\varphi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & f_u & -f_n & 0 & 0 & 0 & R_{11} & R_{12} & R_{13} \\ 0 & 0 & 0 & 0 & 0 & 0 & -f_u & 0 & f_e & 0 & 0 & 0 & R_{21} & R_{22} & R_{23} \\ 0 & 0 & 0 & 0 & 0 & 0 & f_n & -f_e & 0 & 0 & 0 & 0 & R_{31} & R_{32} & R_{33} \\ 0 & 0 & 0 & 0 & \frac{1}{R_M+h} & 0 & 0 & 0 & 0 & R_{11} & R_{12} & R_{13} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{R_N+h} & 0 & 0 & 0 & 0 & 0 & R_{21} & R_{22} & R_{23} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{\tan\varphi}{R_N+h} & 0 & 0 & 0 & 0 & 0 & R_{31} & R_{32} & R_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{\omega x} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{\omega y} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{\omega z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{fx} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{fy} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta_{fz} \end{bmatrix} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \\ \delta V_e \\ \delta V_n \\ \delta V_u \\ \delta p \\ \delta r \\ \delta A \\ \delta\omega_x \\ \delta\omega_y \\ \delta\omega_z \\ \delta f_x \\ \delta f_y \\ \delta f_z \end{bmatrix} + \begin{bmatrix} \sigma_\varphi \\ \sigma_\lambda \\ \sigma_h \\ \sigma_{ve} \\ \sigma_{vn} \\ \sigma_{vu} \\ \sigma_p \\ \sigma_r \\ \sigma_A \\ \sqrt{2\beta_{\omega x}\sigma^2_{\omega x}} \\ \sqrt{2\beta_{\omega y}\sigma^2_{\omega y}} \\ \sqrt{2\beta_{\omega z}\sigma^2_{\omega z}} \\ \sqrt{2\beta_{fx}\sigma^2_{fx}} \\ \sqrt{2\beta_{fy}\sigma^2_{fy}} \\ \sqrt{2\beta_{fz}\sigma^2_{fz}} \end{bmatrix} w \tag{8.14}$$

8.4.2 Measurement Model

The measurement model for a KF in the discrete-time domain is expressed as

$$\delta \mathbf{z}_k = H_k \delta \mathbf{x}_k + \boldsymbol{\eta}_k \tag{8.15}$$

Since the state vector $\delta \mathbf{x}$ of our KF contains the errors in the INS, the corresponding measurement vector$\delta \mathbf{z}_k$ consists of the differences between the position coordinates and velocities predicted by the INS and the corresponding values measured by the GPS, as follows:

$$\delta \mathbf{z}_k = \begin{bmatrix} \mathbf{r}^l_{INS} - \mathbf{r}^l_{GPS} \\ \mathbf{v}^l_{INS} - \mathbf{v}^l_{GPS} \end{bmatrix} = \begin{bmatrix} \varphi_{INS} - \varphi_{GPS} \\ \lambda_{INS} - \lambda_{GPS} \\ h_{INS} - h_{GPS} \\ v_{e,INS} - v_{e,GPS} \\ v_{n,INS} - v_{n,GPS} \\ v_{u,INS} - v_{u,GPS} \end{bmatrix} \tag{8.16}$$

In Eq. (8.15) $\delta \mathbf{x}_k$ is the state vector and $\boldsymbol{\eta}_k$ is a vector of measurement noise which is zero-mean with covariance R_k. The term H_k is the measurement design matrix at time t_k and it describes the linear combinations of state variables that comprise $\mathbf{z}_k$ in the absence of noise. Since in loosely coupled integration the measurements directly correspond to the position and velocity errors states, H_k therefore has the following simple form

$$H_k = \begin{bmatrix} I_{6\times 6} & 0_{6\times 9} \end{bmatrix} \tag{8.17}$$

The full measurement model for the loosely coupled INS/GPS integration can therefore be written as

$$\begin{bmatrix} \mathbf{r}^l_{INS} - \mathbf{r}^l_{GPS} \\ \mathbf{v}^l_{INS} - \mathbf{v}^l_{GPS} \end{bmatrix} = \begin{bmatrix} I_{3\times 3} & 0_{3\times 3} & 0_{3\times 9} \\ 0_{3\times 3} & I_{3\times 3} & 0_{3\times 9} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\eta}_r \\ \boldsymbol{\eta}_v \end{bmatrix} \tag{8.18}$$

the expanded form of which is

$$\begin{bmatrix} \varphi_{INS} - \varphi_{GPS} \\ \lambda_{INS} - \lambda_{GPS} \\ h_{INS} - h_{GPS} \\ v_{e,INS} - v_{e,GPS} \\ v_{n,INS} - v_{n,GPS} \\ v_{u,INS} - v_{u,GPS} \end{bmatrix}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_k + \begin{bmatrix} \eta_\varphi \\ \eta_\lambda \\ \eta_h \\ \eta_{ve} \\ \eta_{vn} \\ \eta_{vu} \end{bmatrix}_k \tag{8.19}$$

We also have two important covariance matrices R_k and P_k.R_k contains the variances of the measured states on its diagonal and is defined as

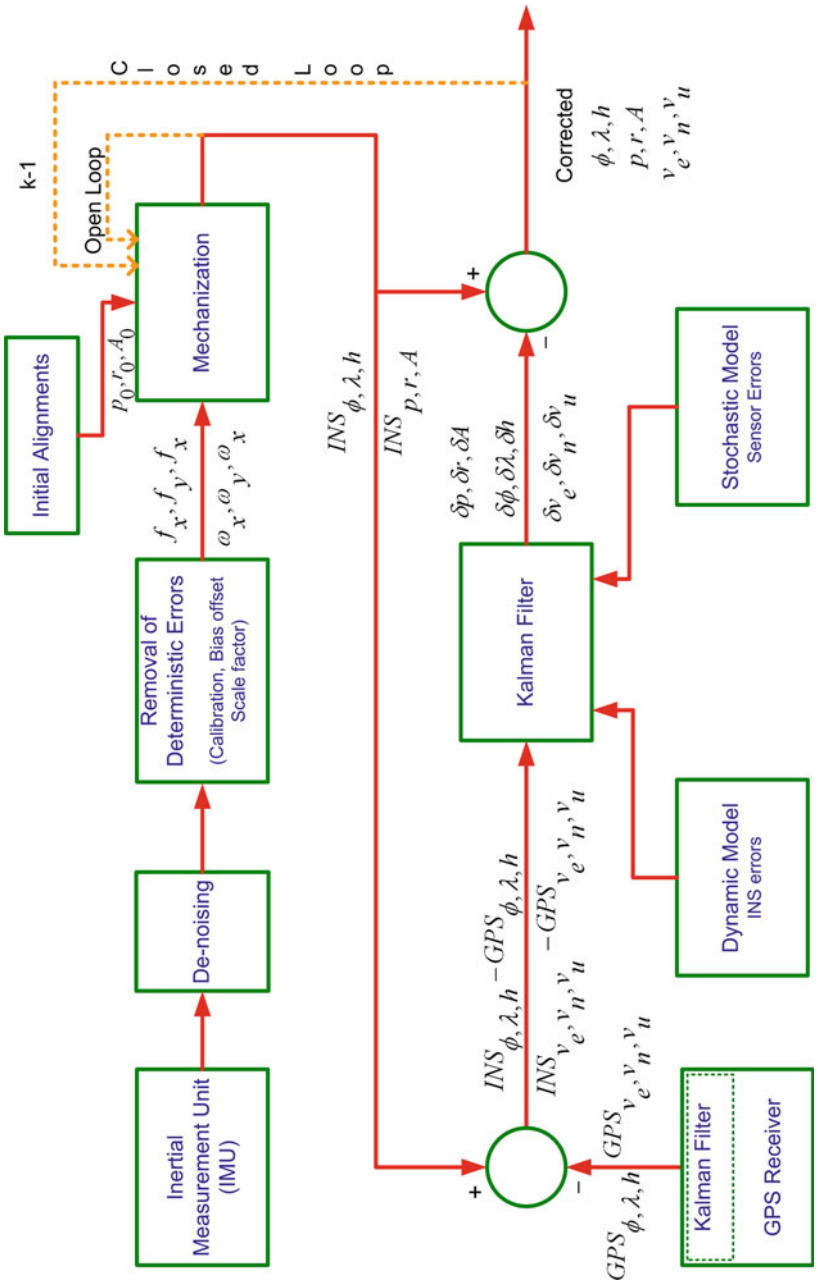


Fig. 8.7 A block diagram for the overall implementation for a loosely coupled GPS/INS integration

$$R_k = \begin{bmatrix} \sigma_{\varphi}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\lambda}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_h^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{ve}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{vn}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{vu}^2 \end{bmatrix} \tag{8.20}$$

The state prediction covariance matrix P_k has variances of predicted states along its diagonal. Cross diagonal elements are cross correlations between various states. In our case it is 15 × 15 element square matrix which (ignoring the cross diagonal elements) is

$$P_k = \begin{bmatrix} \sigma_{r,3\times3}^2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \sigma_{v,3\times3}^2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \sigma_{\varepsilon,3\times3}^2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \sigma_{\omega,3\times3}^2 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \sigma_{f,3\times3}^2 \end{bmatrix} \tag{8.21}$$

where all the sigma terms are also 3 × 3 diagonal matrices that are associated with the position, velocity, attitude, gyroscopes bias drift and accelerometers bias drift.

8.4.3 The Overall Implementation Block Diagram of the Loosely Coupled INS/GPS Integration

The block diagram for the overall implementation of a loosely coupled GPS/INS integration is shown in Fig. 8.7.

8.5 Modeling Tightly Coupled INS/GPS Integration

This section deals with KF for a tightly coupled integration of INS/GPS. It starts with INS dynamic error and measurement models in the l-frame, follows up with a GPS error and measurement error model, and finally presents a combined overall system and measurement model for a KF implementation (Hide et al. 2006; Wang et al. 2005; Wei et al. 2006, 2007; Yang May 2008; Salycheve 2004).

8.5.1 System Model

INS part

The system model for continuous INS KF is

$$\delta\dot{\mathbf{x}} = F_I\delta\mathbf{x}_I + G_I w_I \tag{8.22}$$

where the state vector $\delta\mathbf{x}_I$, dynamic coefficient matrix F_I, noise distribution matrix G_I, and white noise w_I have exactly the same form as the corresponding quantities of the system model of loosely coupled INS/GPS integration outlined in Sect. 8.4.1. These terms will be used later, when we develop the expanded form of the combined INS/GPS system model.

GPS part

The equation for the KF system model for GPS is

$$\delta\dot{\mathbf{x}}_G = F_G\delta\mathbf{x}_G + G_G w_G \tag{8.23}$$

where the bias of the GPS receiver clock δb_r and its drift δd_r are both included as states. They are modeled by random walk, therefore

$$\delta\dot{b}_r = \delta d_r + w_b \tag{8.24}$$

$$\delta\dot{d}_r = w_d \tag{8.25}$$

The state vector of the system model in Eq. (8.23) can be defined as

$$\delta\mathbf{x}_G = [\,\delta b_r \quad \delta d_r\,]^T \tag{8.26}$$

and the dynamic coefficient matrix is

$$F_G = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{8.27}$$

The noise distribution vector for the GPS system model is

$$G_G = [\,\sigma_b \quad \sigma_d\,]^T \tag{8.28}$$

where
σ_b is the standard deviation of the white noise for the clock bias
σ_d is the standard deviation of the white noise for the clock drift

Equation (8.23) therefore can be written as

$$\begin{bmatrix} \delta\dot{b}_r \\ \delta\dot{d}_r \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \delta b_r \\ \delta d_r \end{bmatrix} + \begin{bmatrix} \sigma_b \\ \sigma_d \end{bmatrix} w_G \tag{8.29}$$

where

w_G is the unity variance white Gaussian noise.

Combining INS and GPS system models

The INS and GPS system models can be combined as

$$\delta\dot{\mathbf{x}} = F\delta\mathbf{x} + Gw \tag{8.30}$$

$$\begin{bmatrix} \delta\dot{\mathbf{x}}_I \\ \delta\dot{\mathbf{x}}_G \end{bmatrix} = \begin{bmatrix} F_I & 0 \\ 0 & F_G \end{bmatrix} \begin{bmatrix} \delta x_I \\ \delta x_G \end{bmatrix} + \begin{bmatrix} G_I \\ G_G \end{bmatrix} w \tag{8.31}$$

By inserting the parameters from Eq. (8.11) for the INS part and from Eq. (8.29) for the GPS part we get

$$\begin{bmatrix} \delta\dot{\mathbf{r}}^l_{3\times1} \\ \delta\dot{\mathbf{v}}^l_{3\times1} \\ \dot{\boldsymbol{\varepsilon}}^l_{3\times1} \\ \delta\dot{\boldsymbol{\omega}}_{3\times1} \\ \delta\dot{\mathbf{f}}^l_{3\times1} \\ \delta\dot{b}_r \\ \delta\dot{d}_r \end{bmatrix} = \begin{bmatrix} I_{3\times3} & F_r & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0 & 0 \\ 0_{3\times3} & I_{3\times3} & F_v & 0_{3\times3} & R^l_b & 0 & 0 \\ 0_{3\times3} & F_\varepsilon & I_{3\times3} & R^l_b & 0_{3\times3} & 0 & 0 \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} + F_\omega & 0_{3\times3} & 0 & 0 \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} + F_f & 0 & 0 \\ 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0 & 1 \\ 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\mathbf{r}^l_{3\times1} \\ \delta\mathbf{v}^l_{3\times1} \\ \boldsymbol{\varepsilon}^l_{3\times1} \\ \delta\boldsymbol{\omega}_{3\times1} \\ \delta\mathbf{f}_{3\times1} \\ \delta b_r \\ \delta d_r \end{bmatrix} + \begin{bmatrix} \boldsymbol{\sigma}_{r,3\times1} \\ \boldsymbol{\sigma}_{v,3\times1} \\ \boldsymbol{\sigma}_{\varepsilon,3\times1} \\ \boldsymbol{\sigma}_{\omega,3\times1} \\ \boldsymbol{\sigma}_{f,3\times1} \\ \sigma_b \\ \sigma_d \end{bmatrix} w \tag{8.32}$$

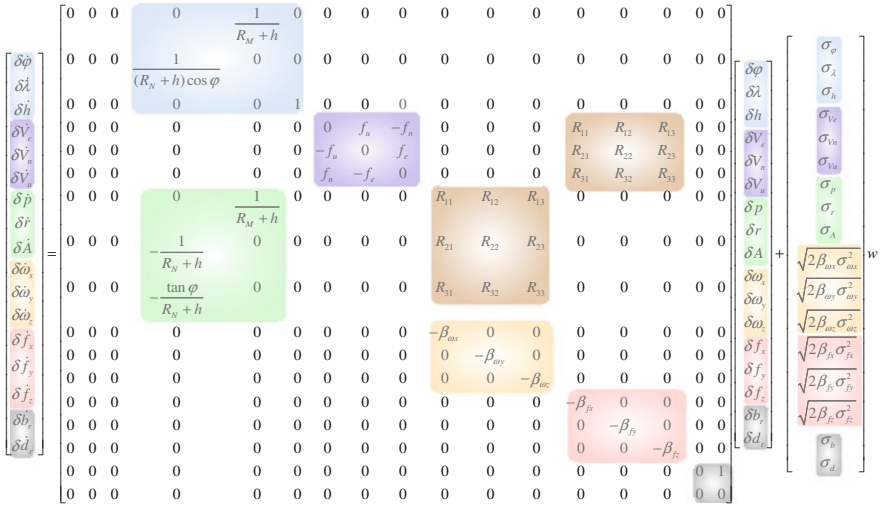
The discrete form of Eq. (8.30) can be written as

$$\delta\mathbf{x}_k = (I + F\Delta t)\delta\mathbf{x}_{k-1} + G\Delta t w_{k-1} \tag{8.33}$$

and hence Eq. (8.32) can be written in the discrete form as

$$\begin{bmatrix} \delta\mathbf{r}^l_{3\times1} \\ \delta\mathbf{v}^l_{3\times1} \\ \boldsymbol{\varepsilon}^l_{3\times1} \\ \delta\boldsymbol{\omega}_{3\times1} \\ \delta\mathbf{f}_{3\times1} \\ \delta b_r \\ \delta d_r \end{bmatrix}_k = \begin{bmatrix} I_{3\times3} & F_r\Delta t & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0 & 0 \\ 0_{3\times3} & I_{3\times3} & F_v\Delta t & 0_{3\times3} & R^l_b\Delta t & 0 & 0 \\ 0_{3\times3} & F_\varepsilon\Delta t & I_{3\times3} & R^l_b\Delta t & 0_{3\times3} & 0 & 0 \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} + F_\omega\Delta t & 0_{3\times3} & 0 & 0 \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} + F_f\Delta t & 0 & 0 \\ 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 1 & \Delta t \\ 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta\mathbf{r}^l_{3\times1} \\ \delta\mathbf{v}^l_{3\times1} \\ \boldsymbol{\varepsilon}^l_{3\times1} \\ \delta\boldsymbol{\omega}_{3\times1} \\ \delta\mathbf{f}_{3\times1} \\ \delta b_r \\ \delta d_r \end{bmatrix}_{k-1} + \begin{bmatrix} \boldsymbol{\sigma}_{r,3\times1} \\ \boldsymbol{\sigma}_{v,3\times1} \\ \boldsymbol{\sigma}_{\varepsilon,3\times1} \\ \boldsymbol{\sigma}_{\omega,3\times1} \\ \boldsymbol{\sigma}_{f,3\times1} \\ \sigma_b \\ \sigma_d \end{bmatrix} w_{k-1} \tag{8.34}$$

Equation (8.35) is the completely expanded form of the combined (INS and GPS) system model.



