



Royal Institute of Technology

GPS/IMU Integrated System for Land Vehicle Navigation based on MEMS

Yueming Zhao

Licentiate thesis in Geodesy

Royal Institute of Technology (KTH)
Division of Geodesy and Geoinformatics
10044 Stockholm Sweden

September 2011

TRITA-SoM 2011-16

ISSN 1653-6126

ISRN KTH/SoM/11-16/SE

ISBN 978-91-7501-126-4

Abstract

The Global Positioning System (GPS) and an Inertial Navigation System (INS) are two basic navigation systems. Due to their complementary characters in many aspects, a GPS/INS integrated navigation system has been a hot research topic in the recent decade. Both advantages and disadvantages of each individual system are analyzed.

The Micro Electrical Mechanical Sensors (MEMS) successfully solved the problems of price, size and weight with the traditional INS. Therefore they are commonly applied in GPS/INS integrated systems. The biggest problem of MEMS is the large sensor errors, which rapidly degrade the navigation performance in an exponential speed. By means of different methods, i.e. autoregressive model, Gauss-Markov process, Power Spectral Density and Allan Variance, we analyze the stochastic errors within the MEMS sensors. Real tests on a MEMS based inertial measurement unit for each method are carried out. The results show that different methods give similar estimates of stochastic error sources. These error coefficients can be used further in the Kalman filter for better navigation performance and in the Doppler frequency estimate for faster acquisition after the GPS signal outage.

Three levels of GPS/IMU integration structures, i.e. loose, tight and ultra tight GPS/IMU navigation, are introduced with a brief analysis of each character. The loose integration principles are given with detailed equations as well as the basic INS navigation principles.

The Extended Kalman Filter (EKF) is introduced as the basic data fusion algorithm, which is also the core of the whole navigation system to be presented. The kinematic constraints of land vehicle navigation, i.e. velocity constraint and height constraint, are presented. These physical constraints can be used as additional information to further reduce the navigation errors. The theoretical analysis of the Kalman filter with constraints are given to show the improvement on the navigation performance. As for the outliers in practical applications, the equivalent weight is introduced to adaptively reduce the influence on positioning accuracy.

A detailed implementation process of the GPS/IMU integration system is given. Based on the system model, we show the propagation of position standard errors with the tight integration structure under different scenarios. Even less than 4 observable satellites can contribute to the integrated system. Especially 2 satellites can maintain the orientation errors at a reasonable level due to the benefit of the tight integration. A real test with loose integration structure is carried out, and the EKF performance as well as the physical constraints are analyzed in detail. Also a test with random outliers at the resolution level is carried out to show the effectiveness of the equivalent weight. Finally some suggestions on future research are proposed.

Keywords: GPS, IMU, MEMS, integration, Kalman filter, physical constraint, outlier

Acknowledgments

First of all, I would like to express my sincere gratitude to my supervisors, Professor Lars E. Sjöberg and Docent Milan Hormuz, for their inspiring guidance, valuable suggestions during my study time at KTH. Thanks to the strict requirements and kind help from Professor Sjöberg, I gradually learned how to write a formal paper and thesis in many aspects. Docent Hormuz is always so nice and kind when I encounter many detailed problems. Without the supports from both of you, this thesis could not be finished.

I wish to thank the members of our division for their kind help, especially Dr. H. Fan. For many times I felt puzzled with different problems about both my study and my life. Dr. H. Fan is always willing to help both as one of the faculty and as an elder friend from the same country. Thanks also go to Mr. M. Shirazian for the discussions and helps on GPS during this period, as well as Mr. E. Asenjo, Docent M. Eshagh, Dr. M. Bagherbandi and Ms. U. Danila for the useful advices during the discussions and the everyday life.

Finally, I would like to express special thanks to my family members, especially my dear sister, for their limitless love and support during my whole life no matter where I am, no matter what I am doing.

I also want to thank my friends at KTH and other universities in Sweden as well as my friend in China for our friendship.

Yueming Zhao
August 2011, Stockholm

Contents

Abstract	i
Acknowledgement	ii
List of abbreviations	v

1. Introduction

1.1 GNSS positioning and its features	1
1.2 INS and its errors	2
1.3 Integrated navigation system and physical constraints	3
1.4 My thesis work	4

2. MEMS-based IMU error modeling

2.1 MEMS and error sources	6
2.2 Methodology of MEMS error modeling	8
2.2.1 Autoregressive model.....	8
2.2.2 Gauss-Markov process	9
2.2.3 Power Spectral Density	11
2.2.4 Allan Variance	12
2.3 Testing results	17
2.4 Discussion and conclusion	22

3. Integrated navigation structures

3.1 Three levels of integration	23
3.2 principle of loose integration for LVN.....	25
3.2.1 GPS observations.....	26
3.2.2 INS and navigation principle.....	28
3.2.2.1 Coordinate frames	28
3.2.2.2 Rotation matrix and Euler angles	30
3.2.2.3 IMU and navigation equations	32
3.2.3 INS sensors calibration.....	37
3.2.4 INS Initialization and Alignment	38
3.2.4.1 Coarse alignment and fine alignment	39

4. Robust Kalman filter with constraints

4.1 Extended Kalman filter	41
4.2 Land vehicle model	42
4.2.1 Velocity constraint.....	43
4.2.2 Height constraint	45
4.2.3 Lever arm effect	47
4.3 Constrained Kalman filter	48
4.4 Outliers and robust filter	51

5. Tests and analysis

5.1 Implement of the integrated system with EKF.....	53
5.1.1 State vector and system dynamic equations.....	53
5.1.2 Differenced GPS observation equations	56
5.1.3 System observation equations.....	58
5.1.4 Loose integration implementation.....	60
5.2 Variances of tight integration on simulated observations.....	63
5.3 Positioning of loose integration on real observations	67
5.4 Analysis on Kalman filter performance and the physical constraints	69
5.5 Tests on robust Kalman filter	72

6. Summary, conclusion and further research

6.1 Summary and conclusion	75
6.2 Recommendations for further research	77

References.....	78
------------------------	-----------

Appendix

Appendix A	84
------------------	----

List of abbreviations

ACF	Autocorrelation function
AD	Allan Deviation
AR	Autoregressive
AV	Allan Variance
b-frame	body frame
CG	Center of Gravity
DPR	Delta Pseudorange
e-frame	Earth-Centered-Earth-fixed Cartesian frame
EKF	Extended Kalman Filter
GM process	Gauss-Markov process
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
i-frame	inertial frame
INS	Inertial Navigation System
IMU	Inertial Navigation Unit
LVN	Land Vehicle navigation
M-estimation	maximum likelihood estimation
MCU	Micro Controller Unit
MEMS	Micro Electrical Mechanical Sensor
NED	North-East-Down
n-frame	navigation frame
RMS	Root Mean Square
RTK	Real Time Kinematic
PDOP	Position Dilution of Precision
PPP	Precise Point Positioning
PR	Pseudorange
PSD	Power Spectral Density
SBC	Schwarz's Bayesian Criterion
SINS	Strapdown Inertial Navigation System
TDF	Two-degree of Freedom
ZVUP	Zero Velocity Update

1. Introduction

Navigation deals with trajectory determination and guidance. Trajectory determination relates to the derivation of the state vector of an object at any given time. The state vector includes position, velocity, and attitude. On the other hand guidance forces the moving object onto a predetermined route to reach a given destination.

GNSS and INS are two commonly used systems for vehicle navigations. In this chapter, the advantages and disadvantages of both systems are introduced. There different common integration structures are explained. Within the integration system the core filtering algorithm is introduced as well as the stochastic error modeling of Inertial Measurement Unit (IMU).

1.1 GNSS positioning and its features

There are two main working global navigation satellite system (GNSS) system now, i.e. GPS and GLONASS, plus another two under construction, i.e. Galileo and Compass. GPS is the most commonly used system, therefore we are going to use GPS in this thesis, but generally any other GNSS could also be used.

The Global Positioning System (GPS) is a space-based GNSS that provides reliable positioning and time information in all weather and at all times and anywhere on or near the Earth. GPS was established in 1973 by the U.S. Department of Defense. It is composed of three segments: the space segment, the control segment and the user segment. The space segment consists of 24 satellites in six planes.

The most important advantage of GPS positioning is the limited positioning errors. Once signals from more than 4 satellites are received with suitable PDOP (Position Dilution of Precision), the quality of positioning results can be guaranteed, i.e. the errors are limited.

Other important advantages are the light weight and cheap price. As a result, since the beginning GPS positioning is becoming more and more popular. Nowadays GPS is used almost everywhere, in cell phones, cars, laptops and so on. The use of GPS (or GNSS) is only limited by our imaginations. Current GPS receiver chips are reaching a unit price of \$5, which is predicted to drop to about \$1 in the future.

However, GPS positioning has many inherent shortcomings. The most important problem is the signal outage. It is quite common that satellite signals are blocked in urban and mountainous areas, as well as indoor and underground situations. Another

situation is that for high dynamic movements the signal acquisition and tracking is difficult. The availability of satellite signals must be considered in these situations.

Moreover, the security of GPS signals is also a big problem. As we all know that the satellite signals are vulnerable to interference and spoofing. The security of wireless signal including satellite signal is always a vital problem for electronics engineers, especially when it is related with communication and military applications. Theoretically a 1-W jammer located 100 km from the GPS antenna could prevent the acquisition of the C/A code (Schmidt, 2010).

Another problem related with high dynamics applications is the low update frequency. The update frequency of GPS receivers is normally 1-10 Hz, which is 1-2 orders of magnitude lower than that of Inertial Navigation System (INS). For sport cars and air planes this update frequency is too low to be acceptable. With the fast development of CPU and the ongoing research, this problem is hoped to be solved gradually in the future.

In real applications, if there is only one GPS receiver or one antenna there is no possibility to get the attitude information from GPS navigation. This is a great loss of sufficient information in many cases like airplane navigation.

1.2 INS and its errors

INS is based on the Newton's second law and has many advantages as a means of navigation. The most important one is that it does not rely on external information and does not radiate any energy when operating. Therefore it is a kind of autonomous or self-contained navigation system, which is quite suitable for military applications.