(8.35)

8.5.2 Measurement Model

The measurement model for a KF in the discrete-time domain is expressed as

$$\delta \mathbf{z}_k = H_k \delta \mathbf{x}_k + \boldsymbol{\eta}_k \tag{8.36}$$

In the tightly coupled implementation, the available observations are pseudo-ranges and pseudo-range rate measurements. Therefore the measurement vector is the differences between these values as predicted by the INS and as measured by the GPS

$$\delta \mathbf{z} = \begin{bmatrix} \delta \mathbf{z}_\rho \\ \delta \mathbf{z}_{\dot{\rho}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\rho}_{INS} - \boldsymbol{\rho}_{GPS} \\ \mathbf{p}_{INS} - \dot{\boldsymbol{\rho}}_{GPS} \end{bmatrix} \tag{8.37}$$

For M satellites this equation can be written as

$$\begin{bmatrix} \delta z_\rho^1 \\ \delta z_\rho^2 \\ \vdots \\ \delta z_\rho^M \\ \delta z_{\dot{\rho}}^1 \\ \delta z_{\dot{\rho}}^2 \\ \vdots \\ \delta z_{\dot{\rho}}^M \end{bmatrix} = \begin{bmatrix} \rho_{INS}^1 - \rho_{GPS}^1 \\ \rho_{INS}^2 - \rho_{GPS}^2 \\ \vdots \\ \rho_{INS}^M - \rho_{GPS}^M \\ \dot{\rho}_{INS}^1 - \dot{\rho}_{GPS}^1 \\ \dot{\rho}_{INS}^2 - \dot{\rho}_{GPS}^2 \\ \vdots \\ \dot{\rho}_{INS}^M - \dot{\rho}_{GPS}^M \end{bmatrix} \tag{8.38}$$

Pseudo-range measurements

Because we want to implement an error state KF, the pseudo-range measurement vector has the following form

$$\delta \mathbf{z}_{\rho} = \boldsymbol{\rho}_{INS} - \boldsymbol{\rho}_{GPS} \tag{8.39}$$

The pseudo-range for the m^{th} satellite to the GPS receiver can be modeled (Misra and Enge Dec 2001) by the following expression

$$\rho_{GPS}^{m} = r^{m} + c\delta t_r - c\delta t_s + cI^{m} + cT^{m} + \tilde{\varepsilon}_{\rho}^{m} \tag{8.40}$$

where

ρ_{GPS}^{m}	is the measured pseudo-range from the m^{th} satellite to the GPS receiver (meters)
r^{m}	is the actual distance between the receiver antenna at the reception time t_r and the satellite's antenna at the transmit time t^{t} (meters)
δt_r	is the receiver's clock offset (sec)
c	is the speed of light (meters/sec)
δt_s	is the satellite's clock offset (sec)
I^{m}	is the ionospheric delay (sec)
T^{m}	is the tropospheric delay (sec)
ε_{ρ}^{m}	is the error due to inexact modeling, receiver noise and multipath.

Satellite clock bias and ionospheric errors can be calculated from the satellite's navigation message, and tropospheric error can also be estimated using an appropriate model. Hence after correcting all the errors except receiver errors (noise and clock bias), we can write the corrected pseudo-range as

$$\rho_{GPS}^{m} = r^{m} + c\delta t_r + \tilde{\varepsilon}_{\rho}^{m} \tag{8.41}$$

where, $\tilde{\varepsilon}_{\rho}^{m}$ represents the total effect of residual errors.

The true geometric range from the m^{th} satellite to the receiver is

$$r^{m} = \sqrt{(x - x^{m})^2 + (y - y^{m})^2 + (z - z^{m})^2} = \|\mathbf{x} - \mathbf{x}^{m}\| \tag{8.42}$$

where

$\mathbf{x} = [x, y, z]^{T}$	is the true receiver position in the e-frame
$\mathbf{x}^{m} = [x^{m}, y^{m}, z^{m}]^{T}$	is the m^{th} satellite's position in the e-frame.

Equation (8.41) can be rewritten as

$$\rho_{GPS}^{m} = \sqrt{(x - x^{m})^2 + (y - y^{m})^2 + (z - z^{m})^2} + \delta b_r + \tilde{\varepsilon}_{\rho}^{m} \tag{8.43}$$

where $\delta b_r = c\delta t_r$ is the error in range (in meters) due to the receiver's clock bias.

The corrected position of the receiver is defined as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_{INS} - \delta x \\ y_{INS} - \delta y \\ z_{INS} - \delta z \end{bmatrix} \tag{8.44}$$

where

$\mathbf{x}_{INS} = [x_{INS},\ y_{INS},\ z_{INS}]^T$ is the output of the mechanization
$\delta\mathbf{x} = [\delta x,\ \delta y,\ \delta z]^T$ is the estimated position error.

Equation (8.43) is non-linear and so must be linearized for KF by applying a Taylor series around x_{INS}. For any function $f(x, y, z)$,, we have the Taylor series expansion around (x_i, y_i, z_i) as

$$f(x,y,z) = f(x_i,y_i,z_i) + \left.\frac{\partial f}{\partial x}\right|_{x_i,y_i,z_i}(x - x_i) + \left.\frac{\partial f}{\partial y}\right|_{x_i,y_i,z_i}(y - y_i) + \left.\frac{\partial f}{\partial z}\right|_{x_i,y_i,z_i}(z - z_i) + \text{Higher order terms(H.O.T)} \tag{8.45}$$

and linearizing Eq. (8.43) around the current best estimate $(x_{INS},\ y_{INS},\ z_{INS})$ yields

$$\rho^m_{GPS} = \sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2} + \frac{(x_{INS} - x^m)(x - x_{INS}) + (y_{INS} - y^m)(y - y_{INS}) + (z_{INS} - z^m)(z - z_{INS})}{\sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2}} + \delta b_r + \tilde{\varepsilon}^m_\rho \tag{8.46}$$

By defining the pseudo-range for the output of the INS to be

$$\rho^m_{INS} = \sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2} \tag{8.47}$$

we have

$$\rho^m_{INS} - \rho^m_{GPS} = -\frac{(x_{INS} - x^m)(x - x_{INS}) + (y_{INS} - y^m)(y - y_{INS}) + (z_{INS} - z^m)(z - z_{INS})}{\sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2}} - \delta b_r + \tilde{\varepsilon}^m_\rho \tag{8.48}$$

$$\rho^m_{INS} - \rho^m_{GPS} = -\frac{(x_{INS} - x^m)(x_{INS} - x) + (y_{INS} - y^m)(y_{INS} - y) + (z_{INS} - z^m)(z_{INS} - z)}{\sqrt{(x_{INS} - x^m)^2 + (y_{INS} - y^m)^2 + (z_{INS} - z^m)^2}} - \delta b_r + \tilde{\varepsilon}^m_\rho \tag{8.49}$$

By defining the line of sight unit vector $\mathbf{1}^m_{INS}$ from the m^{th} satellite to the receiver's position based upon the output of mechanization, we have

$$\mathbf{1}^m_{INS} = \begin{bmatrix} 1^m_{x,INS} \\ 1^m_{y,INS} \\ 1^m_{z,INS} \end{bmatrix} = \begin{bmatrix} \frac{(x_{INS}-x^m)}{\sqrt{(x_{INS}-x^m)^2+(y_{INS}-y^m)^2+(z_{INS}-z^m)^2}} \\ \frac{(y_{INS}-y^m)}{\sqrt{(x_{INS}-x^m)^2+(y_{INS}-y^m)^2+(z_{INS}-z^m)^2}} \\ \frac{(z_{INS}-z^m)}{\sqrt{(x_{INS}-x^m)^2+(y_{INS}-y^m)^2+(z_{INS}-z^m)^2}} \end{bmatrix} \tag{8.50}$$

and

$$\rho^m_{INS} - \rho^m_{GPS} = 1^m_{x,INS}\,\delta x + 1^m_{y,INS}\,\delta y + 1^m_{z,INS}\,\delta z - \delta b_r + \tilde{\varepsilon}^m_\rho \tag{8.51}$$

where

$$\begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} = \begin{bmatrix} x_{INS} - x \\ y_{INS} - y \\ z_{INS} - z \end{bmatrix}$$

and therefore

$$\delta z^m_\rho = \begin{bmatrix} 1^m_{x,INS} & 1^m_{y,INS} & 1^m_{z,INS} \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} - \delta b_r + \tilde{\varepsilon}^m_\rho \tag{8.52}$$

For M visible satellites, this equation can be expressed as

$$\delta \mathbf{z}_\rho = \begin{bmatrix} \rho^1_{INS} - \rho^1_{GPS} \\ \rho^2_{INS} - \rho^2_{GPS} \\ \vdots \\ \rho^M_{INS} - \rho^M_{GPS} \end{bmatrix} = \underbrace{\begin{bmatrix} 1^1_{x,INS} & 1^1_{y,INS} & 1^1_{z,INS} \\ 1^2_{x,INS} & 1^2_{y,INS} & 1^2_{z,INS} \\ \vdots & \vdots & \vdots \\ 1^M_{x,INS} & 1^M_{y,INS} & 1^M_{z,INS} \end{bmatrix}_{M\times 3}}_{G} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}_{3\times 1} - \begin{bmatrix} \delta b_r \\ \delta b_r \\ \vdots \\ \delta b_r \end{bmatrix}_{M\times 1} + \begin{bmatrix} \tilde{\varepsilon}^1_\rho \\ \tilde{\varepsilon}^1_\rho \\ \vdots \\ \tilde{\varepsilon}^m_\rho \end{bmatrix}_{M\times 1} \tag{8.53}$$

By defining a geometry matrix G in the above equation, we get

$$\delta \mathbf{z}_\rho = G_{M\times 3} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix}_{3\times 1} - \delta \mathbf{b}_{r,M\times 1} + \tilde{\boldsymbol{\varepsilon}}_{\rho,M\times 1} \tag{8.54}$$

Because we have the position in ECEF geodetic coordinates, we must convert them to ECEF rectangular coordinates using the following relation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (R_N + h)\cos(\varphi)\cos(\lambda) \\ (R_N + h)\cos(\varphi)\cos(\lambda) \\ \left\{R_N(1 - e^2) + h\right\}\sin(\varphi) \end{bmatrix} \tag{8.55}$$

To be used in KF, the above set of equations has to be linearized through a Taylor series

$$\begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} = \underbrace{\begin{bmatrix} -(R_N + h)\sin\varphi\cos\lambda & -(R_N + h)\cos\varphi\sin\lambda & \cos\varphi\cos\lambda \\ -(R_N + h)\sin\varphi\sin\lambda & (R_N + h)\cos\varphi\cos\lambda & \cos\varphi\sin\lambda \\ \left\{R_N(1 - e^2) + h\right\}\cos\varphi & 0 & \sin\varphi \end{bmatrix}}_{T_{3\times3}} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix} \tag{8.56}$$

By defining a matrix $T_{3\times3}$, the above equation can be expressed as

$$\begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} = T_{3\times3} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix}_{3\times1} \tag{8.57}$$

and substituting this into Eq. (8.54) gives

$$\delta\mathbf{z}_\rho = \underbrace{G_{M\times3}T_{3\times3}}_{H^\rho_{M\times3}} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix}_{3\times1} - \delta\mathbf{b}_{r,M\times1} + \boldsymbol{\varepsilon}_{\rho,M\times1} \tag{8.58}$$

By defining a matrix $H^\rho_{M\times3}$ such that $H^\rho_{M\times3} = G_{M\times3}T_{3\times3}$, we finally obtain the following pseudo-range measurement model

$$\delta\mathbf{z}_\rho = H^\rho_{M\times3} \begin{bmatrix} \delta\varphi \\ \delta\lambda \\ \delta h \end{bmatrix}_{3\times1} - \delta\mathbf{b}_{r,M\times1} + \tilde{\boldsymbol{\varepsilon}}_{\rho,M\times1} \tag{8.59}$$

Pseudo-range rate measurements

The measurement vector for the pseudo-range rate is the difference between the pseudo-range rate predicted by the INS and the value measured by the GPS

$$\delta\mathbf{z}_{\dot\rho} = \dot{\boldsymbol{\rho}}_{INS} - \dot{\boldsymbol{\rho}}_{GPS} \tag{8.60}$$

The Doppler shift produced by satellite and receiver motion is the projection of the relative velocities onto the line of sight, scaled by the transmission frequency and divided by speed of light

$$D^m = \frac{[(\mathbf{v}^m - \mathbf{v}) \cdot \mathbf{1}^m]L_1}{c} \tag{8.61}$$

where

$\mathbf{v}^m = \left[v_x^m, v_y^m, v_z^m\right]^T$ is the velocity of the m^{th} satellite in the e-frame

$\mathbf{v} = \left[v_x, v_y, v_z\right]^T$ is the true receiver velocity in the e-frame

L_1 is the satellite's transmission frequency

c is the speed of light

$\mathbf{1}^m$ is line of sight unit vector from the m^{th} satellite to the GPS receiver, which can be expressed as

$$\mathbf{1}^m = \frac{[(x - x^m), (y - y^m), (z - z^m)]^T}{\sqrt{(x - x^m)^2 + (y - y^m)^2 + (z - z^m)^2}} = \begin{bmatrix} 1_x^m & 1_y^m & 1_z^m \end{bmatrix}^T \tag{8.62}$$

Given the Doppler, the pseudo-range rate $\dot{\rho}^m$ can be computed as

$$\dot{\rho}^m = -\frac{D^m c}{L_1} \tag{8.63}$$

and the true pseudo-range rate is

$$\dot{r}^m = 1_x^m.(v_x - v_x^m) + 1_y^m.(v_y - v_y^m) + 1_z^m.(v_z - v_z^m) \tag{8.64}$$

This pseudo-range rate can be modeled as

$$\dot{\rho}_{GPS}^m = 1_x^m.(v_x - v_x^m) + 1_y^m.(v_y - v_y^m) + 1_z^m.(v_z - v_z^m) + c\delta\dot{t}_u + \varepsilon_{\dot{\rho}}^m \tag{8.65}$$

$$\dot{\rho}_{GPS}^m = 1_x^m.(v_x - v_x^m) + 1_y^m.(v_y - v_y^m) + 1_z^m.(v_z - v_z^m) + \delta d_r + \varepsilon_{\dot{\rho}}^m \tag{8.66}$$

where $\delta d_r = c\delta\dot{t}_u$ is the receiver's clock drift in meters/sec and $\varepsilon_{\dot{\rho}}^m$ is the error in observation (meters/sec).

As we want our measurement model as a function of δv_x, δv_y, δv_z, Eq. (8.66) must be converted to the error state.

The INS estimated pseudo-range rate is given as

$$\dot{\rho}_{INS}^m = 1_{x,INS}^m.(v_{x,INS} - v_x^m) + 1_{y,INS}^m.(v_{y,INS} - v_y^m) + 1_{z,INS}^m.(v_{z,INS} - v_z^m) \tag{8.67}$$

where $v_{x,INS}$, $v_{y,INS}$, $v_{z,INS}$ are the receiver's velocities in the e-frame estimated by the INS.

Taking the difference of INS-estimated pseudo-range rate Eq. (8.67) and GPS-measured pseudo-range rate Eq. (8.66) we get

$$\dot{\rho}^m_{INS} - \dot{\rho}^m_{GPS} = -1^m_{x,INS}.(v_x - v_{x,INS}) - 1^m_{y,INS}.(v_y - v_{y,INS}) + 1^m_{z,INS}.(v_z - v_{z,INS}) - \delta d_r + \varepsilon^m_{\dot{\rho}} \tag{8.68}$$

$$\dot{\rho}^m_{INS} - \dot{\rho}^m_{GPS} = 1^m_{x,INS}.(v_{x,INS} - v_x) + 1^m_{y,INS}.(v_{y,INS} - v_y) + 1^m_{z,INS}.(v_{z,INS} - v_z) - \delta d_r + \varepsilon^m_{\dot{\rho}} \tag{8.69}$$

$$\delta z^m_{\dot{\rho}} = 1^m_{x,INS}.\delta v_x + 1^m_{y,INS}.\delta v_y + 1^m_{z,INS}.\delta v_z - \delta d_r + \varepsilon^m_{\dot{\rho}} \tag{8.70}$$

where $\delta z^m_{\dot{\rho}} = \dot{\rho}^m_{INS} - \dot{\rho}^m_{GPS}$ and $\begin{bmatrix} \delta v_x \\ \delta v_y \\ \delta v_z \end{bmatrix} = \begin{bmatrix} v_{x,INS} - v_x \\ v_{y,INS} - v_y \\ v_{z,INS} - v_z \end{bmatrix}$.

Writing Eq. (8.70) in state space form gives

$$\delta z^m_{\dot{\rho}} = \begin{bmatrix} 1^m_{x,INS} & 1^m_{y,INS} & 1^m_{z,INS} \end{bmatrix} \begin{bmatrix} \delta v_x \\ \delta v_y \\ \delta v_z \end{bmatrix} - \delta d_r + \varepsilon^m_{\dot{\rho}} \tag{8.71}$$

and for M satellites the pseudo-range rate measurement model is

$$\delta \mathbf{z}_{\dot{\rho}} = \begin{bmatrix} \dot{\rho}^1_{INS} - \dot{\rho}^1_{GPS} \\ \dot{\rho}^2_{INS} - \dot{\rho}^2_{GPS} \\ \vdots \\ \dot{\rho}^M_{INS} - \dot{\rho}^M_{GPS} \end{bmatrix} = \underbrace{\begin{bmatrix} 1^1_{x,INS} & 1^1_{y,INS} & 1^1_{z,INS} \\ 1^2_{x,INS} & 1^2_{y,INS} & 1^2_{z,INS} \\ \vdots & \vdots & \vdots \\ 1^M_{x,INS} & 1^M_{y,INS} & 1^M_{z,INS} \end{bmatrix}_{M\times 3}}_{G} \begin{bmatrix} \delta v_x \\ \delta v_y \\ \delta v_z \end{bmatrix} - \begin{bmatrix} \delta d_r \\ \delta d_r \\ \vdots \\ \delta d_r \end{bmatrix}_{M\times 1} + \begin{bmatrix} \varepsilon^1_{\dot{\rho}} \\ \varepsilon^2_{\dot{\rho}} \\ \vdots \\ \varepsilon^M_{\dot{\rho}} \end{bmatrix}_{M\times 1} \tag{8.72}$$

$$\delta \mathbf{z}_{\dot{\rho}} = G_{M\times 3} \begin{bmatrix} \delta v_z \\ \delta v_y \\ \delta v_z \end{bmatrix}_{M\times 1} - \delta \mathbf{d}_{r,M\times 1} + \boldsymbol{\varepsilon}_{\dot{\rho},M\times 1} \tag{8.73}$$

The relationship between the velocity in the l-frame and the e-frame is

$$\begin{bmatrix} \delta v_z \\ \delta v_y \\ \delta v_z \end{bmatrix} = R^e_l \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} \tag{8.74}$$

Substituting for R^e_l gives

$$\begin{bmatrix} \delta v_z \\ \delta v_y \\ \delta v_z \end{bmatrix} = \begin{bmatrix} -\sin\lambda & \sin\varphi\cos\lambda & \cos\varphi\cos\lambda \\ \cos\lambda & \sin\varphi\sin\lambda & \cos\varphi\sin\lambda \\ 0 & \cos\varphi & \sin\varphi \end{bmatrix} \begin{bmatrix} \delta v_e \\ \delta v_n \\ \delta v_u \end{bmatrix} \tag{8.75}$$

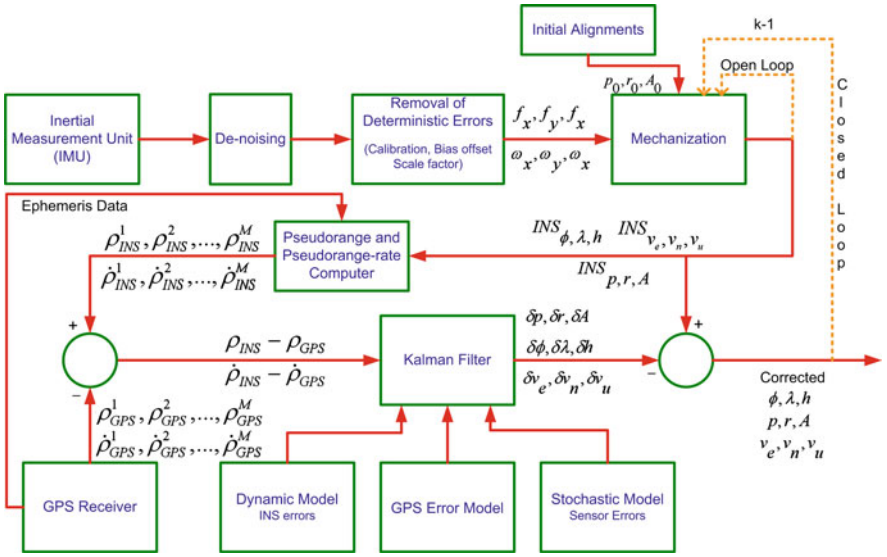


Fig. 8.8 An overall implementation diagram of tightly coupled INS/GPS integration

and the pseudo-range rate measurement model becomes

$$\delta \mathbf{z}_{\dot{\rho}} = \underbrace{G_{M\times 3} R_l^e}_{H^{\rho}_{M\times 3}} \begin{bmatrix} \delta v_E \\ \delta v_N \\ \delta v_U \end{bmatrix} - \delta \mathbf{d}_{r,M\times 1} + \boldsymbol{\varepsilon}_{\dot{\rho},M\times 1} \tag{8.76}$$

By defining a matrix $H^{\dot{\rho}}_{M\times 3} = G_{M\times 3} R_l^e$, we finally have the expression for the pseudo-range rate measurement model

$$\delta \mathbf{z}_{\dot{\rho}} = H^{\dot{\rho}}_{M\times 3} \begin{bmatrix} \delta v_E \\ \delta v_N \\ \delta v_U \end{bmatrix} - \delta \mathbf{d}_{r,M\times 1} + \boldsymbol{\varepsilon}_{\dot{\rho},M\times 1} \tag{8.77}$$

8.5.3 The Overall Measurements Model

The measurement model for pseudo-range errors provided by Eq. (8.59) and pseudo-range rate errors provided by Eq. (8.77) can be combined to create an overall measurement model as

$$\underbrace{\begin{bmatrix} \rho_{INS}^1 - \rho_{GPS}^1 \\ \rho_{INS}^2 - \rho_{GPS}^2 \\ \vdots \\ \rho_{INS}^M - \rho_{GPS}^M \\ \dot{\rho}_{INS}^1 - \dot{\rho}_{GPS}^1 \\ \dot{\rho}_{INS}^2 - \dot{\rho}_{GPS}^2 \\ \vdots \\ \dot{\rho}_{INS}^M - \dot{\rho}_{GPS}^M \end{bmatrix}_{2M\times 1}}_{\delta z} = \underbrace{\begin{bmatrix} H_{M\times 3}^{\rho} & 0_{M\times 3} & 0_{M\times 9} & -1_{M\times 1} & 0_{M\times 1} \\ 0_{M\times 3} & H_{M\times 3}^{\dot{\rho}} & 0_{M\times 9} & 0_{M\times 1} & -1_{M\times 1} \end{bmatrix}_{2M\times 17}}_{H} \delta \mathbf{x}_{17\times 1} + \underbrace{\begin{bmatrix} \widetilde{\boldsymbol{\varepsilon}}_{\rho,M\times 1} \\ \boldsymbol{\varepsilon}_{\dot{\rho},M\times 1} \end{bmatrix}_{2M\times 1}}_{\eta} \tag{8.78}$$

The implementation of the tightly coupled INS/GPS integration is summarized as a block diagram in Fig. 8.8.

References

Alban S, Akos DM, Rock SM, Gebre-Egziabher D (2003) Performance analysis and architectures for INS-aided GPS tracking loops. In: Proceedings of ION national technical meeting (NTM), U.S. Institute of navigation, Fairfax, Anaheim 11–24 Jan 2003, pp 611–622

Gebre-Egziabher D (2007) What is the difference between 'loose', 'tight' and 'ultra-tight' and 'deep' integration strategies for INS and GNSS. InsideGNSS

Groves PD (2008) Principles of GNSS, inertial and multi-sensor integrated navigation systems. Artech House, Boston

Hide C, Moore T, Hill C, Park D (2006) Low cost, high accuracy positioning in urban environments. J Navig 59(3):365–379

Maybeck PS (1979) Stochastic models, estimation, and control, vol 1. Academic, New York

Minkler G, Minkler J (1993) Theory and application of Kalman filtering. Magellan Book Company, Palm Bay

Misra P, Enge P (2001) Global positioning system: signals, measurements and performance. Ganga-Jamuna Press, Lincoln

Salycheve OS (2004) Applied inertial navigation: problem and solutions. BMSTU Press, Moscow

Wagner JF, Wieneke T (2003) Integrating satellite and inertial navigation-conventional and new fusion approaches. Control Eng Pract 11(5):543–550

Wang W, Liu Z, Xie R (2005) An improved tightly coupled approach for GPS/INS integration. In: Piscataway NJ (ed) 2004 IEEE conference on robotics, automation and mechatronics (IEEE Cat. No.04EX913). IEEE, pp 1164–1167

Wei G, Qi N, Guofu Z, Hui J (2007) Gyroscope drift estimation in tightly-coupled INS/GPS navigation system. In: Piscataway NJ (ed) 2007 second IEEE conference on industrial electronics and applications. IEEE, pp 391–396

Wei W, Zong-yu L, Rong-rong X (2006) Quadratic extended Kalman filter approach for GPS/INS integration. Aerosp Sci Technol 10(8):709–713

Yang Y (2008) Tightly coupled MEMS INS/GPS integration with INS aided receiver tracking loops. Ph.D. dissertation, University of Calgary, Calgary

Chapter 9
Three-Dimensional Reduced Inertial Sensor System/GPS Integration for Land-Based Vehicles

This chapter discusses a dead reckoning (DR) solution which is suitable for any wheel-based platform integrated with GPS. It eliminates several error sources that exist when using a traditional full IMU, especially low cost MEMS grade sensors. After discussing and analyzing the performance of a full IMU system, the theory of the methods employed to tackle sources of errors will be outlined. The reduced inertial sensor system is introduced and compared to a traditional full IMU, and its mechanization equations derived. This is followed by a description of both loosely and tightly coupled KF-based integration of this reduced inertial sensor system with GPS, including the linearized system model and measurement model for each integration scheme.

9.1 Performance Analysis of 3D Positioning Utilizing a MEMS Grade Full IMU

To demonstrate the sources of performance degradation during GPS outages in a three-dimensional (3D) navigation solution based on a low cost MEMS grade full IMU integrated with GPS, three uncompensated biases will be investigated: an uncompensated bias in any of the three accelerometers; an uncompensated bias in one of the horizontally aligned gyroscopes; and an uncompensated bias in the vertically aligned gyroscope. As was discussed in Chap. 7, both the LKF or EKF solutions attempt to compensate for these biases by using the error models for the sensors that are used within the KF-based solution. These models are indirectly updated by the KF-based solution when GPS is available. During GPS outages the filter runs in prediction mode with the predicted values of the INS errors being subtracted from the mechanization solution to correct it. But this compensation is not perfect and the residual uncompensated parts of these errors will remain. The effect of these uncompensated biases will be discussed below.

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_9,

As pointed out in Chap. 4, when calculating position and velocity from accelerometers, an uncompensated accelerometer bias error b_f will introduce

1. An error in velocity proportional to t (the GPS outage duration, when the KF is working in prediction mode) $\delta v_1 = \int b_f dt = b_f t$.
2. An error in position proportional to t^2 which can be expressed as $\delta p_1 = \int \delta v_1 dt = \int b_f t dt = \frac{1}{2} b_f t^2$.

For the second source of errors, an uncompensated bias b_w in one of the two horizontally aligned gyroscopes introduces

1. An angle error in either pitch or roll that is proportional to time, $\delta\theta = \int b_w dt = b_w t$. This small angle will result in a misalignment of the INS. Therefore when projecting the acceleration from the body frame to the local-level frame, the acceleration vector is projected incorrectly owing to this misalignment.
2. An error in acceleration in one of the horizontal channels in the local-level frame $\delta a_2 = g \sin(\delta\theta) \approx g\delta\theta \approx g b_w t$ (for the bounded duration t during a GPS outage, $\delta\theta$ will be a small angle because b_w is a very small value and the approximation $\sin(\delta\theta) \approx \delta\theta$ holds).
3. This causes an error in velocity $\delta v_2 = \int b_w g t dt = \frac{1}{2} b_w g t^2$ and an error in position $\delta p_2 = \int \delta v_2 dt = \int \frac{1}{2} b_w g t^2 dt = \frac{1}{6} b_w g t^3$.

Thus a gyroscope bias introduces a third order error in position. As confirmed in (Skog and Handel 2009) the effect of these gyroscope errors will be the most influential part of position and velocity errors when the INS works alone, such as during GPS outages.

A third source of error arises because any uncompensated bias in the vertically aligned gyroscope, b_w, will cause an error in azimuth $\delta A = \int b_{w_z} dt = b_{w_z} t$ that is proportional to time. The resulting positional error will be proportional to vehicle speed, time and this azimuth error (which is in turn proportional to time and the uncompensated bias).

9.2 The Proposed Techniques for Overcoming MEMS Grade IMU Shortcomings for Land-Based Vehicles

The navigation solution outlined in this chapter is suitable for land-based vehicles that have a source of speed readings, such as an odometer or wheel encoders. This will address uncompensated errors whilst also minimizing the cost by using fewer inertial sensors.

To overcome the first error discussed above due to accelerometer biases, the odometer-derived speed will be used in the navigation solutions. To overcome the second error due to horizontal gyroscope biases, the pitch and roll angles will be obtained by combining measurements from the accelerometers and the odometer.

The calculation of pitch and roll angles from accelerometers was suggested before for the Measurement-While-Drilling surveying application (Noureldin et al. 2002, 2004). There now follows a brief derivation of the pitch and roll equations.

When the vehicle is stationary the accelerometers measure components due to gravity owing to the tilt from the horizontal plane by the pitch and roll angles. The accelerometer measurements are

$$\begin{bmatrix} f^x \\ f^y \\ f^z \end{bmatrix} = R_\ell^b \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} = \left(R_b^\ell\right)^T \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} = \begin{bmatrix} -g\cos p \sin r \\ g \sin p \\ g \cos p \cos r \end{bmatrix} \tag{9.1}$$

Since only two accelerometers along the X and Y directions are used, the pitch and the roll angles can be expressed as

$$p = \sin^{-1}\left(\frac{f^y}{g}\right) \tag{9.2}$$

$$r = \sin^{-1}\left(\frac{-f^x}{g\cos p}\right) \tag{9.3}$$

When the vehicle is moving the forward accelerometer (corrected for sensor errors) measures the forward vehicle acceleration as well as the component due to gravity. In order to calculate the pitch angle, the vehicle acceleration derived from the odometer measurements must be subtracted from the forward accelerometer measurements

$$p = \sin^{-1}\left(\frac{f^y - a^{od}}{g}\right) \tag{9.4}$$

where g is the gravity acceleration, f^y is the forward accelerometer measurement (corrected for sensor errors) and a^{od} is the vehicle acceleration derived from the odometer measurements.

Similarly the transversal accelerometer (corrected for sensor errors) measures the normal component of the vehicle acceleration as well as the component due to gravity. To calculate the roll angle, the transversal accelerometer measurement must be compensated for the normal component of acceleration

$$r = -\sin^{-1}\left(\frac{f^x + v^{od}\omega^z}{g\cos p}\right) \tag{9.5}$$

where f^x is the transversal accelerometer measurement (corrected for sensor errors), v^{od} is the speed derived from the odometer measurements, w^z is the angular rate measured by the vertically aligned gyro (corrected for sensor errors) and $v^{od}w^z$ is the normal component of the vehicle acceleration.

For low cost MEMS grade sensors the calculation of pitch and roll from the two accelerometers is better than that from the two gyroscopes for several reasons. Firstly, calculating pitch and roll from the two horizontal gyroscopes involves an

integration operation, whereas calculating them from the accelerometers does not. The drawback of integration is the accumulation of errors due to uncompensated sensor bias errors, impairing the positional accuracy. Another reason to eliminate these two gyroscopes is that accelerometers are less expensive.

The third source of error that is due to the vertical gyroscope bias needs to be modeled by the integration filter that integrates these sensors with GPS.

9.3 Three-Dimensional Reduced Inertial Sensor System

To cut the cost of a navigation solution for land vehicles, a reduced inertial sensor system (RISS) which employs one single-axis gyroscope and the odometer was suggested for two-dimensional (2D) navigation (Iqbal and Noureldin 2009; Iqbal et al. 2008, 2009) and integrated with GPS using a KF. This 2D RISS/GPS solution assumes that the vehicle moves mostly in the horizontal plane, which is generally plausible. This was called a RISS because it also contained a pair of accelerometers for calculating pitch and roll (Iqbal et al. 2008), but these angles were to be obtained as separate quantities and not utilized to estimate the off-plane motion. The single gyroscope, which had its sensitive axis aligned with the vertical axis of the vehicle, was used along with the odometer, and the entire system was integrated with GPS. On the assumption that the vehicle operated in the horizontal plane, its speed was derived from the odometer measurements and used with the heading information obtained from the gyroscope to determine the velocities along the east and north directions and thereby enable the longitude and latitude to be tracked. This reduced multi-sensor system provides an effective 2D navigation solution for land vehicles.

As described in (Iqbal et al. 2008), two additional accelerometers may be used if it is necessary to determine the pitch and roll angles. As described earlier, two accelerometers pointing in the forward and transverse directions of the vehicle, a reliable model for the Earth's gravity and speed readings can provide the roll and pitch angles. These angles were not included in the dynamic model used by KF to determine when a system is off the horizontal plane.

9.3.1 Overview of 3D RISS

For the 3D RISS solution discussed here, the measurements provided by the two accelerometers used to calculate pitch and roll are incorporated to estimate the off-plane motion and allow this solution to target a 3D navigation solution. It has been proven (Brandt and Gardner 1998) that due to the non-holonomic constraints on the land vehicle the forward accelerometer and three gyroscopes are sufficient (as opposed to a full IMU with three accelerometers and three gyroscopes) to provide a navigation solution that calculates 3D position, velocity and attitude.

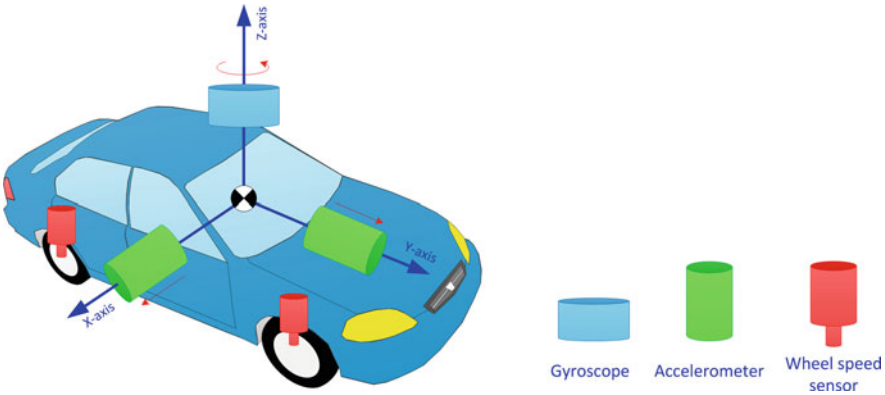


Fig. 9.1 The sensors used in a 3D RISS depicted with respect to the body frame of a land vehicle

This also means that the three gyroscopes together with the forward speed derived from the vehicle's odometer are sufficient to achieve 3D navigation. The work presented in this chapter is a 3D navigation solution using the vehicle's odometer, one single-axis gyroscope and two accelerometers. The gyroscope is aligned with the vertical axis of the vehicle. The pitch and roll angles that would have been provided by the two eliminated gyroscopes are now calculated using the two accelerometers. The 3D RISS integrates the measurements from the vertically aligned gyroscope and the two horizontal accelerometers with speed readings provided by an odometer or wheel encoders. The 3D RISS system is depicted in Fig. 9.1. Using a 3D RISS to obtain a navigation solution of 3D position, 3D velocity and 3D attitude was first proposed in (Georgy et al. 2010).

9.3.2 Advantages of 3D RISS for Wheel-Based Land Vehicles

The advantages of the 3D RISS over 2D vehicular dead reckoning solutions are based on the fact that the measurements from the two accelerometers are used to calculate the off-plane motion. The first benefit is the calculation of the correct azimuth angle. This is because the gyroscope (vertically aligned to the body frame of the vehicle) is tilted with the vehicle when it is not purely horizontal and thus is not measuring the angular rate in the horizontal E-N plane. Because the azimuth angle is in the E-N plane, detecting and correcting the gyroscope tilt yields a more accurate calculation of the azimuth angle than 2D vehicular dead reckoning that neglects this effect. The second benefit of 3D RISS is more accurate 2D horizontal positioning than 2D RISS for two reasons, namely: (1) incorporation of the pitch angle in calculating the two horizontal velocities from the speed measured by the odometer and a more accurate velocity provides better position estimates; and (2) the more accurate azimuth calculation leads to better estimates of velocities along

east and north and thus a better 2D position. The third advantage of 3D RISS over 2D vehicular dead reckoning is the calculation of the upward velocity and altitude, quantities which were not previously available.

The advantages of the proposed 3D RISS over a full IMU are the calculation of pitch and roll from accelerometers instead of gyroscopes, and the calculation of the vehicle's velocity from an odometer-derived speed instead of accelerometers. To demonstrate the superiority of calculating pitch and roll from accelerometers rather than gyroscopes consider the uncompensated bias in one of the eliminated horizontal gyroscopes, which will introduce an angle error in pitch or roll that is proportional to time arising from integration and thus cause a misalignment of the INS that will incorrectly project the acceleration vector from the body frame to the local-level frame, in turn introducing an error in the acceleration in one of the horizontal channels of the local-level frame that will produce an error in velocity proportional to t^2 and an error in position proportional to t^3. When pitch and roll are calculated from accelerometers, the first integration is eliminated and the error in these angles is not proportional to time. Furthermore, the portion of the position error that arises from these angle errors will be proportional to t^2 rather than t^3.