INS is mainly divided into two types: platform INS and strapdown INS (SINS). Now SINS is becoming more and more dominated with the powerful computation ability of embedded MCU (Micro Controller Unit). The presence of so-called MEMS (Micro-Electro-Mechanical Systems) significantly reduces both price and weight. Although the accuracy of current MEMS is still at a low level, it becomes a hot research topic soon after adopted in the navigation field. The accuracy performance of MEMS improves fast all the time and the price is very low.

INS is very accurate over short periods with high update frequency. Usually the update frequency is at least 100 Hz. But meanwhile the cost is also high and the high accurate INS is heavy and with big volume. Furthermore, the fine initial alignment is tricky and time-consuming.

The biggest problem of INS is the sensor errors, the mean value of which is not zero but keeps on increasing with time. We cannot completely eliminate this physical phenomenon, but the errors can be modeled and reduced thereafter. To get the precise

result of the noise parameters, usually a long time testing and recording the error data are needed beforehand. This part is crucial to reduce the effects of different kind of error sources.

There are different methods to model and estimate the inertial sensor stochastic errors, such as Autoregressive (AR) Model (Babu et al., 2004, 2008; Nassar S., 2005), Gauss-Markov (GM) model (Mohammed D. and Spiros P., 2009) and Allan Variance (AV) (Kim et al., 2004; EI-Sheimy, 2008). Hou (2004) has verified these methods on modeling inertial sensors errors. In Chapter 2 these methods are applied to estimate the coefficients of IMU stochastic errors.

1.3 Integrated navigation system and physical constraints

For many aspects, GPS and INS are two complementary navigation systems, including the advantages each has and the positioning errors of each system. INS has almost no high frequency errors but the errors grow up with time, while GPS, on the other hand, has high frequency noise but with good long-term accuracy (i.e., small bias errors).

There are several different architectures as to the GPS/INS integration, namely the loose, tight and ultra tight architecture. Initially, two broad classes of integration structure: loose and tight coupling were developed (Grewal et al., 2007). However, in the recent years a third class has been proposed, i.e. deep integration or ultra tight integration (Sun, 2010).

The salient difference among these couplings is the different levels of combining INS and GPS observables. The deeper GPS and INS are integrated, the more information we can get, but meanwhile the more dependently they rely on each other. The detailed information about GPS/INS integration architectures will be given in Chapter 3.

Traditionally GPS and INS are coupled through a Kalman filter for the processing of raw observables to obtain position, velocity and time. It is the core algorithm of the navigation system. As a filter we hope it to be accurate and robust. The robustness of the system is the ability to get rid of outliers. To avoid the influence of outliers, the equivalent weights are applied to the Kalman filter along with kinematic constraints to give smooth weights for measurement errors (Yang et al., 2010). In this way the outliers can be rejected and the contributions of measurements with larger errors are suppressed to avoid causing significant rise of the standard errors.

To further reduce the navigation errors, the physical model of the Land Vehicle Navigation (LVN) can provide additional navigation information besides the GPS and IMU, which is quite helpful in specific situations. For example it is acceptable to adopt the height constraint, i.e. assuming that the height is constant, when there are only 3 observable GPS satellites in relatively flat areas. We study the kinematic model

of LVN under several assumptions in Chapter 4. Aiming at specific applications of LVN, the velocity and height constraints as well as the lever arm effect correction are derived and analyzed in details. All these constraints have clear physical meanings and contributions to improve the accuracy of the whole system in different aspects (see Section 4.2).

With these physical models the constrained Kalman filter can be implemented, and the additional physical information of LVN is helpful to improve the accuracy in different situations as shown in Section 4.3, since Kalman filtering with state constraints is verified to significantly improve the estimation accuracy (Simon and Chia, 2002).

1.4 My thesis work

This thesis is aiming at designing an accurate and robust GPS/IMU navigation system for LVN. There are also many other types of navigation systems, e.g. image-based navigation, Terrestrial radio navigation, etc. In this thesis, only the MEMS based IMU and GPS receiver are used in the integrated system for the LVN in urban areas.

First, the MEMS-based IMU (Inertial Measurement Unit) is modeled and tested with different methods. The detailed analysis and experiments on IMU error sources and estimating the coefficients are shown in Chapter 2.

Among different integrated structures, three levels of GPS/IMU integrated navigation system are analyzed. As for the data fusion algorithm in the integrated navigation system, the Extended Kalman Filter (EKF) will be adopted. For the LVN applications, we will study the kinematic models of LVN and apply them in the constrained Kalman filter to further reduce the positioning errors. To eliminate the influence of outliers, the maximum likelihood estimation is also applied to the Kalman filter.

To test the effectiveness of the methodology above, both computational simulations and an actual LVN experiment are carried out. We will first carry out the variance analysis on the tight integrated system based on double-difference GPS observations and IMU observations to show the influence of different numbers of observable satellites. Then with an MEMS based IMU and differenced GPS solutions we implement a loose integrated LVN system for both real-time navigation and post-processing analysis. The EKF as the data fusion algorithm of the integration system is introduced and implemented. Velocity and height constraints are analyzed based on the EKF integration implementation. Finally the equivalent weight is applied to the navigation filter to suppress the influence of outliers at the resolution level with some manually added outliers.

Although all these methods have been discussed in parts by different papers or literatures as mentioned in the related references, we believe that this is the first time

to integrate all of them together in one specific LVN application. Here we will study the integrated results of the various parts of the LVN. After analysis of the methodology and test results, a summary is composed and some proposals for further research are proposed.

2. MEMS-based IMU error modeling

Since the gyroscope behaviors of precession and nutation are known, gyroscopes are used to construct gyrocompasses which can replace magnetic compasses on different vehicles to assist in stability or be used as part of an inertial navigation system. The IMU are composed of two parts, i.e. three single-degree of freedom gyros and three mutually orthogonal accelerometers based on Newton's second law.

2.1 MEMS and error sources

For a long time the IMU used in navigation systems were heavy and expensive. Then gradually the Micro Electrical Mechanical Sensor (MEMS) with great advantages of price and volume is widely used in GPS/IMU integrated systems, as shown in Fig. 2.1. Although the accuracy of MEMS is improved rapidly during latest years as shown in Fig. 2.2, the MEMS stand-alone navigation during GPS signal blockage could still only maintain the reasonable error within a very short time due to quickly growing sensor errors as shown in Fig. 2.3.



Fig. 2.1 The Crista IMU produced by Cloud Cap Technology (Brown and Lu, 2004)

The systemic errors of MEMS are estimated before used, which means that we roughly estimated the bias and scale errors. General stochastic error sources existing in inertial sensors include (IEEE STD 647, 2006): Quantization Noise, Random Walk, Bias Instability, Rate Random Walk and Rate Ramp.

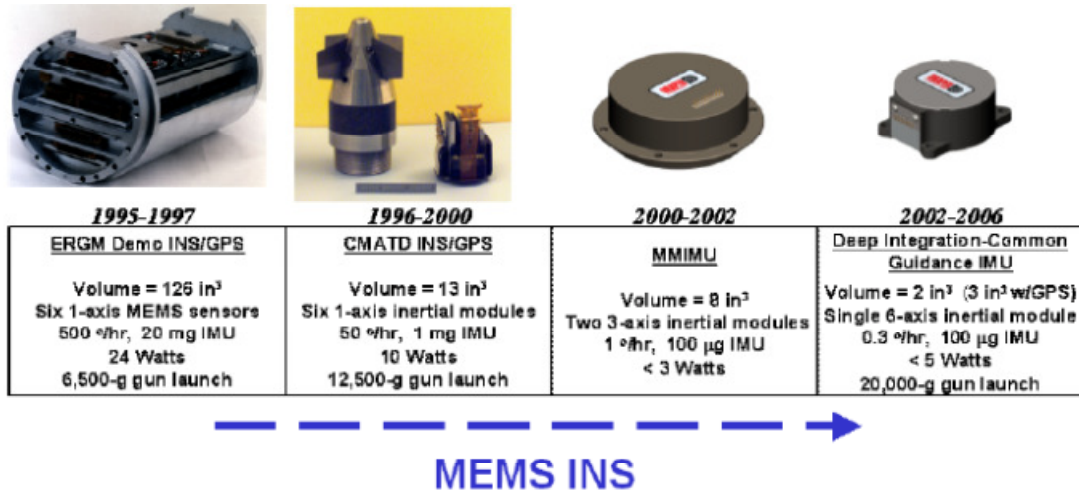


Fig. 2.2 The development of MEMS INS (Schmidt, 2010)

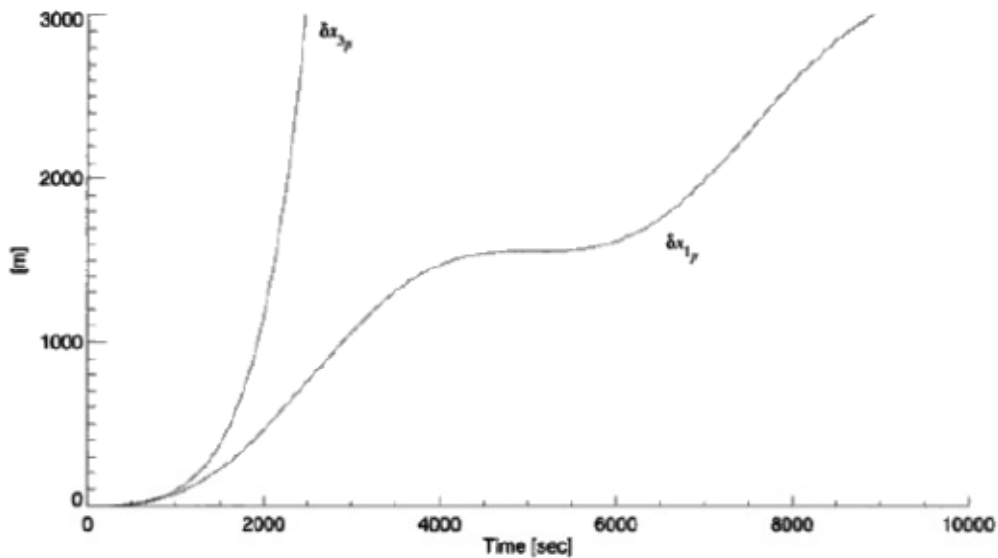


Fig. 2.3 Position errors due to unknown constant gyro drifts (Jekli, 2000, p. 147).