In addition to these advantages of using two accelerometers rather than two gyroscopes for calculating pitch and roll, a RISS provides a further improvement in the velocity calculations. Calculating velocity using the forward speed derived from an odometer rather than from accelerometers (in accordance with the non-holonomic constraints imposed on land vehicles) achieves better performance than calculating it from the accelerometers. This is because when calculating velocity from accelerometers any uncompensated accelerometer bias error will introduce an error proportional to t in velocity and an error proportional to t^2 in position. The calculation of velocity from an odometer avoids the first integration, thereby making the position calculation require a single integration. This means that the position is obtained after a single integration when using odometer measurements as opposed to two consecutive integrations using accelerometer measurements. In long GPS outages the error when using accelerometers will be proportional to the square of the outage duration, which will be drastic for long outages.

As a result of the two abovementioned improvements, a further improvement of the position calculation will follow. The errors in pitch and roll calculated from accelerometers (no longer proportional to time) will cause a misalignment of the inertial system that will influence the projection of velocity (for a RISS) instead of acceleration (for a full IMU) from the body frame to the local-level frame, making the portion of the position error arising from pitch and roll errors proportional to t rather than to the t^2 that was discussed in the first improvement to be gained from eliminating the two gyroscopes.

The only remaining main source of error in 3D RISS is the azimuth error due to the vertically aligned gyroscope. An uncompensated bias in this gyroscope will cause an error in azimuth that is proportional to time. The position error due to this azimuth error will be proportional to vehicle speed, time, and the azimuth error (in

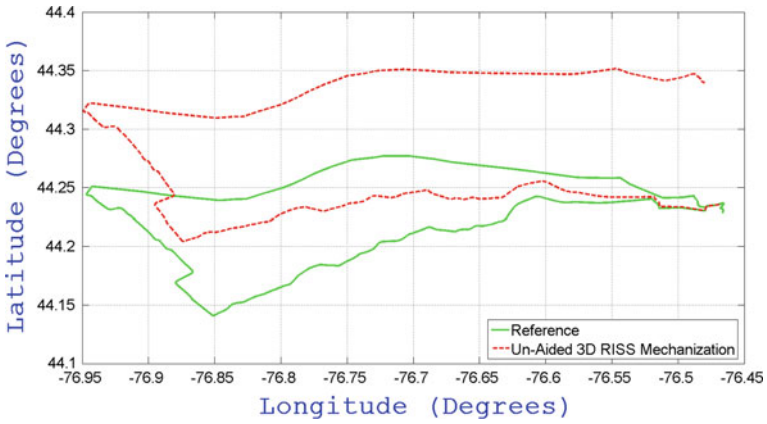


Fig. 9.2 Positioning results for a 3D RISS mechanization versus a high-end reference solution

turn proportional to time and the uncompensated bias). This must be tackled inside the integration filter by modeling the stochastic drift of the gyroscope.

To demonstrate the advantages of 3D RISS over a full IMU for land vehicles, the positioning output of the standalone unaided mechanization for each system is presented for driving a land vehicle around the Kingston area in Ontario, Canada, on a trajectory that traveled a distance of around 96 km over a duration of almost 100 min. The positioning results for both systems are presented in Figs. 9.2, 9.3 and 9.4. Figure 9.2 shows a 3D RISS standalone mechanization in comparison to a high-end reference solution, while Fig. 9.3 shows a full IMU standalone mechanization in comparison to this same reference. It is evident from Fig. 9.3 that the full IMU standalone solution drifts significantly over time and soon deviates from the true trajectory. Comparing the two standalone solutions of Figs. 9.2 and 9.3 demonstrates the significance and the importance of the 3D RISS solution. Figure 9.4 shows a zoom-in at the beginning of the trajectory to show how the standalone full IMU solution degrades more rapidly than that of the 3D RISS. The MEMS sensors used in both 3D RISS and full IMU came from the same IMU, a model IMU300CC-100 manufactured by Crossbow. The gyroscope bias of this IMU is 2°/s as listed in Table 10.1. It must be noted that for the mechanization laboratory calibration values were used in both cases to remove the accelerometer bias offsets from the readings and the gyroscope biases offset were obtained by averaging the readings during a static period at the start of the trajectory and they were removed upon finishing the trajectory. The high-end reference is an off-the-shelf solution by NovAtel called a synchronized position attitude navigation (SPAN) HG, incorporating a high-end tactical grade IMU from Honeywell (HG 1,700) with a dual frequency NovAtel GPS receiver (OEM4).

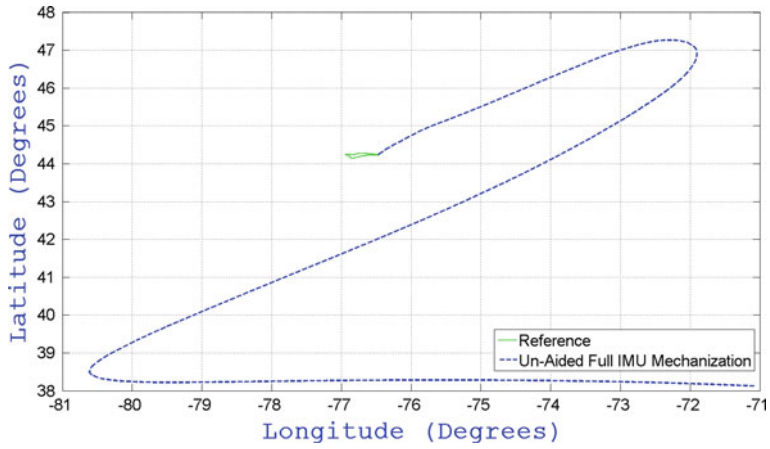


Fig. 9.3 Positioning results for a full IMU mechanization versus a high-end reference solution

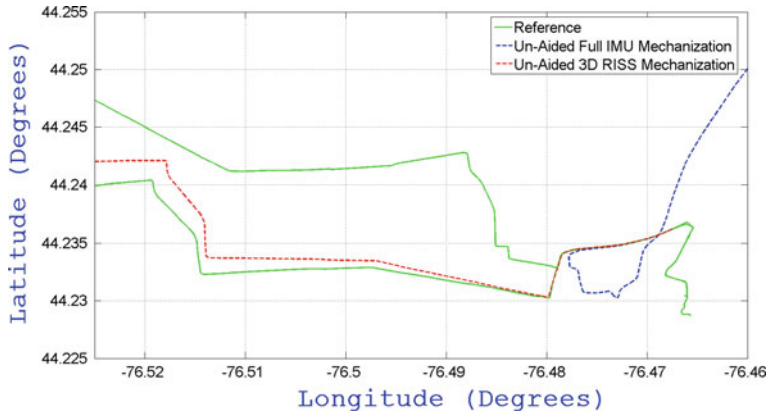


Fig. 9.4 A close-up of the positioning results at the beginning of the trajectory for a 3D RISS mechanization, a full IMU mechanization and a high-end reference solution

9.3.3 Derivation of the 3D RISS Motion Equations

The non-linear motion model for a 3D RISS involving the position, velocity and attitude states is presented in this section.

The common reference frames are used. The body frame of the vehicle has its X-axis along the transversal direction, its Y-axis along the forward longitudinal direction, and its Z-axis along the vertical direction of the vehicle. The local-level frame is the ENU frame with axes along east, north and vertical (up) directions. The rotation matrix that transforms from the vehicle body frame to the local-level frame at time $k-1$ is

$$R^{l}_{b,k-1} = \begin{bmatrix} \cos A_{k-1}\cos r_{k-1} + \sin A_{k-1}\sin p_{k-1}\sin r_{k-1} & \sin A_{k-1}\cos p_{k-1} & \cos A_{k-1}\sin r_{k-1} - \sin A_{k-1}\sin p_{k-1}\cos r_{k-1} \\ -\sin A_{k-1}\cos r_{k-1} + \cos A_{k-1}\sin p_{k-1}\sin r_{k-1} & cosA_{k-1}\cos p_{k-1} & -\sin A_{k-1}\sin r_{k-1} - \cos A_{k-1}\sin p_{k-1}\cos r_{k-1} \\ -\cos p_{k-1}\sin r_{k-1} & \sin p_{k-1} & \cos p_{k-1}\cos r_{k-1} \end{bmatrix} \tag{9.6}$$

In order to describe the motion model, we first require the control inputs. The sensors measurements provided by the gyroscope, the two accelerometers, and the odometer comprise the control inputs represented by the vector

$$\mathbf{u}_k = \begin{bmatrix} v_k^{od} & a_k^{od} & f_k^x & f_k^y & \omega_k^z \end{bmatrix}^T \tag{9.7}$$

where

v_k^{od} is the speed derived from the vehicle's odometer
a_k^{od} is the acceleration derived from the vehicle's odometer
f_k^x is the transversal accelerometer measurement
f_k^y is the forward accelerometer reading
ω_k^z is the angular rate obtained from the vertically aligned gyroscope.

The control inputs with the suffix k in this discussion denote those that cause the system state to change between the time epoch $k-1$ and k.

Before deriving the motion equations, it is important to define the navigation state of the system. One possible state vector is

$$\mathbf{x}_k = \left[\varphi_k, \lambda_k, h_k, v_k^f, p_k, r_k, A_k,\right]^T \tag{9.8}$$

where

φ_k is the latitude of the vehicle
λ_k is the longitude
h_k is the altitude
v_k^f is the forward speed
p_k is the pitch angle
r_k is the roll angle
A_k is the azimuth angle.

Another possible state vector is $\left[\varphi_k, \lambda_k, h_k, v_k^E, v_k^N, v_k^U, p_k, r_k, A_k,\right]^T$ where v_k^E, v_k^N, and v_k^U are the velocity components along the east, north and up directions.

9.3.3.1 Pitch and Roll Calculation

For a land vehicle the pitch angle is the angle that it makes with respect to level ground (i.e. rotation about the transversal axis) and the roll angle is its rotation about the longitudinal axis, as depicted in Fig. 9.5.

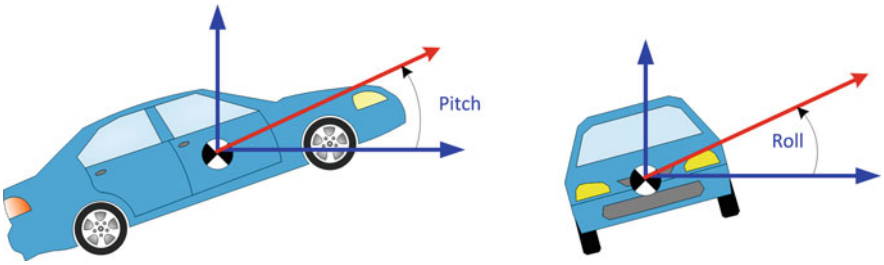


Fig. 9.5 Pitch and roll angles of a land vehicle

When the vehicle is moving the forward accelerometer measures the forward vehicle acceleration as well as the component due to gravity, and to calculate the pitch angle the vehicle acceleration derived from the odometer measurements is removed from the forward accelerometer measurements

$$p_k = \sin^{-1}\left(\frac{f_k^y - a_k^{od}}{g}\right) \tag{9.9}$$

Similarly the transversal accelerometer measures the normal component of the vehicle acceleration as well as the component due to gravity, and to calculate the roll angle the transversal accelerometer measurement must be compensated for the normal component of acceleration

$$r_k = -\sin^{-1}\left(\frac{f_k^x + v_k^{od}\omega_k^z}{g\cos p_k}\right) \tag{9.10}$$

9.3.3.2 Azimuth Calculation

The azimuth angle of a land vehicle is its deviation from north, as in Fig. 9.6.

In time interval Δt between time epochs $k-1$ and k, the counterclockwise angle of rotation around the vertical axis of the body frame of the vehicle is

$$\gamma_k^z = \omega_k^z \Delta t \tag{9.11}$$

The aim is to get the corresponding angle when projected on the E-N plane; i.e. the corresponding angle about the up direction of the local-level frame. The unit vector along the forward direction of the vehicle at time k observed from the body frame at time k is $U_{k|k}^b = [0 \quad 1 \quad 0]^T$. It is necessary to get this unit vector, which is along the forward direction of the vehicle at time k, observed from the body frame at time $k-1$; i.e. $U_{k|k-1}^b$. The rotation matrix from the body frame at time $k-1$ to that frame at time k due to rotating γ_k^z around the vertical axis of the vehicle is $R_z(\gamma_k^z)$. The relationship between $U_{k|k}^b$ and $U_{k|k-1}^b$ is

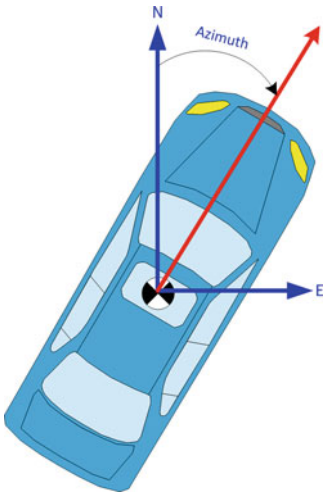


Fig. 9.6 The azimuth angle for a land vehicle

$$U^b_{k|k} = R_z\left(\gamma^z_k\right) U^b_{k|k-1} \tag{9.12}$$

and because $R_z\left(\gamma^z_k\right)$ is an orthogonal rotation matrix

$$U^b_{k|k-1} = \left(R_z\left(\gamma^z_k\right)\right)^T U^b_{k|k} = \begin{bmatrix} \cos\gamma^z_k & -\sin\gamma^z_k & 0 \\ \sin\gamma^z_k & \cos\gamma^z_k & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\sin\gamma^z_k \\ \cos\gamma^z_k \\ 0 \end{bmatrix} \tag{9.13}$$

The unit vector along the forward direction of the vehicle at time k seen from the local-level frame at time $k-1$ can be obtained from

$$U^\ell_{k|k-1} = \begin{bmatrix} U^E \\ U^N \\ U^{Up} \end{bmatrix} = R^\ell_{b,k-1} U^b_{k|k-1} = R^\ell_{b,k-1} \begin{bmatrix} -\sin\gamma^z_k \\ \cos\gamma^z_k \\ 0 \end{bmatrix} \tag{9.14}$$

Consequently the new heading from the north direction due to the angle γ^z_k is $\tan^{-1}\left(\frac{U^E}{U^N}\right)$, where

$$\begin{aligned} U^E &= \sin A_{k-1}\cos p_{k-1}\cos\gamma^z_k - \left(\cos A_{k-1}\cos r_{k-1} + \sin A_{k-1}\sin p_{k-1}\sin r_{k-1}\right)\sin\gamma^z_k \\ U^N &= \cos A_{k-1}\cos p_{k-1}\cos\gamma^z_k - \left(-\sin A_{k-1}\cos r_{k-1} + \cos A_{k-1}\sin p_{k-1}\sin r_{k-1}\right)\sin\gamma^z_k \end{aligned} \tag{9.15}$$

The azimuth angle defined by $\tan^{-1}\left(\frac{U^E}{U^N}\right)$ is the angle relative to the north direction and is positive in the clockwise direction.

In addition to the rotations performed by the vehicle, the angle γ^z_k has two additional components. These are due to the Earth's rotation and the change of orientation of the local-level frame. The part due to the Earth's rotation around the up direction is equal to $(\omega^e \sin\varphi_{k-1})\Delta t$ counterclockwise in the local-level frame;

ω^e is the Earth's rotation rate. This component is compensated directly from the new calculated heading to give the azimuth angle. It is worth mentioning that this component should be subtracted if the calculation is for the yaw angle (which is positive in the counterclockwise direction). In this study we calculate the azimuth angle directly (which is positive in the clockwise direction) and so the component for the Earth's rotation is added. The part of γ_k^z due to the change of orientation of the local-level frame with respect to the Earth from time epoch $k-1$ to k is in the counterclockwise direction and can be expressed as

$$\left.\frac{d\lambda}{dt}\right|_{k-1}(\sin\varphi_{k-1})\Delta t = \frac{v_{k-1}^E \sin\varphi_{k-1}}{(R_N + h_{k-1})\cos\varphi_{k-1}}\Delta t \tag{9.16}$$

The relationship between the vehicle's velocity in the body frame and in the local-level frame is

$$\begin{bmatrix} v_k^E \\ v_k^N \\ v_k^{Up} \end{bmatrix} = R_{b,k}^1 \begin{bmatrix} 0 \\ v_k^f \\ 0 \end{bmatrix} = \begin{bmatrix} v_k^f \sin A_k \cos p_k \\ v_k^f \cos A_k \cos p_k \\ v_k^f \sin p_k \end{bmatrix} \tag{9.17}$$

which enables Eq. (9.16) to be rewritten as

$$\left.\frac{d\lambda}{dt}\right|_{k-1}(\sin\varphi_{k-1})\Delta t = \frac{v_{k-1}^f \sin A_{k-1}\cos p_{k-1}\tan\varphi_{k-1}}{(R_N + h_{k-1})}\Delta t \tag{9.18}$$

and this must also be added in calculating the azimuth angle.

Finally, the model for the azimuth angle is

$$A_k = \tan^{-1}\left(\frac{U^E}{U^N}\right) + (\omega^e \sin\varphi_{k-1})\Delta t + \frac{v_{k-1}^f \sin A_{k-1}\cos p_{k-1}\tan\varphi_{k-1}}{(R_N + h_{k-1})}\Delta t \tag{9.19}$$

9.3.3.3 3D Position and Velocity Calculations

Before describing the system equations for position and velocity, take note that Eq. (9.17) gave the relation between the vehicle's velocity in the body frame and in the local-level frame. As a result of the non-holonomic constraint on land vehicles the velocity in the body frame consists only of the forward longitudinal speed of the vehicle; the transversal and vertical components are zeros. Hence the latitude can be expressed as

$$\varphi_k = \varphi_{k-1} + \left.\frac{d\varphi}{dt}\right|_k \Delta t = \varphi_{k-1} + \frac{v_k^N}{R_M + h_k}\Delta t = \varphi_{k-1} + \frac{v_k^f \cos A_k \cos p_k}{(R_M + h_k)}\Delta t \tag{9.20}$$

where R_M is the meridian radius of curvature of the Earth and Δt is the sampling time.

Similarly, the longitude is

$$\lambda_k = \lambda_{k-1} + \left.\frac{d\lambda}{dt}\right|_k \Delta t = \lambda_{k-1} + \frac{v_k^E}{(R_N + h_k)\cos\varphi_k}\Delta t = \lambda_{k-1} + \frac{v_k^f \sin A_k \cos p_k}{(R_N + h_k)\cos\varphi_k}\Delta t \tag{9.21}$$

where R_N is the normal radius of curvature of the Earth.

The altitude is

$$h_k = h_{k-1} + \left.\frac{dh}{dt}\right|_k \Delta t = h_{k-1} + v_k^{Up}\Delta t = h_{k-1} + v_k^f \sin p_k \Delta t \tag{9.22}$$

The forward speed is

$$v_k^f = v_k^{od} \tag{9.23}$$

9.3.4 Overview of 3D RISS Motion Model

The overall motion model is represented as follows

$$\begin{aligned}
\mathbf{x}_k &= \begin{bmatrix} \varphi_k \\ \lambda_k \\ h_k \\ v_k^f \\ p_k \\ r_k \\ A_k \end{bmatrix} = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_k) \\
&= \begin{bmatrix} \varphi_{k-1} + \dfrac{v_k^f \cos A_k \cos p_k}{R_M + h_k}\Delta t \\ \lambda_{k-1} + \dfrac{v_k^f k \cos p_k}{(R_N + h_k)\cos\varphi_k}\Delta t \\ h_{k-1} + v_k^f \sin p_k \Delta t \\ v_k^{od} \\ \sin^{-1}\left(\dfrac{f_k^y - a_k^{od}}{g}\right) \\ -\sin^{-1}\left(\dfrac{f_k^X + v_k^{od}\omega_k^z}{g\cos p_k}\right) \\ \tan^{-1}\left(\dfrac{U^E}{U^N}\right) + \omega^e \sin\varphi_{k-1}\Delta t + \frac{v_{k-1}^f \sin A_{k-1}\cos p_{k-1}\tan\varphi_{k-1}}{(R_N + h_{k-1})}\Delta t \end{bmatrix}
\end{aligned} \tag{9.24}$$

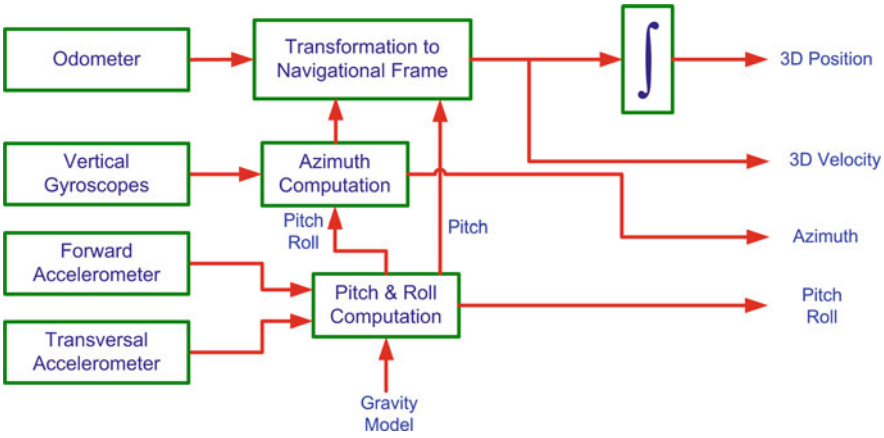


Fig. 9.7 A block diagram of the 3D RISS model

If the velocity components in the local-level frame are used in the navigation state vector instead of the forward speed, the motion model will become

$$\mathbf{x}_k = \begin{bmatrix} \varphi_k \\ \lambda_k \\ h_k \\ v_k^E \\ v_k^N \\ v_k^U \\ p_k \\ r_k \\ A_k \end{bmatrix} = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_k) = \begin{bmatrix} \varphi_{k-1} + \dfrac{v_k^N}{R_M + h_k}\Delta t \\ \lambda_{k-1} + \dfrac{v_k^E}{(R_N + h_k)\cos\varphi_k}\Delta t \\ h_{k-1} + v_k^U \Delta t \\ v_k^{od} \sin A_k \cos p_k \\ v_k^{od} \cos A_k \cos p_k \\ v_k^{od} \sin p_k \\ \sin^{-1}\left(\frac{f_k^y - a_k^{od}}{g}\right) \\ -\sin^{-1}\left(\frac{f_k^x + v_k^{od}\omega_k^z}{g\cos p_k}\right) \\ \tan^{-1}\left(\frac{U^E}{U^N}\right) + (\omega^e \sin\varphi_{k-1})\Delta t + \dfrac{v_{k-1}^E \tan\varphi_{k-1}}{(R_N + h_{k-1})}\Delta t \end{bmatrix} \tag{9.25}$$

The 3D RISS model is summarized in Fig. 9.7.

9.4 KF for Loosely Coupled 3D RISS/GPS Integration

As described in earlier chapters, the non-linear motion model forms the basis of mechanization and the linearized error-state model derived from this serves as the system model for a Kalman filter.

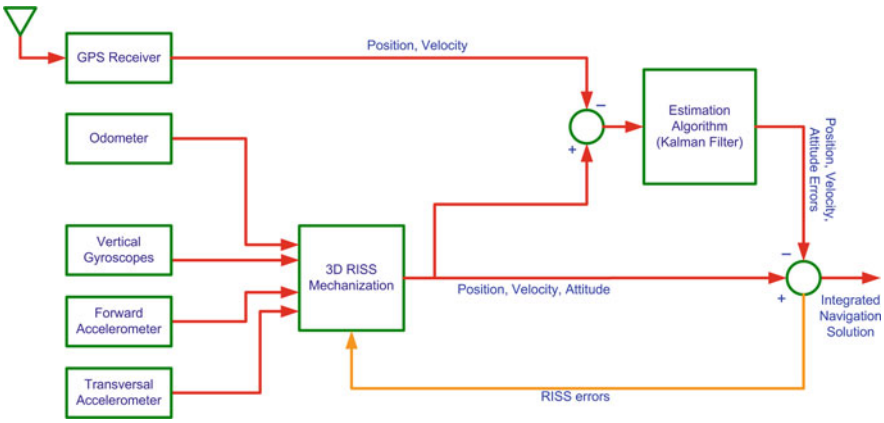


Fig. 9.8 A block diagram of a loosely coupled 3D RISS/GPS integration

Two approximations are made here to simplify the following linearized system model as an error model for a 3D RISS. The first approximation is to exclude the pitch and roll errors as states in the integration filters. This is acceptable because, as discussed earlier, the errors in pitch and roll do not grow with time since they do not involve integration. The second approximation assumes that the gyroscope (along the vertical direction of the body frame of the vehicle) is aligned with the vertical of the local-level frame, simplifying the azimuth calculation. Since this gyroscope measurement includes the component of the Earth's rotation as well as the rotation of the local-level frame on the Earth's curvature these quantities are removed from the measurement prior to integration. Based upon the above 2D approximation for the azimuth, we can write the azimuth angle directly in local-level frame as

$$A_k = A_{k-1} - \omega_z \Delta t + (\omega^e \sin \varphi_{k-1})\Delta t + \frac{v^E_{k-1} \tan \varphi_{k-1}}{(R_N + h_{k-1})}\Delta t \tag{9.26}$$

The block diagram of the loosely coupled 3D RISS/GPS integrated system is depicted in Fig. 9.8.

9.4.1 The Linearized Error Model for 3D RISS

The error-state system model for 3D RISS KF can be written as

$$\boldsymbol{\delta x}_k = \Phi_{k-1}\boldsymbol{\delta x}_{k-1} + G_{k-1}w_{k-1} \tag{9.27}$$

where

$\boldsymbol{\delta x}_k$	is the 9×1 state vector
Φ_{k-1}	is the 9×9 state transition matrix
G_{k-1}	is the 9×1 noise distribution matrix
w_{k-1}	is the unit-variance white Gaussian noise.

The error-state vector that will be used by the Kalman filter is

$$\delta\mathbf{x}_k = \left[\delta\varphi_k, \delta\lambda_k, \delta h_k, \delta v_k^E, \delta v_k^N, \delta v_k^U, \delta A_k, SF_k^{od}, \delta\omega_k^z\right]^T \tag{9.28}$$

where

$\delta\varphi_k$	is the error in latitude
$\delta\lambda_k$	is the error in longitude
δh_k	is the error in altitude
δv_k^E	is the error in the velocity component along east direction
δv_k^N	is the error in the velocity component along north direction
δv_k^U	is the error in the velocity component along up direction
δA_k	is the error in the azimuth angle
SF_k^{od}	is the scale factor error of the odometer-derived speed
$\delta\omega_k^z$	is the stochastic gyroscope drift.

As mentioned earlier, the motion model is non-linear and must be linearized to obtain the error model that will serve as the system model for KF. These equations are linearized by applying a Taylor series expansion and ignoring the higher order terms. The corresponding linearized error state system model is therefore

$$\delta\mathbf{x}_k = \begin{bmatrix}\delta\mathbf{r}_k\\ \delta\mathbf{v}_k\\ \delta\mathbf{e}_k\end{bmatrix} = \begin{bmatrix}I_{3\times3} & F_1 & \mathbf{0}_{3\times3}\\ \mathbf{0}_{3\times3} & I_{3\times3} & F_2\\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & F_3\end{bmatrix}\begin{bmatrix}\delta\mathbf{r}_{k-1}\\ \delta\mathbf{v}_{k-1}\\ \delta\mathbf{e}_{k-1}\end{bmatrix} + \begin{bmatrix}\mathbf{0}_{3\times1}\\ \mathbf{0}_{3\times1}\\ \boldsymbol{\sigma}\end{bmatrix} \tag{9.29}$$

where

$$\delta\mathbf{r}_k = \begin{bmatrix}\delta\varphi_k\\ \delta\lambda_k\\ \delta h_k\end{bmatrix}$$

$$\delta\mathbf{v}_k = \begin{bmatrix}\delta v_k^E\\ \delta v_k^N\\ \delta v_k^U\end{bmatrix}$$

$$\delta\mathbf{e}_k = \begin{bmatrix}\delta A_k\\ SF_k^{od}\\ \delta\omega_k^Z\end{bmatrix}$$

$$\delta\boldsymbol{\sigma}_k = \begin{bmatrix}0\\ \sqrt{2\gamma_{od}\sigma_{od}^2\Delta t}\\ \sqrt{2\beta_Z\sigma_Z^2\Delta t}\end{bmatrix}$$

$$F_1 = \begin{bmatrix}0 & \frac{1}{R_M+h_{k-1}}\Delta t & 0\\ \frac{1}{(R_N+h_{k-1})\cos\varphi_{k-1}}\Delta t & 0 & 0\\ 0 & 0 & \Delta t\end{bmatrix}$$

$$F_2 = \begin{bmatrix} v_{k-1}^{od}\cos A_{k-1}\cos p_{k-1} & v_{k-1}^{od}\sin A_{k-1}\cos p_{k-1} & 0 \\ -v_{k-1}^{od}\sin A_{k-1}\cos p_{k-1} & v_{k-1}^{od}\cos A_{k-1}\cos p_{k-1} & 0 \\ 0 & v_{k-1}^{od}\sin p_{k-1} & 0 \end{bmatrix}$$

$$F_3 = \begin{bmatrix} 1 & 0 & -\Delta t \\ 0 & 1-\gamma_{od}\Delta t & 0 \\ 0 & 0 & 1-\beta_z\Delta t \end{bmatrix}.$$

The stochastic errors associated with the gyroscope and the odometer-derived speed are modeled by a Gauss-Markov model where γ_{od} is the reciprocal of the autocorrelation time for the scale factor of the odometer-derived speed, σ_{od}^2 is the variance of the noise associated with it, β_z is the reciprocal of the autocorrelation time for the gyroscope's stochastic drift, and σ_z^2 is the variance of the noise associated with it.

9.4.2 Measurement Model for Updating 3D RISS

Since we are considering loosely coupled integration, GPS position and velocity updates are used during the update stage of the KF and the measurement model used by the filter is

$$\begin{aligned} \delta\mathbf{z}_k &= \begin{bmatrix} \varphi_k^{RISS} - \varphi_k^{GPS} \\ \lambda_k^{RISS} - \lambda_k^{GPS} \\ h_k^{RISS} - h_k^{GPS} \\ v_k^{E,RISS} - v_k^{E,GPS} \\ v_k^{N,RISS} - v_k^{N,GPS} \\ v_k^{U,RISS} - v_k^{U,GPS} \end{bmatrix} = H\delta\mathbf{x}_k + \delta\boldsymbol{\eta}_k \\ &= \begin{bmatrix} 1&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0 \end{bmatrix} \begin{bmatrix} \delta\varphi_k \\ \delta\lambda_k \\ \delta h_k \\ \delta v_k^E \\ \delta v_k^N \\ \delta v_k^U \\ \delta A_k \\ SF_k^{od} \\ \delta\omega_k^z \end{bmatrix} + \begin{bmatrix} \delta\boldsymbol{\eta}_k^{\varphi} \\ \delta\boldsymbol{\eta}_k^{\lambda} \\ \delta\boldsymbol{\eta}_k^{h} \\ \delta\boldsymbol{\eta}_k^{v^E} \\ \delta\boldsymbol{\eta}_k^{v^N} \\ \delta\boldsymbol{\eta}_k^{v^U} \end{bmatrix} \end{aligned} \tag{9.30}$$

where $\delta\mathbf{z}_k$ is the difference between the 3D positions and velocities provided by RISS mechanization and by GPS, and $\delta\boldsymbol{\eta}_k$ is the measurement noise.

9.5 KF for Tightly Coupled 3D RISS/GPS Integration

The block diagram of the tightly coupled 3D RISS/GPS integrated system is depicted in Fig. 9.9.

9.5.1 Augmenting the System Model

The system model for a tightly coupled system integration is like that for a loosely coupled one, but both it and the state vector also include the GPS receiver's clock bias and drift errors and therefore the RISS and GPS parts are simply combined to provide the full system model

$$\begin{aligned}\delta\mathbf{x}_k &= \Phi_{k-1}\delta\mathbf{x}_{k-1} + G_{k-1}w_{k-1} \\ &= \begin{bmatrix}\delta\mathbf{x}_k^{RISS} \\ \delta\mathbf{x}_k^{GPS}\end{bmatrix} = \begin{bmatrix}\Phi_{k-1}^{RISS} & 0 \\ 0 & \Phi_{k-1}^{GPS}\end{bmatrix}\begin{bmatrix}\delta\mathbf{x}_{k-1}^{RISS} \\ \delta\mathbf{x}_{k-1}^{GPS}\end{bmatrix} + \begin{bmatrix}G_{k-1}^{RISS} \\ G_{k-1}^{GPS}\end{bmatrix}w_{k-1}\end{aligned} \tag{9.31}$$

The equation for the KF system model of a GPS is

$$\delta\mathbf{x}_k^{GPS} = \Phi_{k-1}^{GPS}\delta\mathbf{x}_{k-1}^{GPS} + G_{k-1}^{GPS}w_{k-1} \tag{9.32}$$

The bias of the GPS receiver's clock δb_k^{GPS} and its drift δd_k^{GPS} are included as states. The system model for these states is

$$\begin{bmatrix}\delta b_k^{GPS} \\ \delta d_k^{GPS}\end{bmatrix} = \begin{bmatrix}1 & \Delta t \\ 0 & 1\end{bmatrix}\begin{bmatrix}\delta b_{k-1}^{GPS} \\ \delta d_{k-1}^{GPS}\end{bmatrix} + \begin{bmatrix}\sigma_b\Delta t \\ \sigma_d\Delta t\end{bmatrix}w_{k-1}^{GPS} \tag{9.33}$$

where

w_{k-1}^{GPS} is unit-variance white Gaussian noise
σ_b is the standard deviation of white noise for the clock bias
σ_d is the standard deviation of white noise for the clock drift.

and the state vector becomes

$$\delta\mathbf{x}_k = \left[\delta\varphi_k, \delta\lambda_k, \delta h_k, \delta v_k^E, \delta v_k^N, \delta v_k^U, \delta A_k, SF_k^{od}, \delta\omega_k^z, \delta b_k^{GPS}, \delta d_k^{GPS}\right]^T$$

9.5.2 Raw GPS Measurement Model for Updating 3D RISS

The linearized measurement model for KF in tightly coupled integration is

$$\delta\mathbf{z}_k = H_k\delta\mathbf{x}_k + \boldsymbol{\varepsilon}_k \tag{9.34}$$

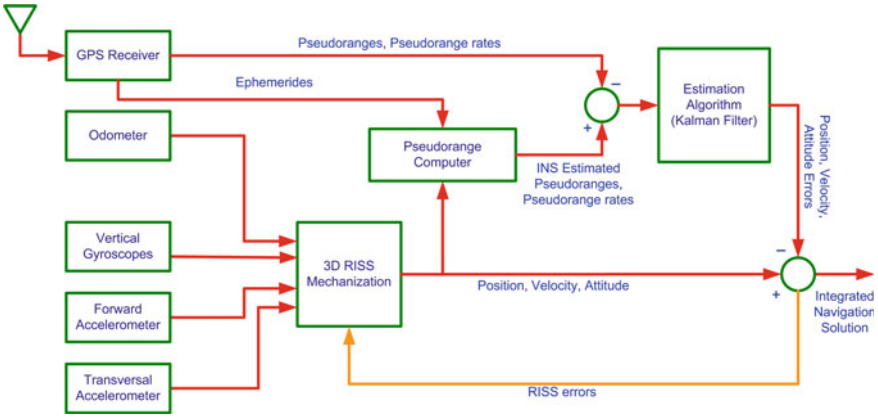


Fig. 9.9 A block diagram of a tightly coupled 3D RISS/GPS integration

where

$\delta \mathbf{z}_k$ is the measurement vector
H_k is the measurement design matrix
$\delta \mathbf{x}_k$ is a state vector of dimension 11×1
$\boldsymbol{\varepsilon}_k$ is a vector of measurement random noise which has a zero-mean and covariance R.

The GPS observations are pseudo-ranges and pseudo-range rate measurements. Hence for the error state KF, the measurement vector is the difference between the RISS and GPS pseudo-ranges $\delta \mathbf{z}_k^{\rho}$ and pseudo-range rates $\delta \mathbf{z}_k^{\dot{\rho}}$, expressed as

$$\delta \mathbf{z}_k = \begin{bmatrix} \delta \mathbf{z}_k^{\rho} \\ \delta \mathbf{z}_k^{\dot{\rho}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\rho}_k^{RISS} - \boldsymbol{\rho}_k^{GPS} \\ \dot{\boldsymbol{\rho}}_k^{RISS} - \dot{\boldsymbol{\rho}}_k^{GPS} \end{bmatrix} \tag{9.35}$$

where

$\boldsymbol{\rho}_k^{RISS}$ is the RISS estimated range between the satellite and the receiver
$\dot{\boldsymbol{\rho}}_k^{RISS}$ is the RISS estimated range rate between the satellite and the receiver
$\boldsymbol{\rho}_k^{GPS}$ is the GPS measured range between the satellite and the receiver
$\dot{\boldsymbol{\rho}}_k^{GPS}$ is the GPS measured range rate between the satellite and the receiver.

As pointed out in Chap. 3, $\boldsymbol{\rho}_k^{GPS}$ is corrected for satellite clock errors and for both ionospheric and tropospheric errors.