(δx_{1_p} and δx_{3_p} are the positioning errors along north and up directions, respectively)

Here we apply different methods to analyze the stochastic sensor errors, i.e. autoregressive (AR) modeling, Gauss-Markov (GM) process, Power Spectral Density (PSD) and Allan Variance (AV). Then the tests on a MEMS based IMU are carried out with these methods. The results show that different methods give similar stochastic error sources and values. These values can be use further in the Kalman filter for better navigation accuracy and in the Doppler frequency estimate for faster acquisition after GPS signal outage.

2.2 Methodology of MEMS error modeling

2.2.1 Autoregressive model (AR model)

An m-variate p-order autoregressive (AR(p)) model for a stationary time series of state vectors \mathbf{v}_v , observed at equally spaced instants v , is defined by

$$\mathbf{v}_v = \mathbf{w} + \sum_{l=1}^p \mathbf{A}_l \mathbf{v}_{v-l} + \boldsymbol{\varepsilon}_v \quad (2.1)$$

or

$$\mathbf{v}_v = \mathbf{B} \mathbf{u}_v + \boldsymbol{\varepsilon}_v \quad (2.2)$$

where $\boldsymbol{\varepsilon}_v$ are uncorrelated random vectors with zero mean and covariance matrix \mathbf{C} , \mathbf{A}_l are the coefficient matrices of the AR model, \mathbf{w} is a vector of intercept terms to allow for a nonzero mean of the time series and $\mathbf{B} = (\mathbf{w} \ \mathbf{A}_1 \ \dots \ \mathbf{A}_p)$, $\mathbf{u}_v^T = (1 \ \mathbf{v}_{v-1} \ \dots \ \mathbf{v}_{v-p})$. From Eq. (2.1) we can see that the current sample \mathbf{v}_v can be estimated by the previous p samples.

The first-order single-variate AR model in discrete time can be simply given as

$$x_k = \phi_1 x_{k-1} + w_k \quad (2.3)$$

where x is the random variable, subscript k and $k-1$ is the discrete time index and w_k is zero-mean Gaussian white noise with variance $\sigma_{w_k}^2$.

To use the AR model we should first determine the order and then the value of each coefficient. There are several methods to determine the order of the AR model. Here we use Schwarz's Bayesian Criterion (SBC) is used (Schwarz, 1978), which states that the AR order p should minimize the criterion

$$SBC(p) = \frac{l_p}{m} - (1 - \frac{n_p}{N}) \log N = \min_p. \quad (2.4)$$

where

$$l_p = \log(\det \Delta_p) \quad (2.5)$$

here Δ_p is the residual cross-product matrix and N is the number of samples, m is the dimension of the state vector, and $n_p = mp + 1$ is the dimension of the predictor \mathbf{v}_v .

Then the AR model parameters can be estimated using least-squares fitting, Yule-Walker equations (Eshel 2010) and Burg's method (Bos et al. 2002). A stepwise least-squares algorithm is used to determine the AR coefficients after the order is fixed (Schneider et al., 2001).

Here the method of least-square fitting is used. By means of the moment matrices

$$\mathbf{U} = \sum_{v=1}^N \mathbf{u}_v \mathbf{u}_v^T, \mathbf{V} = \sum_{v=1}^N \mathbf{v}_v \mathbf{v}_v^T, \mathbf{W} = \sum_{v=1}^N \mathbf{v}_v \mathbf{u}_v^T \quad (2.6)$$

the least-squares estimates of the parameter matrix and the residual covariance matrix can be written as (Neumaier and Schneider, 2001)

$$\hat{\mathbf{B}} = \mathbf{WU}^{-1} \quad (2.7)$$

$$\hat{\mathbf{C}} = \frac{1}{N - n_p} (\mathbf{V} - \mathbf{WU}^{-1} \mathbf{W}^T) \quad (2.8)$$

Where $n_p = mp + 1$ is the dimension of predictor \mathbf{v}_v . The variance of estimation noise can be given as

$$\sigma_{w_k}^2 = \frac{1}{n} \sum_{i=1}^n (x_i^d - \hat{x}_i)^2 \quad (2.9)$$

where n is the size of the sample of the stationary process, x_i^d is the known value of the process (desired output), and \hat{x}_i is the corresponding estimated output.

2.2.2 Gauss-Markov process (GM process)

A first-order GM (GM1) process is the most frequent model used in Kalman filtering of an integrated system due to its simplicity. The GM process is defined by the

exponential Auto Correlation Function (ACF):

$$R_x(\tau) = \sigma^2 e^{-\beta|\tau|} \quad (2.10)$$

where σ^2 is the noise variance, β^{-1} is the correlation time and τ is the time interval. The ACF of Eq. (2.9) is shown in Fig. 2.4. From this figure we can see that there is a peak at zero, and there are two symmetrical descendent slopes at both sides, the gradient of which is getting steep if the value of β goes up, and the gradient at zero is discontinuous.

The first-order GM process in discrete time and the variance are given as

$$x_k = e^{-\beta\Delta t} x_{k-1} + w_k \quad (2.11)$$

$$\sigma_{x_k}^2 = \sigma_{w_k}^2 / (1 - e^{-2\beta\Delta t_k}) \quad (2.12)$$

where Δt is the discrete time sampling interval.

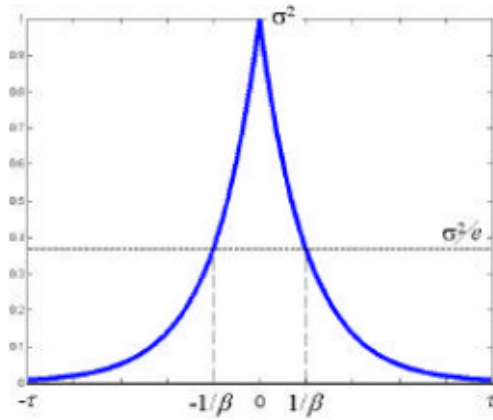


Fig. 2.4 ACF of GM process

and the variance is

$$\sigma_{x_k}^2 = \frac{\sigma_{w_k}^2}{1 - e^{-2\beta\Delta t_k}} \quad (2.13)$$

The GM1 process has been widely used in INS due to its bounded uncertainty characteristic which is suitable for modeling slowly varying sensor errors, i.e. bias and scale factors (El-Diasty and Pagiatakis, 2008).

With fixed sampling rate, the GM1 process is equal to a first order AR model. From equations (2.3) and (2.12), we obtain

$$e^{-\beta\Delta t} = \phi_1 \quad (2.14)$$

Therefore, the relation between these two models can be shown as

$$\beta = -\ln \phi_1 / \Delta t \quad (2.15)$$

However, the principle difference between the AR model and the GM process should be declared. The AR model does not consider the sampling interval, which may be considered as a sub-optimal estimation. The GM-only model needs a long data set, for example 200 times of the expected correlation time for 10% uncertainty (Nassar, 2005).

2.2.3 Power Spectral Density

The PSD is a commonly used and powerful tool for analyzing a signal or time series. In statistical signal processing, the PSD describes the distribution of energy in frequency domain. For a finite-energy signal $f(t)$, the definition of PSD is

$$S(\omega) = \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \right|^2 = \frac{F(\omega)F^*(\omega)}{2\pi} \quad (2.16)$$

where ω is the frequency, $F(\omega)$ and $F^*(\omega)$ are the Fourier transform of $f(t)$ and its complex conjugate respectively. A key point is that the two-sided PSD $S(\omega)$ and autocorrelation function $K(\tau)$ are Fourier transform pairs, if the signal can be treated as a wide-sense stationary random process (IEEE Std. 952, 1997):

$$S(\omega) = \int_{-\infty}^{+\infty} e^{-j\omega\tau} K(\tau) d\tau \quad (2.17)$$

$$K(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{j\omega\tau} S(\omega) d\omega \quad (2.18)$$

This relation also provides an approach to compute the PSD.

The typical characteristic slopes for typical sensor errors of the PSD are shown in Fig. 2.3, where the actual units and frequency range are hypothetical. With real data, gradual

transitions would exist between the different PSD slopes (IEEE Std1293-1998), rather than the sharp transitions in Fig. 2.5, and the slopes might be different from -2 , -1 , 0 , and $+2$ values.

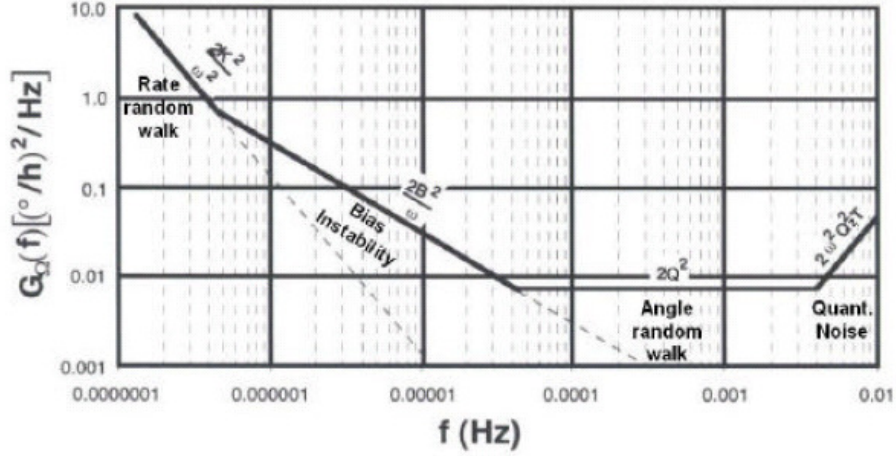


Fig. 2.5 Hypothetical Gyro in Single-sided PSD Form (IEEE Std952-1997)

2.2.4 Allan Variance

The Allan Variance (AV) is a method of representing root mean square (RMS) random drift error as a function of averaging time (Allan 1966). As a time domain analysis technique, it is an accepted IEEE standard for gyro specifications (IEEE STD 647, 2006). If there are N samples of data points with sampling interval Δt , then a group of n data points with $n < (N-1)/2$ can be created. Each group is called a cluster $\tau = n\Delta t$.