For M satellites visible to the receiver, the above equation can be written as

$$\begin{bmatrix} \delta z_k^{\rho,1} \\ \delta z_k^{\rho,2} \\ \vdots \\ \delta z_k^{\rho,M} \\ \delta z_k^{\dot{\rho},1} \\ \delta z_k^{\dot{\rho},2} \\ \vdots \\ \delta z_k^{\dot{\rho},M} \end{bmatrix} = \begin{bmatrix} \rho_k^{RISS,1} - \rho_k^{GPS,1} \\ \rho_k^{RISS,2} - \rho_k^{GPS,2} \\ \vdots \\ \rho_k^{RISS,M} - \rho_k^{GPS,M} \\ \dot{\rho}_k^{RISS,1} - \dot{\rho}_k^{GPS,1} \\ \dot{\rho}_k^{RISS,2} - \dot{\rho}_k^{GPS,2} \\ \vdots \\ \dot{\rho}_k^{RISS,M} - \dot{\rho}_k^{GPS,M} \end{bmatrix} \tag{9.36}$$

These non-linear pseudo-range errors are derived from the measured pseudo-range and the predicted pseudo-range based on the RISS calculated position. The pseudo-range for the mth satellite $\rho_k^{GPS,M}$ can therefore be modeled by

$$\rho_k^{GPS,m} = \left\| \mathbf{r}_k - \mathbf{r}_k^m \right\| + \delta b_k^{GPS} + \varepsilon_k^{\rho,m} \tag{9.37}$$

where

$\mathbf{r}_k$	is the receiver's position in ECEF rectangular coordinates
$\mathbf{r}_k^m$	is position of the mth satellite in ECEF rectangular coordinates
δb_k^{GPS}	is the receiver's clock bias
$\varepsilon_k^{\rho,m}$	is the total effect of residual errors due to atmospheric delays (after correcting for these errors using models inside the receiver itself), receiver noise, etc.

The estimated pseudo-range from the output of the RISS navigation system is defined as

$$\rho_k^{RISS,m} = \left\| \mathbf{r}_k^{RISS} - \mathbf{r}_k^m \right\| \tag{9.38}$$

where $\mathbf{r}_k^{RISS}$ is the position of the vehicle calculated from the position output of the RISS mechanization originally in ECEF geodetic coordinates and transformed to ECEF rectangular coordinates.

The GPS pseudo-range measurements are not linearly related to the vehicle's position but a Taylor expansion series around the latest estimated state allows the difference between the RISS estimate and the GPS measurement to be modeled as being linearly related to the error in position

$$\delta z_k^{\rho,m} = \mathbf{1}_k^{RISS,m} \cdot \left(\mathbf{r}_k^{RISS} - \mathbf{x}_k^m \right) - \delta b_k^{GPS} + \tilde{\varepsilon}_k^{\rho,m} \tag{9.39}$$

where $\mathbf{1}_k^{RISS,m}$ is the line of sight unit vector from the mth satellite to the position of the receiver based on the output of RISS mechanization

$$\mathbf{1}_k^{RISS,m} = \frac{\mathbf{r}_k^{RISS} - \mathbf{r}_k^m}{\left\| \mathbf{r}_k^{RISS} - \mathbf{r}_k^m \right\|} \tag{9.40}$$

In component form, Eq. (9.39) can be written as

$$\delta z_k^{\rho,m} = \begin{bmatrix} \mathbf{1}_{\mathrm{x},k}^{RISS,m} & \mathbf{1}_{y,k}^{RISS,m} & \mathbf{1}_{z,k}^{RISS,m} \end{bmatrix} \begin{bmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{bmatrix} - \delta b_k^{GPS} + \tilde{\varepsilon}_k^{\rho,m} \tag{9.41}$$

where $\delta \mathbf{r}_k = \mathbf{r}_k^{RISS}{}_{RISS} - \mathbf{r}_k^m$.

With M satellites visible to the receiver, this becomes

$$\delta \mathbf{z}_{k,M\times 1}^{\rho} = G_{k,M\times 3} \begin{bmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{bmatrix}_{3\times 1} - \begin{bmatrix} \delta b_k^{GPS} \\ \delta_k^{GPS} \\ \vdots \\ \delta b_k^{GPS} \end{bmatrix}_{M\times 1} + \begin{bmatrix} \tilde{\varepsilon}_k^{\rho,1} \\ \tilde{\varepsilon}_k^{\rho,2} \\ \vdots \\ \tilde{\varepsilon}_k^{\rho,m} \end{bmatrix}_{M\times 1} \tag{9.42}$$

where

$$G_k = \begin{bmatrix} 1_{x,k}^{RISS,1} & 1_{y,k}^{RISS,1} & 1_{z,k}^{RISS,1} \\ 1_{x,k}^{RISS,2} & 1_{y,k}^{RISS,2} & 1_{z,k}^{RISS,2} \\ \vdots & \vdots & \vdots \\ 1_{x,k}^{RISS,M} & 1_{y,k}^{RISS,M} & 1_{Z,k}^{RISS,M} \end{bmatrix}_{M\times 3} \tag{9.43}$$

Since the position components of the state vector are in geodetic coordinates in the ECEF frame, it is necessary to convert them to the corresponding rectangular coordinates

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} (R_N + h_k)\cos\varphi_k \cos\lambda_k \\ (R_N + h_k)\cos\varphi_k \sin\lambda_k \\ \{R_N(1-e^2) + h_k\}\sin\varphi_k \end{bmatrix} \tag{9.44}$$

In matrix notation, the linearized position error form of the above equation is

$$\begin{bmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{bmatrix} = \begin{bmatrix} -(R_N + h_k)\sin\varphi_k \cos\lambda_k & -(R_N + h_k)\cos\varphi_k \sin\lambda_k & \cos\varphi_k \cos\lambda_k \\ -(R_N + h_k)\sin\varphi_k \sin\lambda_k & (R_N + h_k)\cos\varphi_k \cos\lambda_k & \cos\varphi_k \sin\lambda_k \\ \{R_N(1-e^2) + h_k\}\cos\varphi_k & 0 & \sin\varphi_k \end{bmatrix} \begin{bmatrix} \delta\varphi_k \\ \delta\lambda_k \\ \delta h_k \end{bmatrix}$$
$$\begin{bmatrix} \delta x_k \\ \delta y_k \\ \delta z_k \end{bmatrix} = L_{k,3\times 3} \begin{bmatrix} \delta\varphi_k \\ \delta\lambda_k \\ \delta h_k \end{bmatrix} \tag{9.45}$$

Substituting Eq. (9.45) into Eq. (9.42) gives

$$\delta \mathbf{z}_{k,M\times 1}^{\rho} = G_{k,M\times 3} L_{k,3\times 3} \begin{bmatrix} \delta\varphi_k \\ \delta\lambda_k \\ \delta h_k \end{bmatrix} - \delta \mathbf{b}_{k,M\times 1}^{GPS} + \tilde{\boldsymbol{\varepsilon}}_{k,M\times 1}^{\rho} \tag{9.46}$$

By defining a matrix H_k^{ρ} such that $H_{k,M\times 3}^{\rho} = G_{k,M\times 3} L_{k,3\times 3}$, the equation for the linearized pseudo-range measurement model is

$$\delta \mathbf{z}^{\rho}_{k,M\times 1} = H^{\rho}_{k,M\times 3}\begin{bmatrix}\delta\varphi_k\\ \delta\lambda_k\\ \delta h_k\end{bmatrix}_{3\times 1} - \delta \mathbf{b}^{GPS}_{k,M\times 1} + \tilde{\boldsymbol{\varepsilon}}^{\rho}_{k,M\times 1} \tag{9.47}$$

The pseudo-range rate measurement from the mth satellite, $\dot{\rho}_k^{GPS,m}$, is computed by measuring the Doppler shift in the satellite's carrier frequency. As noted in Chap. 3 the pseudo-range rate can be treated as the relative velocity measurement between the receiver and the satellite projected onto the direction of the unit vector from the satellite to the receiver. This measurement model can be mathematically written as

$$\dot{\rho}_k^{GPS,m} = \left(\mathbf{v}_k - \mathbf{v}_k^m\right)\cdot \mathbf{1}_k^m + \delta d_k^{GPS} + \boldsymbol{\varepsilon}_k^{\dot{\rho},m} \tag{9.48}$$

where

$\mathbf{v}_k^m$	is the mth satellite's velocity in the ECEF frame
$\mathbf{v}_k$	is the true velocity of the receiver in the ECEF frame
δd_k^{GPS}	is the receiver's clock drift
$\boldsymbol{\varepsilon}_k^{\dot{\rho},m}$	is the error in observation.

In Eq. (9.48), $\mathbf{1}_k^m$ is the true line of sight unit vector pointing from the mth satellite to the receiver, which is

$$\mathbf{1}_k^m = \frac{\mathbf{x}_k - \mathbf{x}_k^m}{\left\|\mathbf{x}_k - \mathbf{x}_k^m\right\|} \tag{9.49}$$

Equation (9.48) can be rewritten as

$$\dot{\rho}_k^{GPS,m} = \left(\mathbf{v}_k - \mathbf{v}_k^m\right)\cdot \frac{\mathbf{x}_k - \mathbf{x}_k^m}{\left\|\mathbf{x}_k - \mathbf{x}_k^m\right\|} + \delta d_k^{GPS} + \boldsymbol{\varepsilon}_k^{\dot{\rho},m} \tag{9.50}$$

The RISS estimated pseudo-range rate is

$$\dot{\rho}_k^{RISS,m} = \left(\mathbf{v}_k^{RISS} - \mathbf{v}_k^m\right)\cdot \mathbf{1}_k^{RISS,m} \tag{9.51}$$

where $\mathbf{v}_k^{RISS}$ is the velocity of the vehicle estimated by the RISS in rectangular coordinates in the ECEF frame.

The linearized pseudo-range rate measurement equation is therefore

$$\dot{\rho}_k^{RISS,m} - \dot{\rho}_k^{GPS,m} = \left(\mathbf{1}_k^{RISS,m}\right)^T \delta \mathbf{v}_k - \delta d_k^{GPS} + \tilde{\boldsymbol{\varepsilon}}_k^{\dot{\rho},m} \tag{9.52}$$

$$\delta z_k^{\dot{\rho},m} = \begin{bmatrix} 1_{x,k}^{RISS,m} & 1_{y,k}^{RISS,m} & 1_{z,k}^{RISS,m}\end{bmatrix}\begin{bmatrix}\delta v_k^x\\ \delta v_k^y\\ \delta v_k^z\end{bmatrix} - \delta d_k^{GPS} + \tilde{\boldsymbol{\varepsilon}}_k^{\dot{\rho},m} \tag{9.53}$$

where $\delta \mathbf{v}_k = \mathbf{v}_k^{RISS} - \mathbf{v}_k$.

For M satellites visible to the receiver, the pseudo-range rate measurement of Eq. (9.53) can be expressed as

$$\delta \mathbf{z}_{k,M\times 1}^{\dot{\rho}} = G_{k,M\times 3} \begin{bmatrix} \delta v_k^x \\ \delta v_k^y \\ \delta v_k^z \end{bmatrix}_{3\times 1} - \begin{bmatrix} \delta d_k^{GPS} \\ \delta d_k^{GPS} \\ \vdots \\ \delta d_k^{GPS} \end{bmatrix}_{M\times 1} + \begin{bmatrix} \tilde{\varepsilon}_k^{\dot{\rho},1} \\ \tilde{\varepsilon}_k^{\dot{\rho},2} \\ \vdots \\ \tilde{\varepsilon}_k^{\dot{\rho},m} \end{bmatrix}_{M\times 1} \tag{9.54}$$

The relationship between the velocity in the local-level frame and the ECEF frame is

$$\begin{bmatrix} \delta v_k^x \\ \delta v_k^y \\ \delta v_k^z \end{bmatrix} = R_l^e|_k \begin{bmatrix} \delta v_k^E \\ \delta v_k^N \\ \delta v_k^U \end{bmatrix} \tag{9.55}$$

where $R_l^e\big|_k$ transforms from the local-level frame to the ECEF frame, and can be written as

$$R_l^e\big|_k \begin{bmatrix} -\sin\lambda_k & \sin\varphi_k \cos\lambda_k & \cos\varphi_k \cos\lambda_k \\ \cos\lambda_k & \sin\varphi_k \sin\lambda_k & \cos\varphi_k \sin\lambda_k \\ 0 & \cos\varphi_k & \sin\varphi_k \end{bmatrix} \tag{9.56}$$

By substituting Eqs. (9.55) into (9.54) and equating $H_{k,M\times 3}^{\dot{\rho}} = G_{k,M\times 3} R_l^e\big|_k$, the expression for pseudo-range rate measurement model is

$$\delta \mathbf{z}_{k,M\times 1}^{\dot{\rho}} = \mathrm{H}_{k,M\times 3}^{\dot{\rho}} \begin{bmatrix} \delta v_k^E \\ \delta v_k^N \\ \delta v_k^U \end{bmatrix} - \delta \mathbf{d}_{k,M\times 1}^{GPS} + \tilde{\boldsymbol{\varepsilon}}_{k,M\times 1}^{\dot{\rho}} \tag{9.57}$$

The overall measurement model for the pseudo-ranges and pseudo-range rates is therefore

$$\begin{aligned} \delta \mathbf{z}_{k,2M\times 1} = & \begin{bmatrix} H_{k,M\times 3}^{\rho} & 0_{M\times 3} & 0_{M\times 3} & -1_{M\times 1} & 0_{M\times 1} \\ 0_{M\times 3} & H_{k,M\times 3}^{\dot{\rho}} & 0_{M\times 3} & 0_{M\times 1} & -1_{M\times 1} \end{bmatrix}_{2M\times 11} \delta \mathbf{x}_{k,11\times 1} \\ & + \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_{k,M\times 1}^{\rho} \\ \tilde{\boldsymbol{\varepsilon}}_{k,M\times 1}^{\dot{\rho}} \end{bmatrix}_{2M\times 1} \end{aligned} \tag{9.58}$$

References

Brandt A, Gardner JF (1998) Constrained navigation algorithms for strapdown inertial navigation systems with reduced set of sensors. In: Proceedings of the 1998 American control conference (ACC), June 24–26, 1998, Philadelphia, PA, American Automatic Control Council, pp 1848–1852 BN - 1840 7803 4530 1844

Georgy J, Noureldin A, Korenberg M, Bayoumi M (2010) Low-cost three-dimensional navigation solution for RISS/GPS integration using mixture particle filter. IEEE Trans Veh Technol 59(2):599–615

Iqbal U, Karamat TB, Okou AF, Noureldin A (2009) Experimental results on an integrated GPS and multisensor system for land vehicle positioning. Int J Navig Obs, Hindawi Publishing Corporation. doi:10.1155/2009/765010, Article ID 765010

Iqbal U, Noureldin A (2009) Integrated reduced inertial sensor system/GPS for vehicle navigation. VDM Verlag, Germany

Iqbal U, Okou F, Noureldin A (2008) An integrated reduced inertial sensor system-RISS/GPS for land vehicles. In: Proceeding of IEEE/ION position, location and navigation symposium, Monterey, CA, pp 1014–1021, May 2008

Noureldin A, Irvine-Halliday D, Mintchev MP (2004) Measurement-while-drilling surveying of highly inclined and horizontal well sections utilizing single-axis gyro sensing system. Meas Sci Technol 15(12):2426–2434

Noureldin A, Irvine-Halliday D, Mintcheve MP (2002) Accuracy limitations of FOG-based continuous measurement-while-drilling surveying instruments for horizontal wells. IEEE Trans Instrum Meas 51(6):1177–1190

Skog I, Handel P (2009) In-car positioning and navigation technologies—a survey. Intell Transp Syst, IEEE Trans 10(1):4–21

Chapter 10
Two Case Studies: Full IMU/GPS and 3D RISS/GPS Integration

In Chap. 8 the theories of the loosely coupled and tightly coupled integration schemes of a full IMU with GPS were presented. In Chap. 9 the integration of a reduced inertial sensor system (RISS) with GPS was detailed, along with its various advantages. In this chapter we will look at the performance of these integration techniques using real inertial measurements and GPS data collected during road test trajectories. After the data had been obtained, some GPS outages were intentionally simulated during post-processing whereby the number of satellites available was limited to fewer than four and the response of the INS/GPS integration was assessed for accuracy against a reference trajectory. When the number of satellites was reduced to zero in the tightly coupled integration algorithm, a complete blockage of the GPS, its performance resembled that of a loosely coupled integration because the loosely coupled integration only suffers a complete blockage (no GPS solution) when the number of satellites falls below four, denying aiding to the INS. In the strict sense this is not a loosely coupled integration but the performance is so close that we will treat it as if it were. The ensuing sections describe the equipment used for the experiments and an overview of the trajectories. The results of the trajectories are assessed for consistency with the theory and expected improvement or degradation for various satellite configurations. Selected GPS outages from each trajectory are also analyzed.

10.1 Navigation Equipment Used for the Experiments

The full IMU used a tactical grade Honeywell HG1700 model AG58 IMU housed in an IMU-G2 enclosure by NovAtel that provides the power supply and decodes and times the IMU output data. In this enclosure the unit is known as an IMU-G2-H58.

A. Noureldin et al., *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, DOI: 10.1007/978-3-642-30466-8_10,

Table 10.1 Characteristics of the Crossbow IMU300CC (Crossbow 2012) and Honeywell IMU (NovAtel 2005)

	Crossbow IMU300CC	HG1700
Size	7.62 × 9.53 × 3.2 (cm)	15 × 15 × 10 (cm)
Weight	0.59 kg	0.725 kg
Max data rate	200 Hz	100 Hz
Start-up time	<1 s	<0.8 s
	Accelerometer	
Range	±2 g	±50 g
Bias	±30 mg	1.0 mg
Scale factor	<1 %	300 ppm
Random walk	<0.15 m/s/h$^{1/2}$	0.0198 m/s/h$^{1/2}$
	Gyroscope	
Range	±100 g	±1,000 deg/sec
Bias	< ±2.0°/sec	1°/h
Scale factor	<1 %	150 ppm
Random walk	<2.25°/h$^{1/2}$	0.125°/h$^{1/2}$
Electrical		
Input voltage	9–30 V dc	±5 V dc
Power	<3 W	<8 W
Connector	RS-232	RS-422

To assist the reader appreciate low cost sensor performance, a MEMS grade Crossow IMU300CC (Crossbow 2007) was used for the 3D RISS solution. This is an inertial system with six degrees of freedom which employs solid state devices to measure angular rate and linear acceleration. The specifications of these two IMUs are listed in Table 10.1. For the experiment, only data from the vertical gyroscope was needed. For the RISS algorithm we used the speed data from the vehicle collected by a data logger called CarChip (Davis 2012) that was connected to the OBDII interface of the vehicle. A NovAtel ProPak-G2plus GPS receiver housing an OEM4-G2 GPS card was used for external aiding. A single frequency output of the GPS receiver was selected for the algorithm to integrate the IMU with the GPS. For the reference system an off-the-shelf NovAtel SPAN technology system was employed. This integrates the HG1700 IMU and ProPak-G2*plus* GPS receiver through tightly coupled architecture using both L1 and L2 frequencies. The accuracy of the reference system is 2 cm with the aid of RT-2 GPS software and 1.5 m for single point (L1/L2) positioning. For laboratory calibration of the IMU an Ideal Aerosmith 2103HT rate table was used. All of this equipment is shown in Fig. 10.1.

The systems were mounted on a test vehicle and various trajectories carried out to address different road scenarios typically encountered in a road trip. Figure 10.2 shows the data acquisition equipment installed in the vehicle prior to the test.

Fig. 10.1 **a** Honeywell HG1700 IMU enclosed in NovAtel casing (together referred to as IMU-G2-H58). **b** Crossbow IMU300CC. **c** NovAtel OEM4 GPS enclosed in ProPak G-2Plus SPAN unit, and **d** Ideal Aerosmith 2103HT rate table

10.1.1 Partial GPS Outage Criterion

As mentioned, the main advantage of tightly coupled integration is that it provides a GPS solution for aiding INS even when fewer than four satellites are visible. In reality, the satellites with the lowest elevation angles are usually blocked. Figure 10.3 depicts typical urban and rural scenarios. To mimic real life, only the highest elevation angle satellites were selected in the test.

Fig. 10.2 The data collection equipment mounted inside the road test vehicle

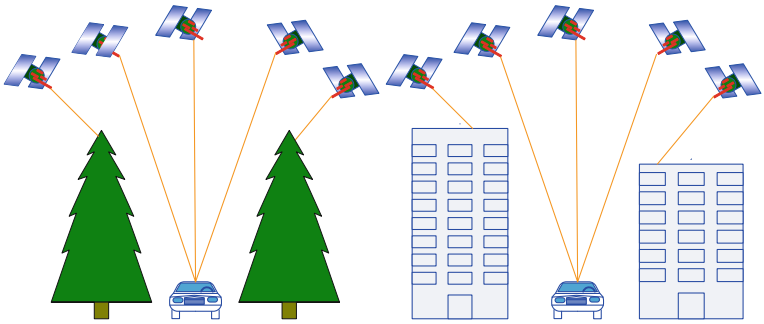


Fig. 10.3 GPS signal blockage in rural (*left*) and urban 'canyons' (*right*)

10.2 Performance of Tightly Coupled Algorithm for Full IMU/GPS

The trajectory selected for the analysis of the full IMU tightly coupled algorithm was performed in downtown Kingston, Ontario, an area that included slow speeds, frequent stops and many sharp turns. The start and end points of the trajectory are indicated by a red triangle and circle respectively in Fig. 10.4. The direction of the trajectory is indicated by a blue arrow. Five intentionally introduced partial GPS outages (with fewer than four visible satellites) were introduced at locations shown by the blue circles. It was ensured that these GPS outages were introduced at

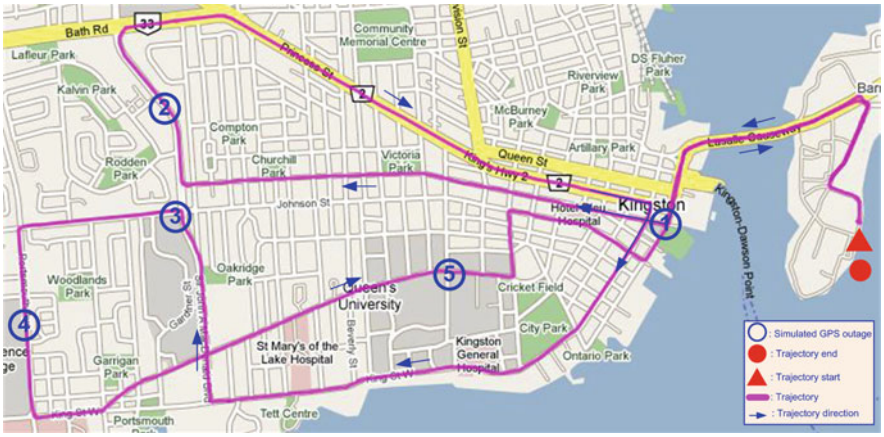


Fig. 10.4 The road test trajectory in and around downtown Kingston, Ontario

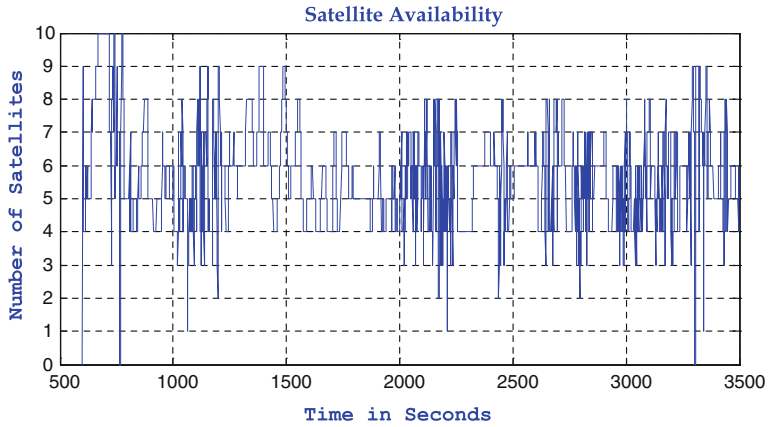


Fig. 10.5 Availability of satellites during the trajectory for the full IMU test

places which contained different features; e.g. slow speeds, normal turns, sharp turns, high speeds, and slopes. The outage duration was fixed at 60 s and each outage was repeated four times with the number of visible satellites being decreased progressively to zero.

Figure 10.5 shows the availability of satellites during the test. It is evident that there are portions of the trajectory where the number of satellites fell below three. Since the outages were simulated in post-processing in order to limit the number of satellites, we could use only those parts of the trajectory for which there were at least three satellites visible.

Table 10.2 lists the position errors during the five partial GPS outages, each simulated by limiting the number of satellite to three, two, one and zero. It shows both the maximum as well as the RMS errors (in meters) for 60 s partial as well as

Table 10.2 Maximum and RMS position errors for partial outages of a GPS that is tightly coupled with a full IMU

Outage number	Maximum and RMS position error (in meters)							
	3 satellites		2 satellites		1 satellite		0 satellite	
	Max	RMS	Max	RMS	Max	RMS	Max	RMS
1	4.72	3.79	12.20	6.51	21.97	11.85	21.99	11.86
2	4.96	2.75	6.15	3.29	11.72	7.52	11.65	7.50
3	8.74	7.10	16.42	10.00	25.38	14.47	25.81	14.69
4	2.81	2.41	7.14	4.06	14.24	8.28	11.20	6.67
5	14.50	11.09	19.62	13.21	37.96	21.61	28.82	16.36
Average	*7.15*	*5.43*	*12.30*	*7.41*	*22.25*	*12.75*	*19.89*	*11.42*

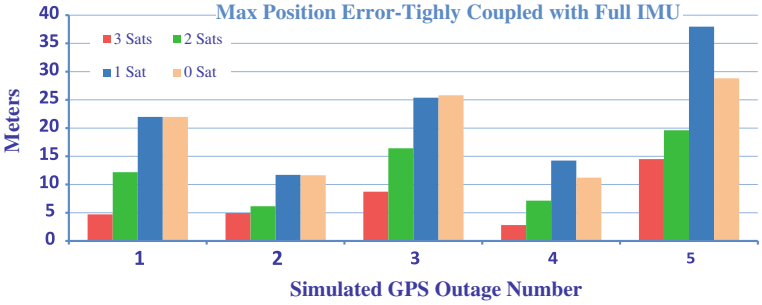


Fig. 10.6 A bar graph of the maximum positional errors for partial outages of a GPS that is tightly coupled with a full IMU

full GPS blockage (zero satellites). The data in Table 10.2 is plotted in the form of a bar graph in Fig. 10.6 for the maximum position error.

As mentioned, in loosely coupled integration anything less than four satellites represents a complete GPS outage and no GPS aiding is provided to the INS. For this reason a loosely coupled integration will always behave like the zero-satellite case when it does not have four satellites available. Hence we can see that when it has three satellites, tightly coupled integration will provide an average positional error of 7 m as compared to 20 m of average positional error in the case of zero satellites mimicking loosely coupled integration. Similar behavior can be seen for two versus zero satellites. One important aspect to notice is that in the majority of cases having one satellite gives poorer performance than having none at all. This is contrary to theory. However, this phenomenon can be explained if having only one satellite with noisy and erratic measurements introduces a greater error than the standalone INS. This is particularly true for the higher grade IMUs. In this case a zero satellite (INS only) solution will provide better accuracy than GPS aiding with one erratic satellite. The average maximum position errors for all the five outages are summarized in Fig. 10.7, where we can clearly see the error trend that increases with a decrease in the number of satellites except for the just-explained case of one satellite versus zero satellites.

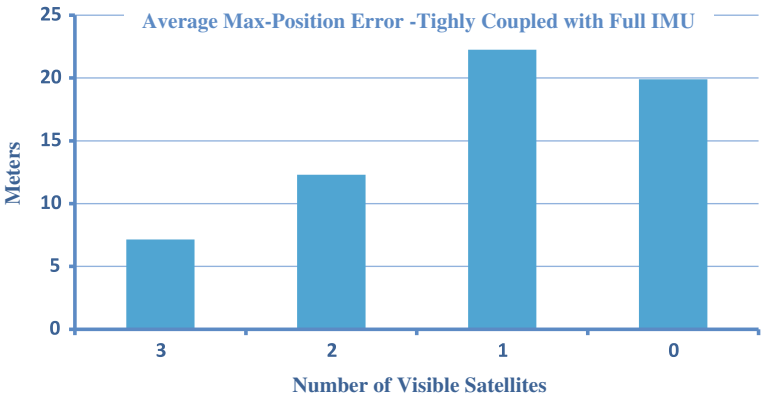


Fig. 10.7 The average maximum position error for all of the GPS outages

10.2.1 Analysis of Selected GPS Outages

The GPS outages were introduced during various phases of the road trip in order to analyze the performance of the algorithm for different scenarios. In this section some of the outages will be analyzed and their behavior compared with various satellite signal outages. The plots of the trajectories have been generated using an online software application called *GPSVisualizer* (2008).

The performance of the tightly coupled KF algorithm during the first outage of the trajectory is shown in Fig. 10.8. It shows the trajectories for partial GPS outages of three (green track), two (blue), one (aqua) and zero (magenta) visible satellites, in addition to the reference trajectory (red). It should be noted that due to a discrepancy in the plotting software the reference solution is slightly off the road. However, for our purposes the trajectory obtained from the NovAtel SPAN system (rather than the road) is considered as the reference.

This GPS outage starts before a sharp turn when the speed of the vehicle at the start of the outage is around 40 km/h, and after the turn it gradually slows to zero for a traffic light. Figure 10.9 shows the speed and azimuth plots of the trajectory for this outage.

We can see from Fig. 10.8 that when the three satellites were available the algorithm followed the reference very closely and the maximum position error was 4.72 m in contrast to 21.99 m for the zero satellite case. During this outage, two satellites also gave satisfactory accuracy with a position error of 12.20 m. The one satellite case gave higher error of 21.97 m, very similar to a complete outage (zero satellites).

Outage #3 was exactly the opposite scenario to outage #1, starting just before a turn while the vehicle was stationary waiting to turn left at a traffic light. After the turn, the vehicle continued to accelerate to the maximum limit of 50 km/h for that street. The track for this outage is depicted in Fig. 10.10.

Fig. 10.8 The performance of a tightly coupled KF algorithm during GPS outage #1

The dynamics during this outage are shown in Fig. 10.11. The response of the algorithm is in line with the theory, as the three-satellite case gives a small position error of 8.74 m with a gradual increase in the position error towards 25.81 m for the zero satellite case, which is approximately 0.43 m worse than the one satellite case. The maximum errors are higher than GPS outage #1 as a consequence of the vehicle undergoing acceleration after making the turn.

The fourth GPS outage was introduced when the vehicle was traveling on a straight road at 50 km/h, and after 45 s it came to a stop at a traffic light. The change in azimuth is minimal at about 4°. For the three-satellite case the position error is only 2.81 m. This is understandable, as the vehicle was traveling on a straight portion of the road where the gyroscope scale factor does not play any role in increasing the position error. In addition, the average velocity was also lower. The other reason for the relatively small error is that vehicle was stationary for the final 15 s or so, when the accelerometer scale factor made no contribution to the error. In this outage the zero satellite case gave a slightly better result than the one satellite case. Figure 10.12 shows the track and Fig. 10.13 shows the vehicle dynamics during outage #4.

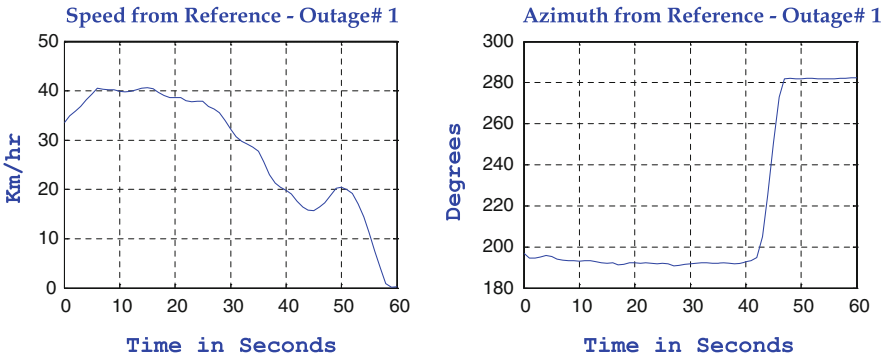


Fig. 10.9 Speed and azimuth dynamics for GPS outage #1 (plotted from the reference solution)

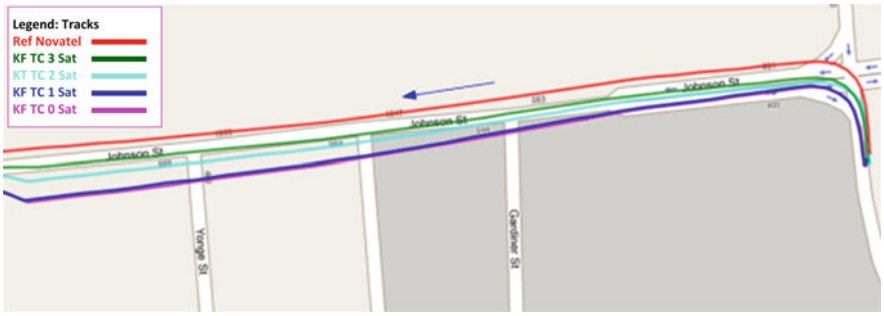


Fig. 10.10 The performance of a tightly coupled KF algorithm during GPS outage #3

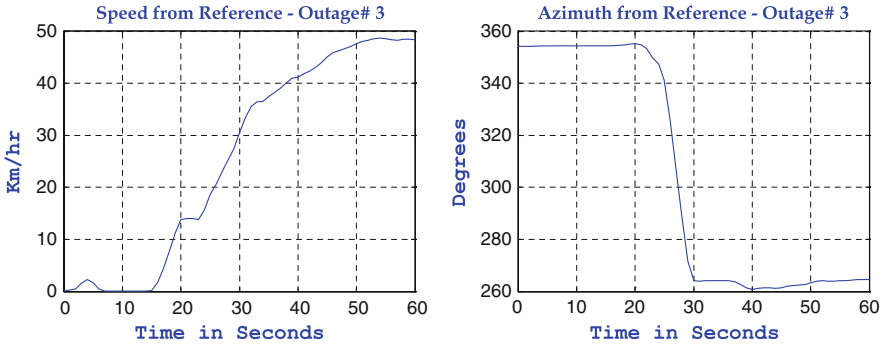


Fig. 10.11 Speed and azimuth dynamics for GPS outage #3 (plotted from the reference solution)

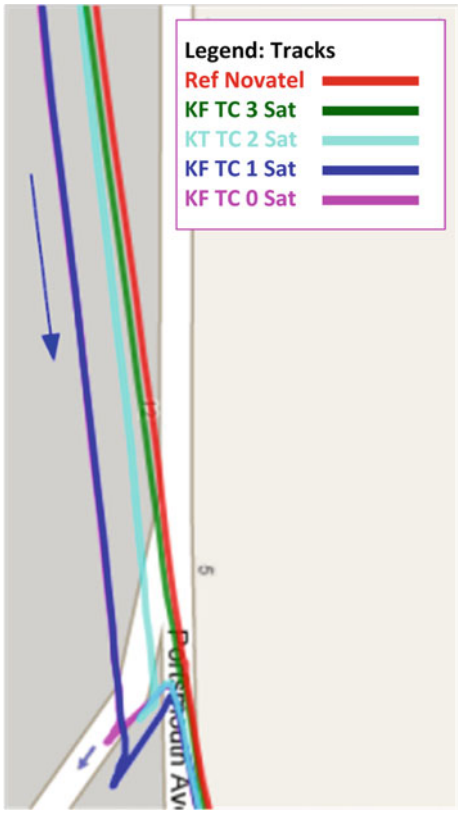


Fig. 10.12 The performance of a tightly coupled KF algorithm during GPS outage #4

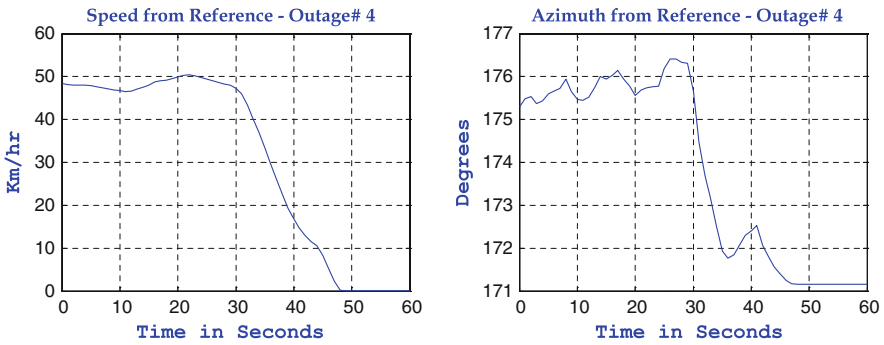


Fig. 10.13 Speed and azimuth dynamics for GPS outage #4 (plotted from the reference solution)

10.3 Performance of Tightly Coupled Algorithm for 3D RISS/GPS

Sensors from a MEMS grade Crossbow IMU300CC were used to assess the performance of the RISS algorithm. This IMU is inexpensive, but has the huge gyro bias of 2 deg/sec as compared to 1 deg/h for the tactical grade IMU that was used for the previous trajectory. This was done to show the superior performance of the RISS algorithm which can achieve an accuracy very close to that of the tactical grade IMU whilst using a MEMS grade IMU.