Assume that the instantaneous sample of the sensor is the angular velocity $\Omega(t)$, and its integration is the angle:

$$\theta(t) = \int_0^t \Omega(t') dt' \quad (2.19)$$

The average angular velocity in the k -th cluster, i.e. between time interval t_k and $t_k + \tau$, is

$$\bar{\Omega}_k(\tau) = \frac{1}{\tau} \int_{t_k}^{t_k+\tau} \Omega(t) dt \quad (2.20)$$

and the AV is defined as

$$\sigma^2 = \frac{1}{2} \left\langle (\bar{\Omega}_{k+n} - \bar{\Omega}_k)^2 \right\rangle = \frac{1}{2\tau^2} \left\langle [(\theta_{k+2n} - \theta_{k+n}) - (\theta_{k+n} - \theta_k)]^2 \right\rangle = \frac{1}{2\tau^2} \left\langle (\theta_{k+2n} - 2\theta_{k+n} + \theta_k)^2 \right\rangle \quad (2.21)$$

where the symbol $\langle \rangle$ is the infinite time average. In practice, the AV can be estimated with a finite number of samples by the so-called overlapped AV.

$$\sigma^2 \approx \frac{1}{2(N-2n)} \sum_{k=1}^{N-2n} (\bar{\Omega}_{k+1}(T) - \bar{\Omega}_k(T))^2 = \frac{1}{2\tau^2(N-2n)} \sum_{k=1}^{N-2n} (\theta_{k+2n} - 2\theta_{k+n} + \theta_k)^2 \quad (2.22)$$

The principle of the overlapped AV is shown by Fig. 2. 6, which has been accepted as the preferred AV estimator in many standards.

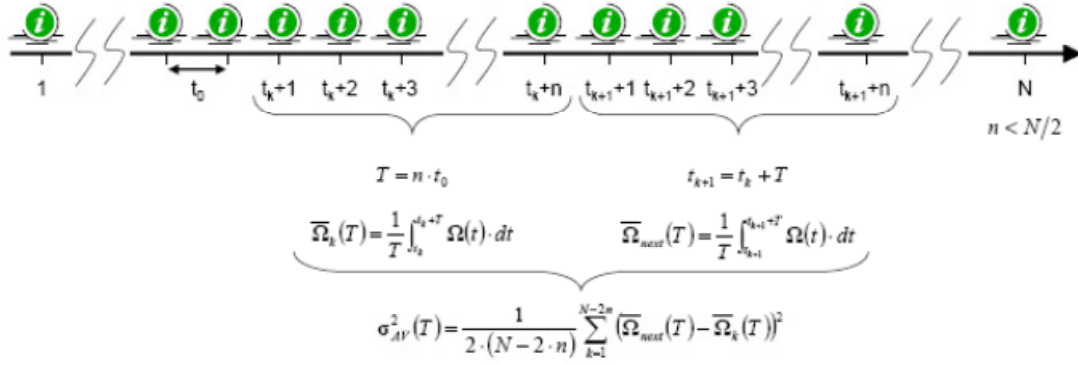


Fig. 2.6 Schematic algorithm of the overlapped AV (S Guerrier, 2008)

There is a unique relation between the AV and the PSD of a stationary process (IEEE Std.952-1997)

$$\sigma^2(\tau) = 4 \int_0^{+\infty} S(\omega) \frac{\sin^4(\pi\omega\tau)}{(\pi\omega\tau)^2} d\omega \quad (2.23)$$

Table 2.1 clearly shows the relationships between the PSD and the AV with Eq.2.23. The detailed derivations of these noises are given in (Tehrani, 1983). Table 2.2 shows the relation between curve slope and coefficient values on the log-log figure of the AV. The detailed derivation can be found in (Hou, 2004). A typical AV plot is shown in Fig. 2.7.

Table 2.1 Summary of PSD and AV (IEEE Std.647-2006)

Error Type	PSD (S_{Ω})	AV ($\sigma^2(\tau)$)	Comments
Quantisation Noise	$(2\pi f)^2 Q^2 T_s$	$\frac{3Q^2}{\tau^2}$	Q: Quantisation noise coefficient
Random Walk	N^2	$\frac{N^2}{\tau}$	N: Random walk coefficient
Bias Instability	$(\frac{B^2}{2\pi}) \frac{1}{f}$ $f \leq f_0$	$\frac{2B^2}{\pi} \left[\ln 2 - \frac{\sin^3 x}{2x^2} (\sin x + 4x \cos x) \right]$ $+C_i(2x) - C_i(4x)$	B: Bias instability coefficient. x: $\pi f_0 \tau$ C_i : Cosine integration function
Rate random walk	$(\frac{K}{2\pi})^2 \frac{1}{f^2}$	$\frac{K^2 \tau}{3}$	K: Rate random walk coefficient
Rate ramp	$\frac{R^2}{(2\pi f)^3}$	$\frac{R^2 \tau^2}{2}$	R: Rate ramp coefficient
Correlated Noise	$\frac{(q_c T_c)^2}{1 + (2\pi f T_c)^2}$	$\frac{(q_c T_c)^2}{\tau}, \tau \gg T_c$ $\frac{q_c^2}{3} \tau, \tau \ll T_c$	q_c : Noise amplitude T_c : correlation time
Sinoidal noise	$\frac{\Omega_0^2}{2} \left[\delta(f - f_0) + \delta(f + f_0) \right]$	$\Omega_0^2 \left(\frac{\sin^2 \pi f_0 \tau}{\pi f_0 \tau} \right)^2$	Ω_0 : point amplitude f_0 : point frequency $\delta(x)$: Delta function

The AV of the total stochastic process can be assumed as the sum of all different error terms, since they are independent on different time regions. The total AV of the system can then be expressed as:

$$\sigma_{tot}^2 = \sigma_{quant}^2 + \sigma_{RW}^2 + \sigma_{Bias}^2 + \sigma_{BRW}^2 + \dots \quad (2.24)$$

Table 2.2 Summary of characteristic curve slope and coefficient value

Error Type	$\log(\sigma)$ vs. $\log(\tau)$	Curve slope	Coefficient value
Quantisation Noise	$\log(\sigma) = -\log(\tau) + \log(\sqrt{3}Q)$	-1	$Q = \sigma(\sqrt{3})$
Random Walk	$\log(\sigma) = -\frac{1}{2}\log(\tau) + \log(N)$	-1/2	$N = \sigma(1)$
Bias Instability	$\log[\sigma(f_0)] = \log\left(\sqrt{\frac{2\ln 2}{\pi}} \cdot B\right) \cong \log(0.664B)$	0	$B = \frac{\sigma(f_0)}{0.664}$
Rate random walk	$\log(\sigma) = \frac{1}{2}\log(\tau) + \log \frac{K}{\sqrt{3}}$	+1/2	$K = \sigma(3)$
Rate ramp	$\log(\sigma) = \log(\tau) + \log \frac{R}{\sqrt{2}}$	+1	$K = \sigma(\sqrt{2})$
Correlated Noise	$\log(\sigma) = \frac{1}{2}\log(\tau) + \log q_c T_c$ $\log(\sigma) = \frac{1}{2}\log(\tau) + \log \frac{q_c}{\sqrt{3}}$	$\pm 1/2$	$q_c T_c = \sigma(1)$ $q_c = \sigma(3)$
Sinsoidal noise		± 1	

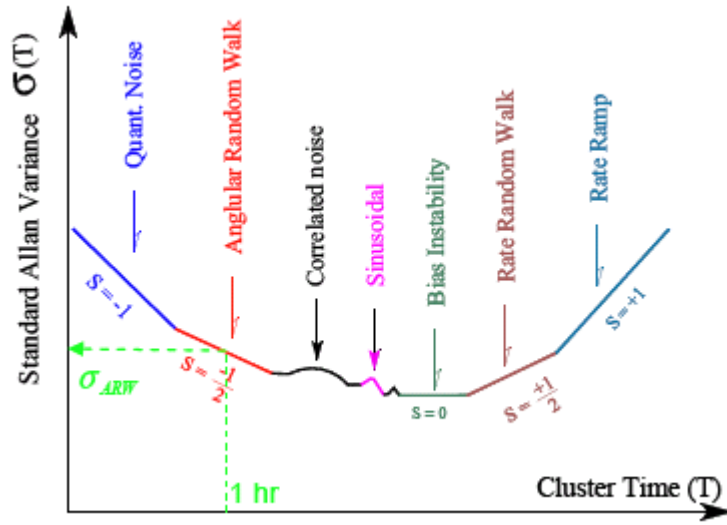


Fig. 2.7 Allan variance analysis noise terms results (IEEE Std. 952, 1997)

From the AV plot, it is straightforward to identify various random processes that exist in the raw data and get the noise coefficients with Table 2.2. For example, it is well known that the random walk is the noise terms in low cost inertial sensors. The angular random walk process can be identified at $T=1h$ and with a straight line of slope $-1/2$ as shown in Fig. 2.7. The noise PSD rate is represented by (IEEE Std.647, 1998):

$$S_x(f) = \sigma_{ARW}^2 \quad (2.25)$$

where σ_{ARW} is the angular random walk coefficient from Fig. 2.7. Substituting Eq. (2.25) in Eq. (2.23), we get

$$\sigma^2(\tau) = \frac{\sigma_{ARW}^2}{T} \quad (2.26)$$

In practice, the AV is based on a finite number of independent clusters that can be formed from any finite length of data. The AV of any noise terms is estimated using the total number of clusters of a given length that can be created. The confidence of the estimation improves as the number of independent clusters is increased. A straightforward calculation (Papoulis, 1991) shows that the percentage error is

$$\sigma = [2(\frac{N}{n} - 1)]^{-1/2} \times 100\% \quad (2.27)$$

where N is the total number of data set points, and n is the number of points in the cluster.

2.3 Testing results

The testing platform is based on ISIS-IMU from Inertial Science Inc., a six-degree of freedom inertial measurement unit designed for commercial use. It consists of three RRS75 (solid state rate sensors) and three solid state accelerometers, as shown in Fig. 2.8. The following testing results are all based on 2 hour static data from the ISIS-IMU.

Since there are 3 gyros and 3 accelerometers in the IMU, there would be $p \times 6$ parameters and 6 noise variances in the whole AR model if the model is all fixed to p -th order. The $\text{ARM}_{\text{FIT}}^{\text{TM}}$ package (Schneider and Neumaier, 2001) is applied to estimate the AR model and parameters. In the test with real data, we found that the order of AR model that minimize the SBC would vary with different length of data samples as well as different sensors, while the residual variance matrix would just change a little.

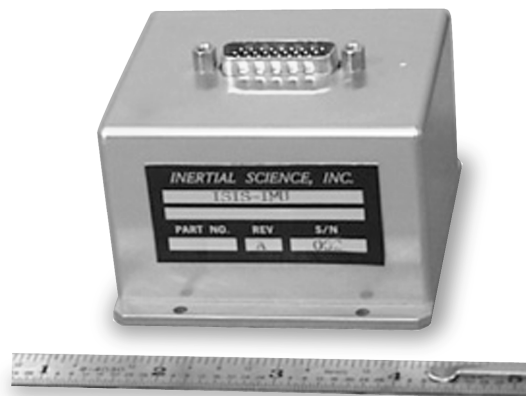


Fig.2.8 ISIS-IMU

One way to check whether the estimated order is valid is to verify the whiteness of the model residuals. From Fig.2.9, we can see that the residuals of AR (2) for the raw data of the accelerometer along y-axis is nearly white noise (within 95% confidence interval between two horizontal lines). Therefore, we can determine a second-order AR model (AR (2)) can fulfill the requirement of accuracy and can be applied for all gyros and accelerometers.

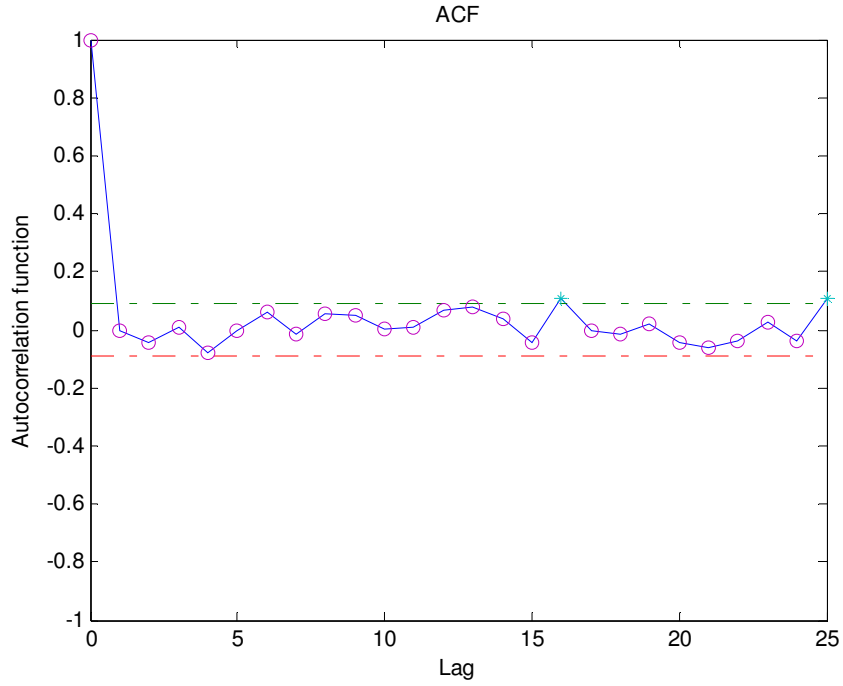


Fig.2.9 ACF of AR(2) residual for Y-Accelerometer

From the collected static data, we can get the PSD plot. A certain amount of noise would exist in the plotted curve due to the uncertainty. The PSD of the gyro along x-axis is shown in Fig.2.10.

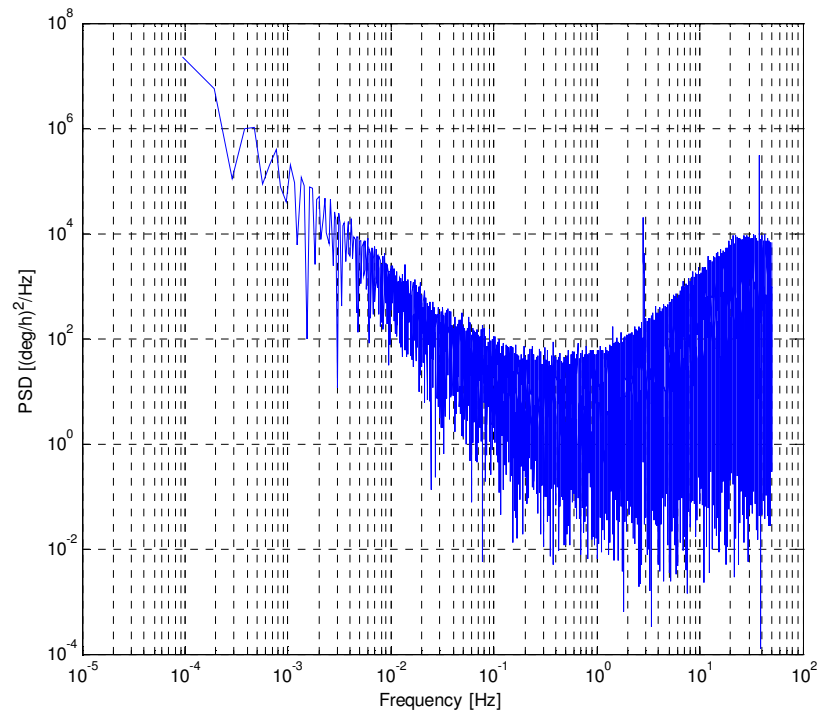


Fig.2.10 X-Gyro PSD

The bunching and flickers in high frequency of the PSD plot prevent us from further analysis. The frequency averaging technique (IEEE Std1293-1998) is used to average the points among the high frequencies of the PSD result in Fig. 2.10.

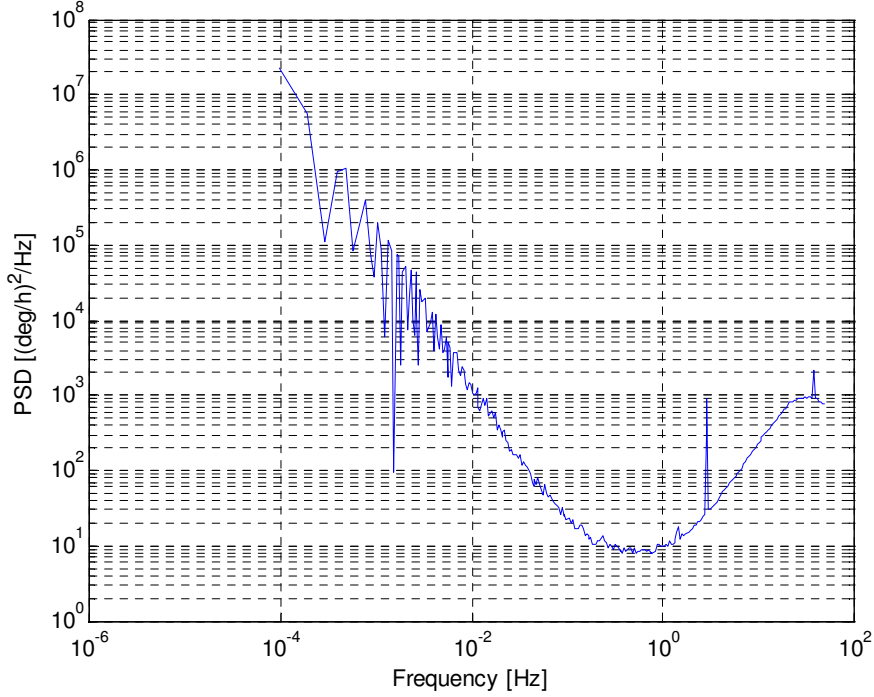


Fig.2.11 X-Gyro PSD after frequency averaging

The slopes of the PSD curve include -2 , 0 , and $+2$, which indicate that the X-axis gyro data contains rate random walk, random walk, and quantization noise, respectively. The acquisition of parameters for noise terms from the PSD result plot is complex. According to IEEE Std 952 1997, the random walk parameter is given by

$$Q(0/\sqrt{h}) = \frac{1}{60} \sqrt{\frac{1}{2} \text{PSD}[(0/\sqrt{h})^2 / \text{Hz}]} \quad (2.28)$$

Applying this equation, we obtain the random walk parameter about $0.024 \text{ deg/h}^{1/2}$. According to the specification, the random walk parameter is less than $0.5 \text{ deg/h}^{1/2}$. The reason why the experiment result is much less may be that these experiments are carried out during static situations. In dynamic situations it may go up.

The AD plot for gyros and accelerometers are shown in Fig. 2.12 and Fig. 2.14 respectively. The standard error of each AD can be evaluated with Eq. (2.27). The AD and its standard errors are shown in Fig. 2.13. With these two figures, we can identify different error coefficients by means of Table 2.2 as listed in Table 2.3 and Table 2.4 for gyros and accelerometers, respectively.

The reason for the big quantization noise for accelerometer Z may be that the specific

force sensed by the accelerometer along the vertical direction is overwhelmingly influenced by the gravity.