The trajectory for analyzing the tightly coupled RISS algorithm is displayed in Fig. 10.14, and included driving in downtown Kingston at a maximum speed of 40–50 km/h, suburban areas with an 80 km/h limit, and a portion of a highway with a 100 km/h limit. Five intentionally introduced GPS outages were designed to address several driving scenarios including high speeds, stops, sharp and mild turns, etc. The start and end points are indicated with a red triangle and red circle respectively. The direction of travel is shown with a blue arrow and the simulated partial GPS outages of fewer than four visible satellites are shown by blue circles. The outage duration was fixed at 60 s (as in the full IMU trial) and every outage was repeated four times in order to progressively decrease the number of visible satellites to zero.

The availability of satellites during the trajectory is shown in Fig. 10.15. It is evident that on average six to seven satellites were available but occasionally the number fell below four due to the blockages typically encountered in this kind of trajectory.

The data for the maximum as well as the RMS error during the partial and full GPS outages is given in Table 10.3. The trend of the maximum positional errors during the partial and full GPS signal blockages is shown in Fig. 10.16.

We can see that when three satellites are available, tightly coupled integration will provide an average positional error of 10.65 m as compared to 36 m for the zero-satellites case. This demonstrates that the tightly coupled integration performs much better by virtue of making use of the information that is available from the visible satellites, even when there are fewer than four satellites visible. Loosely coupled integration does not benefit from GPS if the number of satellites falls below this threshold.

Figure 10.17 shows the maximum positional errors averaged over a total of five GPS outages for the four cases of satellite visibility. The advantage of RISS integration can be appreciated when comparing the results of a MEMS grade IMU in Fig. 10.17 with those of the tactical grade IMU in Fig. 10.7. The average errors of the MEMS grade IMU are not significantly greater than the tactical grade IMU, despite the huge difference in their gyro bias (stated above). Hence it pays to use the RISS when the speed data is available, which is the case for most vehicles.

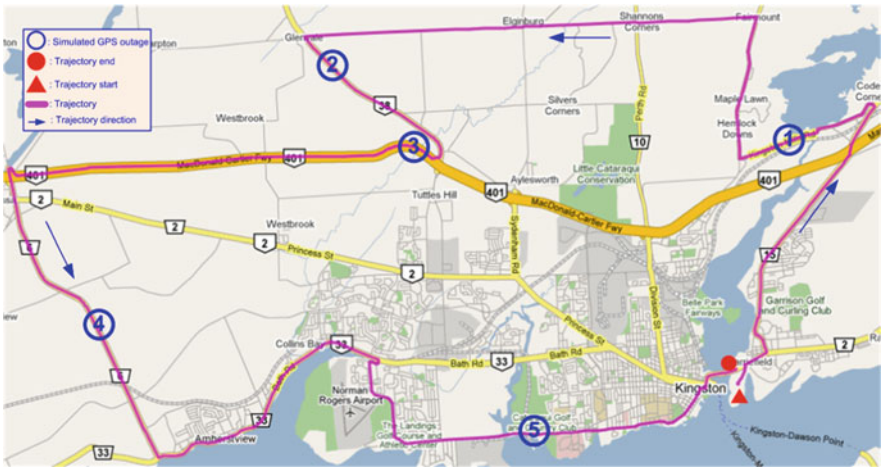


Fig. 10.14 The road test trajectory in a suburban part of Kingston

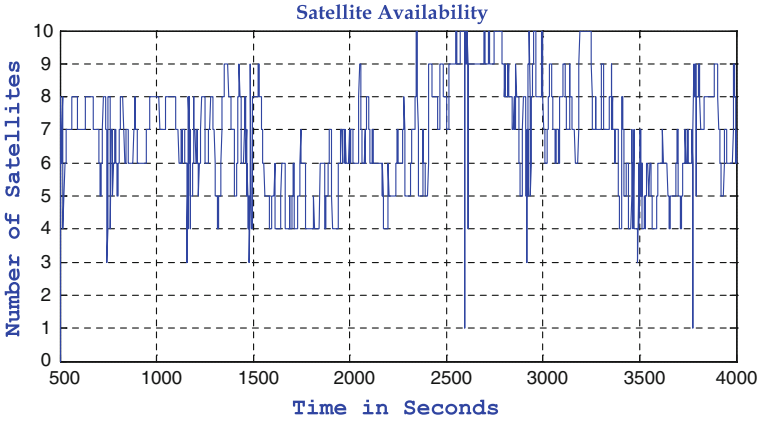


Fig. 10.15 Availability of satellites during the trajectory for the RISS test

Table 10.3 Maximum and RMS position errors for partial outages of a GPS that is tightly coupled with a RISS

Outage number	Maximum and RMS position error (in meters)							
	3 satellites		2 satellites		1 satellite		0 satellite	
	Max	RMS	Max	RMS	Max	RMS	Max	RMS
1	11.77	8.58	22.96	14.44	25.89	15.61	25.22	15.27
2	8.60	7.06	32.20	18.13	38.72	20.33	36.76	19.21
3	13.75	7.81	19.42	12.30	56.30	33.91	57.53	34.52
4	9.39	7.95	10.24	6.85	33.40	19.01	33.59	19.09
5	9.73	7.60	14.97	10.01	31.42	18.36	27.72	16.27
Average	*10.65*	*7.80*	*19.96*	*12.35*	*37.14*	*21.44*	*36.17*	*20.87*

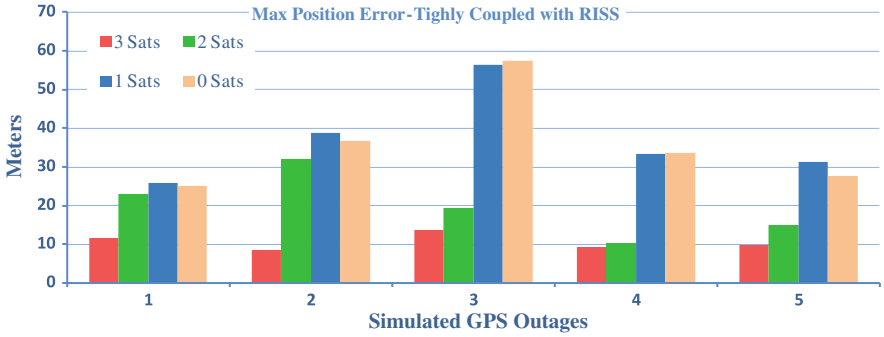


Fig. 10.16 A bar graph of maximum positional errors for partial outages of a GPS that is tightly coupled with a RISS

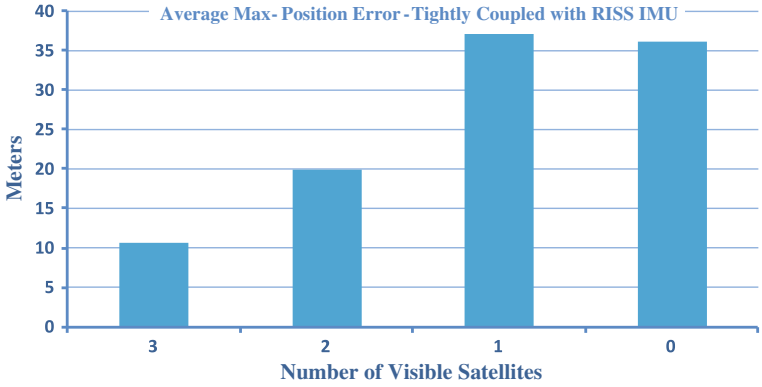


Fig. 10.17 The average maximum position error for all of the GPS outages

10.3.1 Analysis of Selected GPS Outages

Some of the outages will be discussed in order to examine the performance of the 3D RISS algorithm during different phases of the trajectory. GPS outage #2 starts after a sharp turn and continues on a fairly straight portion of the road, as shown in Fig. 10.14. The zoomed GPS outage is shown in Fig. 10.18, and the vehicle dynamics are shown in Fig. 10.19 in terms of speed and azimuth angle. As seen from Fig. 10.18, in the case of three and two satellites the trajectories remain very close to the reference and almost within the road boundaries. For the zero satellite case the trajectory is off the road with an error of about 37 m. This is the typical error for a loosely coupled integration during such an GPS outage, and shows the clear advantage of a tightly coupled 3D RISS algorithm. In the case of one visible satellite the solution does not differ much for the reasons explained earlier.

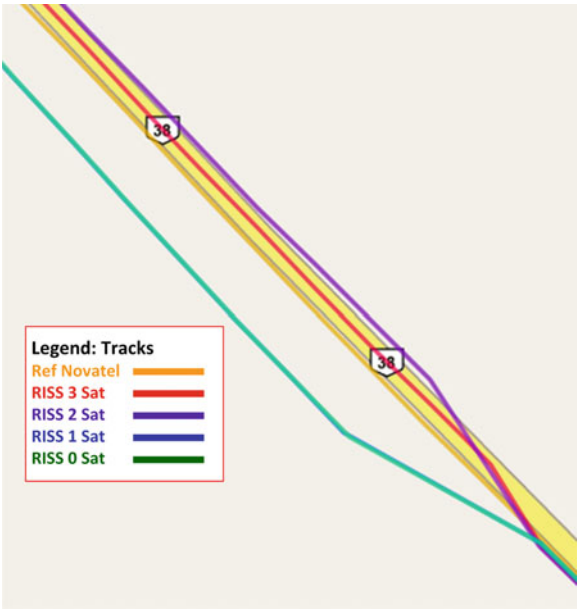


Fig. 10.18 Performance of a tightly coupled 3D RISS during GPS outage #2

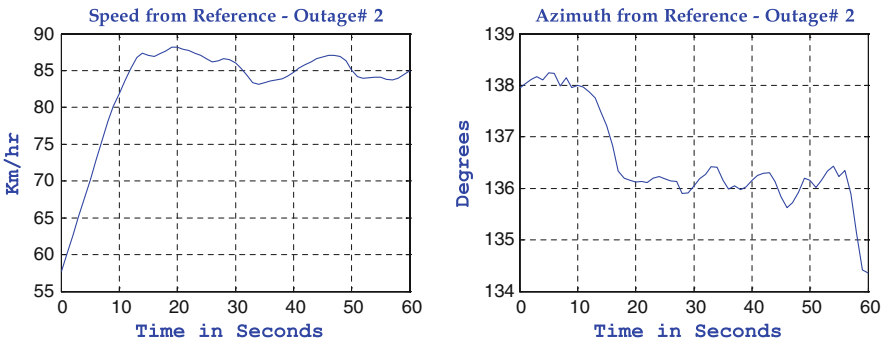


Fig. 10.19 Speed and azimuth dynamics for GPS outage #2 (plotted from the reference solution)

As shown in Fig. 10.20, GPS outage #3 was introduced at a turn where the vehicle entered the highway. The direction of the vehicle changes slowly from an azimuth of 305 to about 240°, and the speed increases from 65 to 105 km/h as depicted in Fig. 10.21.

A zoomed view of GPS outage #3 is shown in Fig. 10.22. It is evident that even during this high speed turn the 3D RISS algorithm performs well, and for the case

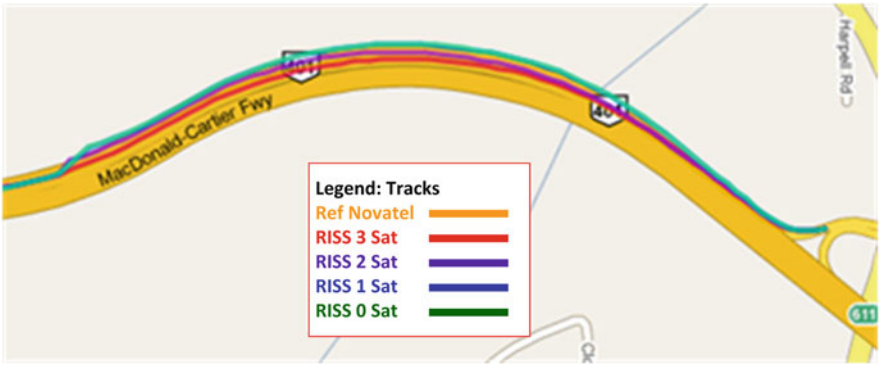


Fig. 10.20 An overview of GPS outage #3 at a turn on the highway

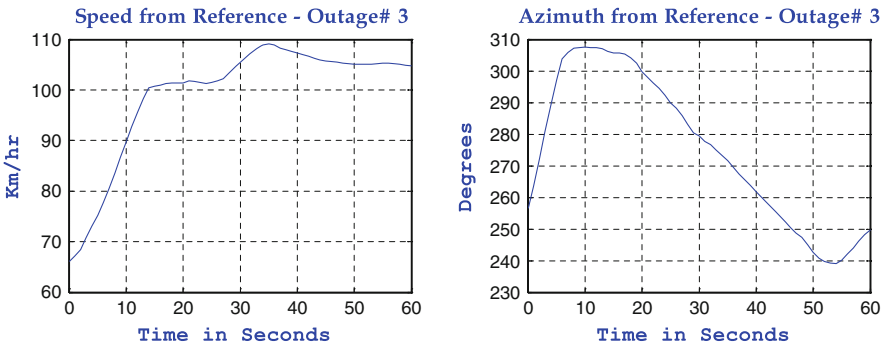


Fig. 10.21 Speed and azimuth dynamics for GPS outage #3 (plotted from the reference solution)

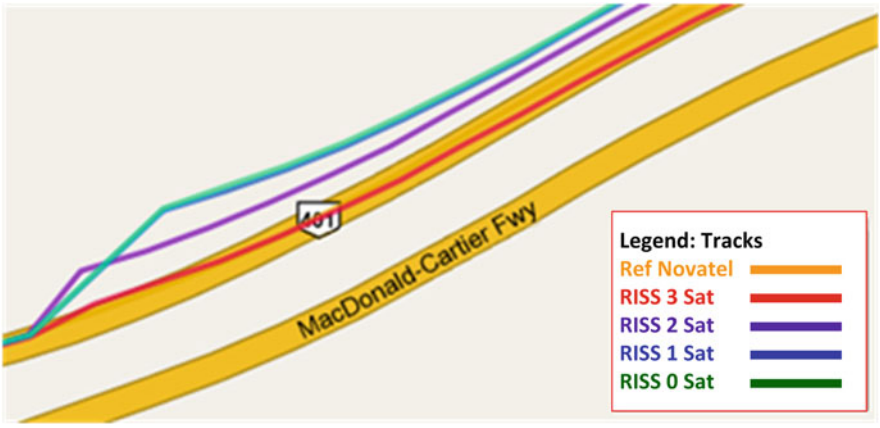


Fig. 10.22 A zoomed view of the performance of a tightly coupled 3D RISS during GPS outage #3

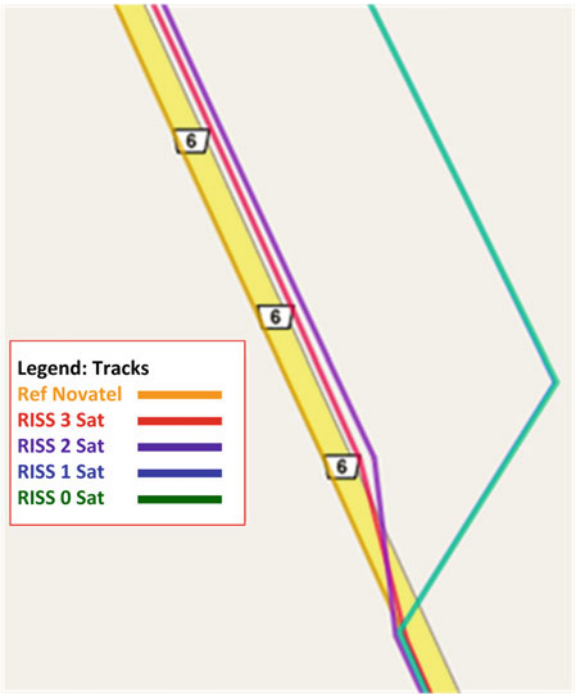


Fig. 10.23 Performance of a tightly coupled 3D RISS during GPS outage #4

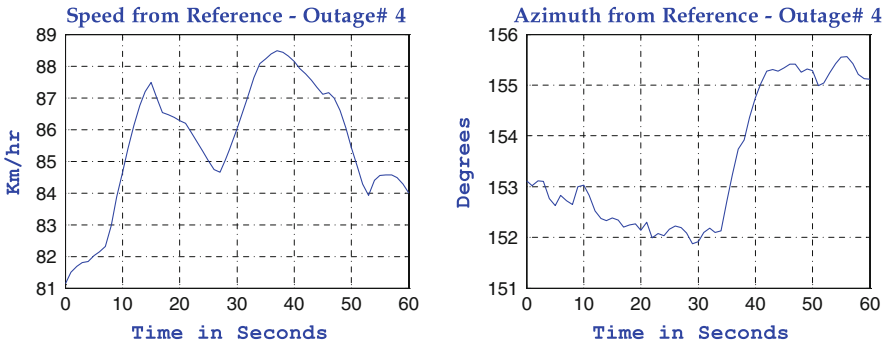


Fig. 10.24 Speed and azimuth dynamics for GPS outage #4 (plotted from the reference solution)

of three satellites the error remains below 14 m. In the zero satellite case the error sometimes exceeds 55 m.

GPS outage #4 was introduced at the moderate average speed of 83 km/h on a section of highway where the azimuth was essentially fixed. The zoomed view of Fig. 10.23 shows the performance and Fig. 10.24 shows the dynamics.

The case of outage #4 also clearly shows the accuracy of the tightly coupled 3D RISS algorithm, with the three and two visible satellite cases providing under 10 m of error and staying very close to the road boundaries. The zero and one satellite cases have positional errors of over 19 m.

This chapter has demonstrated the efficacy of the tightly coupled integration algorithm as compared to loosely coupled integration, which is similar to the case where the number of available satellites is reduced to zero. Also, the results from a 3D RISS/GPS integration using a MEMS grade IMU were contrasted with a fully integrated system using a tactical grade IMU and this algorithm was shown to be capable of improving the positional accuracy of low cost sensors to approximately that of a tactical grade IMU.

References

Crossbow (2007) IMU User's manual-models IMU300CC, IMU400CC,IMU400CD. Crossbow technology, inc. doi:7430-0003-03

Crossbow (2012) IMU300—6DOF inertial measurement unit: crossbow technology inc. http://bullseye.xbow.com:81/Products/productdetails.aspx?sid=208. Accessed 07 Mar 2012

Davis (2012) CarChip OBDII-based vehicle data logger and software: Davis instruments. Davis instruments corp. www.davisnet.com/product_documents/drive/spec_sheets/8211-21-25_carchip_specsB.pdf. Accessed 7 Mar 2012

NovAtel (2005) SPAN technology user manual OM-20000062. NovAtel inc. doi:OM-20000062 Rev7

Schneider A (2008) Draw a Google map from a GPS file. www.gpsvisualizer.com/map_input?form=google. Accessed 29 Sep 2011