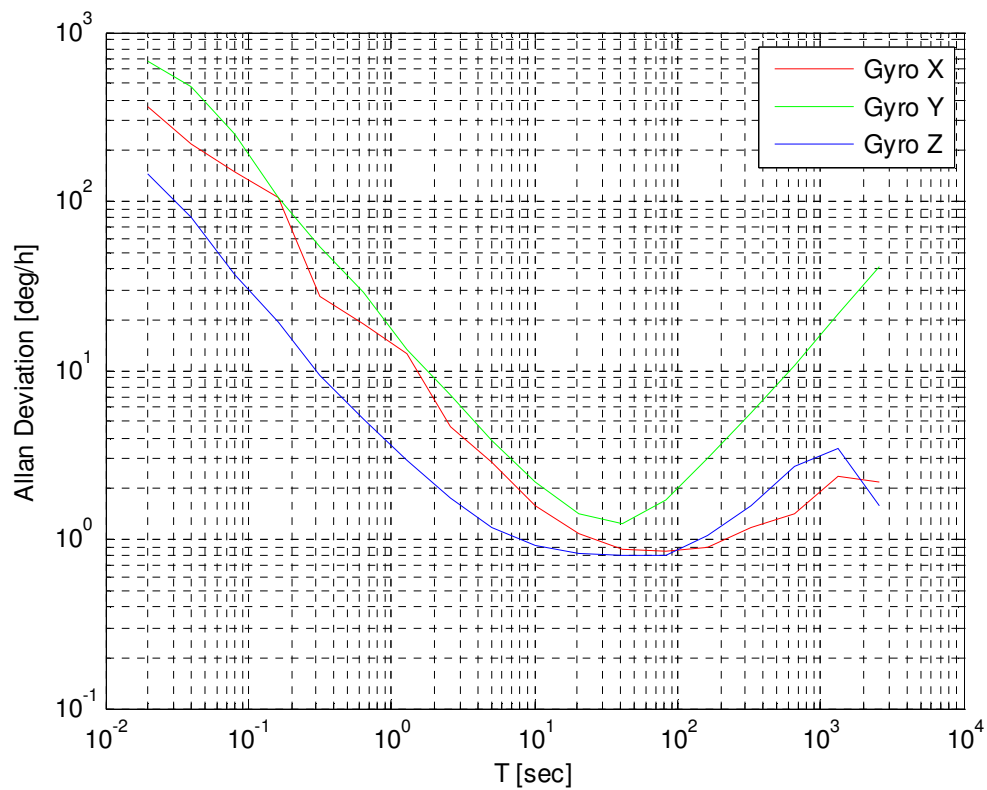


Fig.2.12 AD plot for 3 gyros

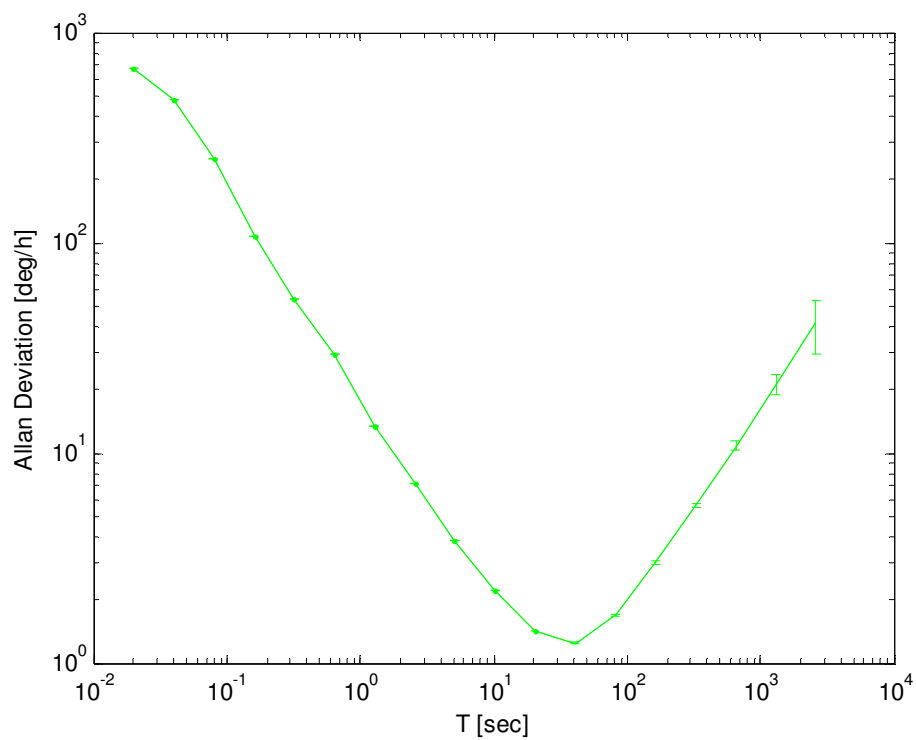


Fig.2.13 AD plot for Y-Gyro with percentage errors (1 bar)

Table 2.3 Error coefficients for gyros

	Gyros X	Gyros Y	Gyros Z
Quantisation noise (deg)	2.5×10^{-3}	7.2×10^{-3}	9.1×10^{-4}
Random walk (m/s/h ^{1/2})	4.8×10^{-2}	N/A	2.6×10^{-2}
Bias instability (deg/h)	1.4	N/A	1.2
Rate random walk (m/s/h ^{3/2})	5.2	N/A	8.1
Rate ramp (deg/h)	N/A	72	N/A

Table 2.4 Error coefficients for accelerometers

	Accelerometer X	Accelerometer Y	Accelerometer Z
Quantisation noise (deg)	4.5×10^{-3}	2.1×10^{-3}	4.8×10^{-1}
Bias instability (deg/h)	1.4	0.9	N/A
Rate ramp (deg/h)	N/A	3.2	N/A

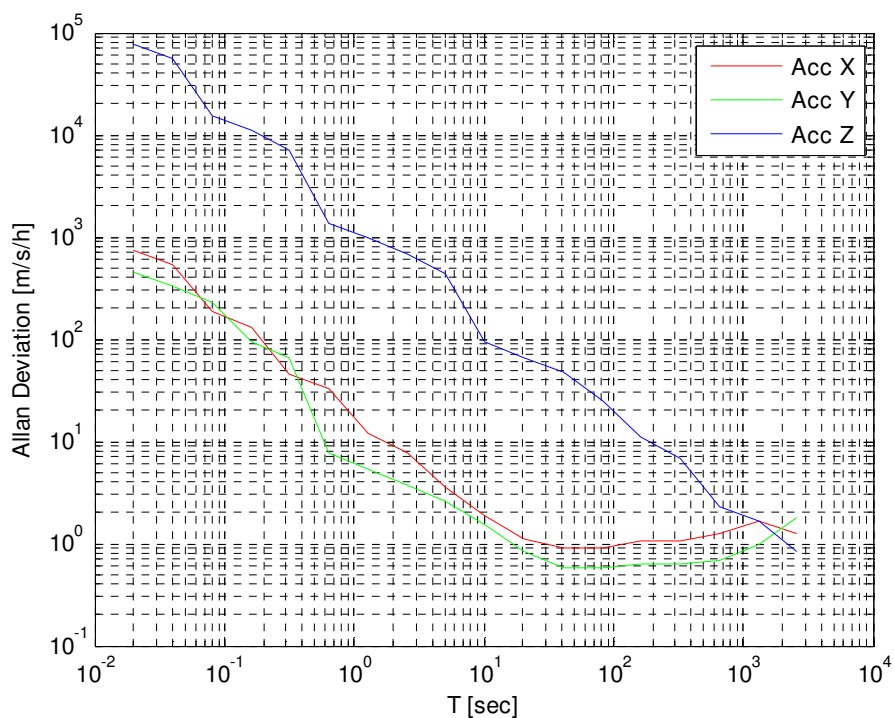


Fig.2.14 AD plot for 3 accelerometers

2.4 Discussion and conclusion

Using different models, we analyzed the sensor errors and got the parameters. We can use the AR model to estimate next consecutive errors, if we do not care much about different error components themselves and treat them totally together as stochastic errors. Although the optimal AR model would change its order with different sensors, a second-order AR model can be accepted as the unique model for all sensors of the IMU. Since the first-order GM process can be treated as a first-order AR model, if the sample interval is fixed, we do not specially analyze this most commonly used model.

The PSD is also a common tool to analyze both periodic and non-periodic signals. Frequency averaging is needed to get the exact slopes of the curve, but the parameters are difficult to compute directly. Also it may be somewhat difficult for the non-system analyst to understand, since it is in the frequency domain.

AV analysis has been widely used for MEMS stochastic modeling, due to its simplicity and effectiveness. It is a great advantage that the noise terms differ from region to region, so that we can obtain different noise coefficients at different regions of the AV plot (see Fig. 2.7). And there is also a relation between the AV and the PSD (see Eq. 2.23). However, a long static data record may be required, even dozens of hours, as we can see that the long-time slope is sometimes not fully determined within several hours of data.

Although different methods may give different values of the noise coefficients, the similar noise sources can be verified with a specific sensor. And it is expected that with longer data sets and many iterations a set of more accurate coefficients can be estimated. After estimating the coefficients for the main stochastic errors, a better error model can be applied in the GPS/INS integrated system, which is of much help for the Doppler frequency estimate and for bounding the error drift during GPS outage.

3. Integrated navigation structures

Due to the complementary character in many aspects, different GPS/INS integrated systems are hot research topics and widely applied during the last decade. With a GPS receiver we can obtain the position with bounded errors, while INS can provide high update frequency of both position and attitude information. Since the GPS positioning has a relatively low update rate, for high dynamic applications the measurements from INS can be used to determine the navigation solution between the intervals of GPS updates. When the new GPS update is available, we can correct the INS sensor errors in real time so that the errors of INS solutions are also bounded.

3.1 Three levels of integration

Initially, there are two broad integration structures, namely the loose and tight integration. In recent years there is a new structure proposed named ultra tight integration or deep integration. The main differences between these three integration structures are the different levels of integration, as shown in Fig. 3.1. There is also another set of terms: centralized and decentralized integration, which are used in many papers or literatures before. These two terms are the same with loose and tight integration. Since the term ultra tight integration came out, the terms loose and tight integration have been adopted in most recent cases.

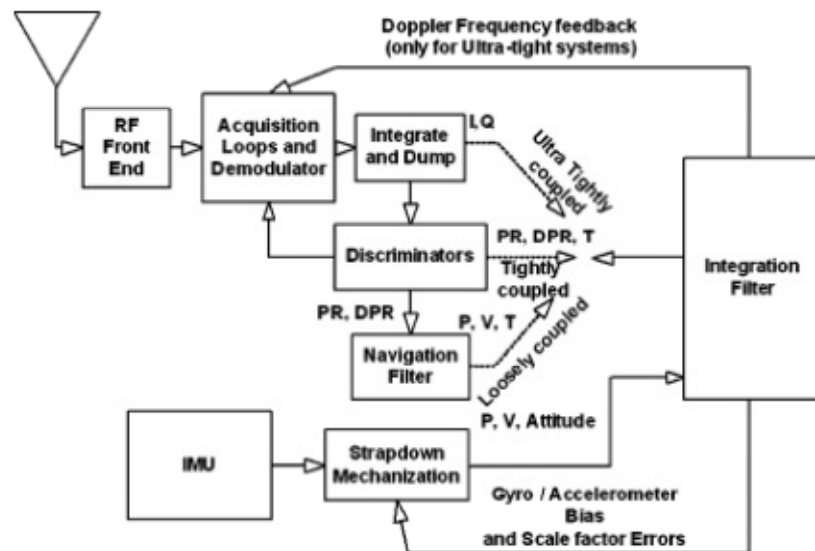


Fig. 3.1 Loose, tight and ultra tight GPS/INS navigation system (Babu and Wang, 2004)

For the tightly coupled system, the measurements from the INS sensors are combined

with the GPS pseudorange (PR) and delta pseudorange (DPR) measurements. The classic tight integration structure is shown in Fig. 3.2. A tightly integrated system can have a more accurate navigation solution, because the basic GPS observable are not as correlated as the position and velocity solutions, and it is possible to implement more sensitive fault detections and to use the GPS data even with less than 4 satellites available. GPS signals can help INS initial alignment, while INS data can be helpful to the GPS ambiguity resolution. But this integration responds more slowly to INS errors than the loosely coupled system.

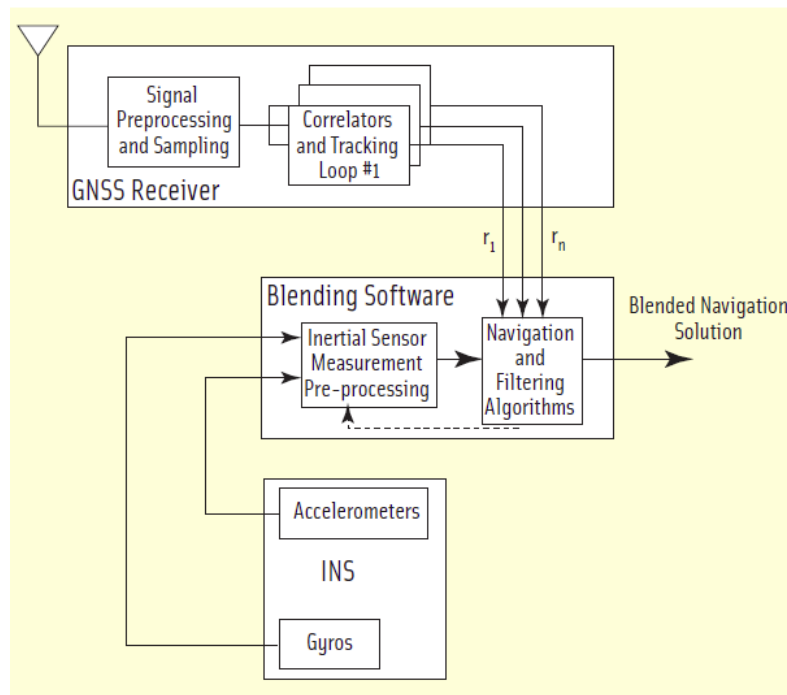


Fig. 3.2 Classic definition of GPS/INS tight integration (Gebre-Egziabher, 2007)
(r_i is the pseudorange between the receiver and satellite i)

There are great differences between the loose and tight architectures. In loosely coupled systems, the position, velocity and time from the GPS receiver are combined with position, velocity and attitude from INS by a mathematical depiction of the error characteristics of the systems by an Extended Kalman Filter (EKF). In such instances, the key character is that INS and GPS are both independent navigation systems. This is a kind of open loop collaboration with or without feedback to correct INS sensor errors and the performance of a loosely coupled integrated system greatly depends on the solution of GPS.

On the other hand in the tightly coupled system, the angular rate and specific force measurements from the INS are combined with the GPS pseudorange (PR) and delta pseudorange (DPR) measurements. In this sense, both INS and GPS are reduced to the basic sensors. A tightly integrated system can have more accurate navigation solution because the basic GPS observable are not as correlated as the position and

velocity solutions and more sensitive fault detection and isolation scheme can be used to verify the quality of PR and Doppler measurements.

Moreover, it is possible to use the GPS data even with less than 4 satellites available. Since only with four or more GPS satellite signals we can perform the GPS positioning, the great benefit of the tightly coupled system is that even less than four tracked satellites can still help to estimate INS errors and improve the precision of position and attitude (Horemuz and Sjöberg, 2001). GPS signals can help INS initial alignment, while INS data can be helpful to the GPS ambiguity resolution. But this integration responds more slowly to INS errors than the loosely coupled system. On the other hand, loose integration has the advantage of redundancy since both INS and GPS can provide their navigation solution independently.

A more complex and potentially more beneficial level of GPS/INS integration occurs at the GPS tracking-loops level named ultra-tight integration. In an ultra-tight coupled system, the inertial sensors are used to aid the GPS phase/frequency and code tracking loops directly. The measurements used are the in-phase and quadrature-phase signal. (For further details, we refer to Tsui, 2004.) The main feature of ultra tight integration is that the GPS signal tracking is combined into the estimation algorithm. Therefore the performances of GPS signal tracking in harsh environment and robustness can be improved. There is also a feedback loop to Doppler frequency estimate, by which it helps to regain the GPS signal tracking faster after the GPS blockage. This integration would offer reduced phase-tracking bandwidth, shorter acquisition time, and better resistance to radio frequency interference or multipath noise. Furthermore, this integration would be helpful to reduce phase-tracking bandwidth, less acquisition time, and more resistance to radio frequency interference or multipath noise. Therefore it has the potential to apply in high performance navigation systems where robustness to cycle slips is of paramount importance (Bhatti et al., 2007a; 2007b).

Along the direction from the loose integration to tight integration and ultra tight integration, the robustness of the integrated system is strengthened. However, the increased robustness is at the sacrifice of system simplicity, redundancy, and independence of the INS and GNSS navigators (Gebre-Egziabher, 2007).

3.2 Principle of loose integration for LVN

In this thesis a loose GPS/INS integration structure is implemented for different kinds of simulations and experiments. The background introduction of the GPS observations and basic INS navigation principles are given below. For detailed implementation of the integrated system, see Chapter 5.

3.2.1 GPS observations and positioning accuracy

As for GPS observations, here only basic phase and code pseudoranges are presented. We assume that the baselines are short, so that the atmospheric and satellite ephemeris errors are negligible. The phase and code observation equations are given as follows (Hofmann-Wellenhof et al., 2001).

$$\begin{aligned}
 L_1^i - \rho_0^i &= a_X^i \delta X + a_Y^i \delta Y + a_Z^i \delta Z + N_1^i \lambda_1 + \varepsilon_{L1}^i \\
 L_2^i - \rho_0^i &= a_X^i \delta X + a_Y^i \delta Y + a_Z^i \delta Z + N_2^i \lambda_2 + \varepsilon_{L2}^i \\
 C_1^i - \rho_0^i &= a_X^i \delta X + a_Y^i \delta Y + a_Z^i \delta Z + \varepsilon_{C1}^i \\
 C_2^i - \rho_0^i &= a_X^i \delta X + a_Y^i \delta Y + a_Z^i \delta Z + \varepsilon_{C2}^i
 \end{aligned} \tag{3.1}$$

where superscript refers to satellite, subscripts 1, 2 are the signal frequencies, L and C are phase and code observables, respectively, ρ_0^i is range between the receiver and satellite computed from approximate receiver coordinates, N_1^i and N_2^i are ambiguities, λ_1 and λ_2 are the wavelengths, ε_{L1}^i , ε_{L2}^i , ε_{C1}^i , ε_{C2}^i are observation random errors for phase and code observables at frequencies L_1 and L_2 , respectively. Please note that all terms in Eq. (3.1) are double-differences when used in tight integration. These double-difference phase and code pseudoranges would be derived in Chapter 5.

Assuming uncorrelated data, the covariance matrix of the double-difference observations is given by (Leick, 1999)

$$\mathbf{R} = 4\sigma_{L1}^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda_2^2 / \lambda_1^2 & 0 & 0 \\ 0 & 0 & k^2 & 0 \\ 0 & 0 & 0 & k^2 \lambda_2^2 / \lambda_1^2 \end{bmatrix} \tag{3.2}$$

where σ_{L1} , σ_{L2} , σ_{C1} , σ_{C2} are the standard error of the uncorrelated and undifferenced phase and code observables on two frequencies, k is the ratio between code and phase standard error, i.e.

$$k = \frac{\sigma_{C1}}{\sigma_{L1}} = \frac{\sigma_{C2}}{\sigma_{L2}} \quad (3.3)$$

We also assume that the phase ambiguities in Eq. (3.1) are fixed, so that these terms can be moved to the left side of the equations. Otherwise we should add two ambiguities into the state vector each time one satellite is locked-on.

For the GPS positioning accuracy, here we give out two example positioning curves with two typical GPS applications: Precise Point Positioning (PPP) and Real Time Kinematic (RTK). The PPP is to use a single receiver to determine the absolute position at one point, and an example of PPP positioning error curve is shown in Fig. 3.3.

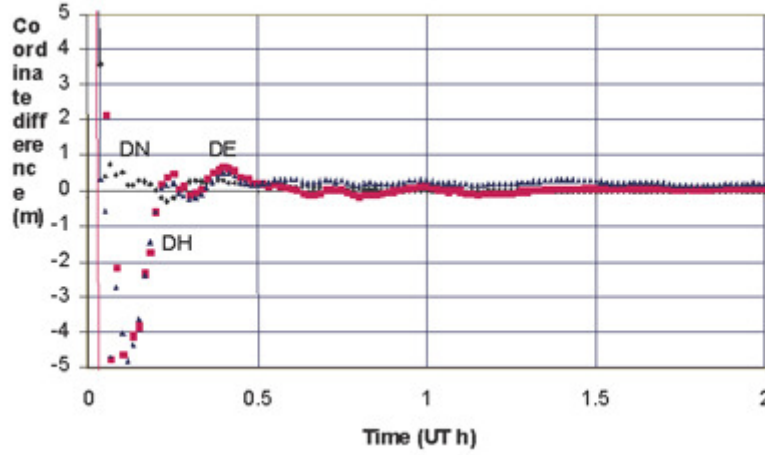


Fig. 3.3 Static PPP positioning error (Chen et al., 2006).
(DN,DE and DH are coordinate differences in north, east and height directions, respectively)

RTK is another common technique used in land survey based on phase observation. The RTK configuration contains a static reference station and moving rover receivers. It starts with the initialization stage, during which the phase ambiguities are solved. Then the reference station provides the real-time corrections to the rovers, providing up to centimetre level accuracy. The standard RTK configuration with one reference station works satisfactory for rover stations within 10 km from the reference station (Sjöberg, 2007, p. 68). An example of RTK positioning error curve is shown in Fig. 3.4.

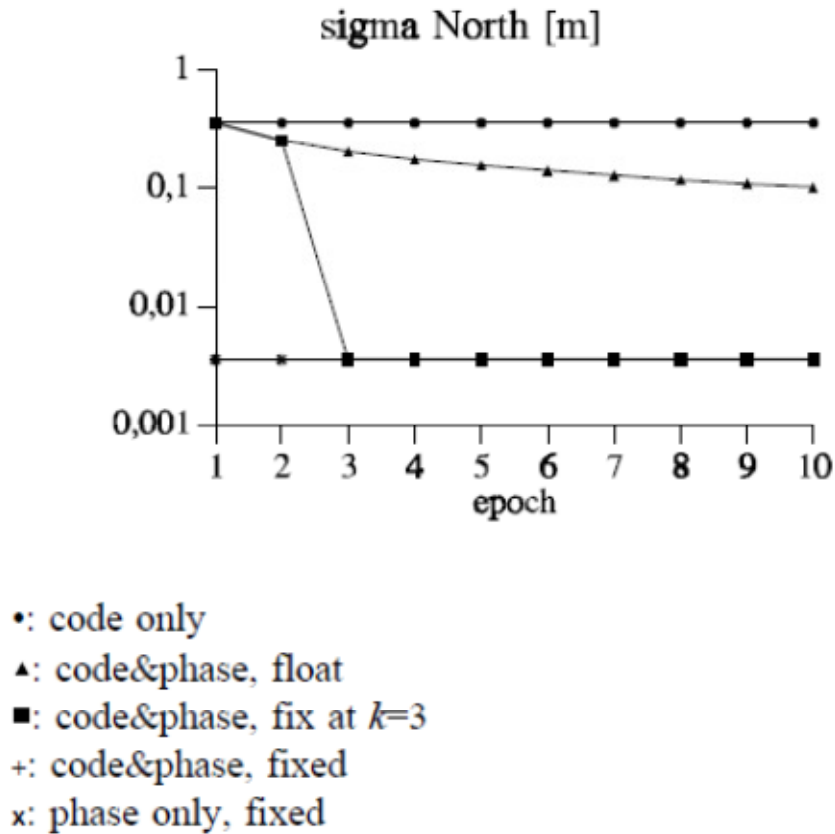


Fig. 3.4 Precision of North coordinates in kinematic mode (Tiberius, 1998, p. 174).

Please note that the “+” and “x” symbols are coincident with each other at the first two epochs and then with the square symbols in the following epochs.

3.2.2 INS and navigation principle

An INS usually contains an embedded microcomputer unit and a platform or module containing a set of accelerometers and gyroscopes. The initial position and attitude of the INS should be provided from another sources (GPS, compass etc.), and thereafter computes the updated position and velocity by integrating measurements from the accelerometers and gyroscopes. INS requires no external information to determine its position, velocity and orientation, once it has been initialized. Therefore it is autonomous, immune to jamming or deception and quite suitable for military applications.

Before introducing the INS navigation principle, two basic concepts of coordinate frames and rotation matrix are explained.

3.2.2.1 Coordinate frames

When describing the position on or near the Earth’s surface, we need to define the

coordinate system. The commonly used frames in inertial navigation systems are inertial frame (i-frame), conventional terrestrial frame (e-frame), navigation frame (n-frame) and body frame (b-frame).

By the physical definition of inertial frame, it is the coordinate system in which the Newton's laws of motion hold. In real world, a global inertial system is at best an abstraction, since any frame in vicinity of the solar system is permeated by a gravitational field that possesses spatially varying gradients (Jekeli, 2001). For the applications, the operational i-frame is defined as:

Origin - at the center of mass of the Earth
z-axis - parallel to the mean spin axis of the Earth
x-axis - pointing towards the mean vernal equinox
y-axis - completing a right-handed orthogonal frame

The e-frame is an Earth-Centered-Earth-fixed Cartesian frame, and the WGS-84 used by GPS as the reference coordinate system is one of the realization of e-frame. It is defined as:

Origin - at the center of mass of the Earth
z-axis - parallel to the mean spin axis of the Earth
x-axis - pointing towards the mean meridian of Greenwich
y-axis - completing a right-handed orthogonal frame

The n-frame is commonly used to describe the navigation of a vehicle in a local coordinate frame. It is defined as a set of Cartesian coordinate axes too, usually in the direction of north-east-down (NED) as shown in Figure 3.5:

Origin - at the origin of the sensor frame
z-axis – aligned with the ellipsoidal normal at a point, in the down direction
x-axis - pointing to north (parallel to the tangent to the meridian)
y-axis – pointing to the east and completing a right-handed orthogonal frame

The b-frame refers to the vehicle to be navigated and the conventional definition is along the forward, right and through-the-floor directions:

Origin - at the origin of the sensor frame
z-axis – aligned with the ellipsoidal normal at a point, in the down direction
x-axis - pointing to forward (the direction of instant motion)
y-axis – pointing to the right and completing a right-handed orthogonal frame

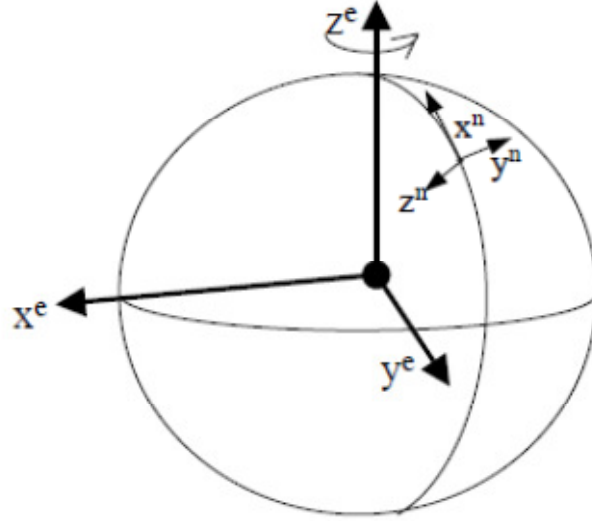


Fig. 3.5 The conventional terrestrial frame and the navigation frame (Shin, 2001)

3.2.2.2 Rotation matrix and Euler angles

The rotation matrix is also named direction cosine matrix which is an orthogonal matrix. It is the matrix that transforms the coordinates from one frame to another. That can be shown as

$$\mathbf{p}^A = \mathbf{R}_B^A \mathbf{p}^B \quad (3.4)$$

where \mathbf{p}^A and \mathbf{p}^B are the same vector expressed in arbitrary frames A and B , respectively, and matrix \mathbf{R}_B^A transforms the coordinates of a vector expressed in the B -frame into the coordinates of the same vector expressed in the A -frame.

The relative orientation between two frames can be described by a sequence of rotations. The rotation matrix around each axis are given by

$$\mathbf{R}(\theta_x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & \sin \theta_x \\ 0 & -\sin \theta_x & \cos \theta_x \end{bmatrix} \quad (3.5 \text{ a})$$

$$\mathbf{R}(\theta_y) = \begin{bmatrix} \cos \theta_y & 0 & -\sin \theta_y \\ 0 & 1 & 0 \\ \sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \quad (3.5 \text{ b})$$

and

$$\mathbf{R}(\theta_z) = \begin{bmatrix} \cos \theta_z & \sin \theta_z & 0 \\ -\sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.5 \text{ c})$$

where $\mathbf{R}(\theta_i)$ represents the rotation about the i -axis by the angle θ_i , positive in the counterclockwise sense as viewed along the axis toward the origin (right-hand rule). Each rotation matrix is orthogonal, i.e. $\mathbf{R}^{-1}(\theta_i) = \mathbf{R}^T(\theta_i)$. If the A-frame is the result of rotating the B-frame, first about the z-axis by θ_z , then about the new y-axis by θ_y , and finally about the new x-axis by θ_x , we get the total transformation matrix given by

$$\mathbf{R}_B^A = \mathbf{R}(\theta_x)\mathbf{R}(\theta_y)\mathbf{R}(\theta_z) \quad (3.6)$$

And this total transformation matrix is also orthogonal. It is particularly important that the transformation matrix depends on the sequence of the rotations, i.e.

$$\mathbf{R}(\theta_x)\mathbf{R}(\theta_y) \neq \mathbf{R}(\theta_y)\mathbf{R}(\theta_x) \quad (3.7)$$

The explicit form of Eq. (3.6) is given by

$$\mathbf{R}_B^A = \begin{bmatrix} \cos \theta_z \cos \theta_y & \sin \theta_z \cos \theta_y & -\sin \theta_y \\ -\sin \theta_z \cos \theta_x + \cos \theta_z \sin \theta_y \sin \theta_x & \cos \theta_z \cos \theta_x + \sin \theta_z \sin \theta_y \sin \theta_x & \cos \theta_y \sin \theta_x \\ \sin \theta_z \sin \theta_x + \cos \theta_z \sin \theta_y \cos \theta_x & -\cos \theta_z \sin \theta_x + \sin \theta_z \sin \theta_y \cos \theta_x & \cos \theta_y \cos \theta_x \end{bmatrix} \quad (3.8)$$

Then the Euler angles can be computed from Eq. (3.8) as

$$\begin{aligned} \theta_x &= \arctan \frac{\mathbf{R}(2,3)}{\mathbf{R}(3,3)} \\ \theta_y &= \arcsin[-\mathbf{R}(1,3)] \\ \theta_z &= \arctan \frac{\mathbf{R}(1,2)}{\mathbf{R}(1,1)} \end{aligned} \quad (3.9)$$

where $\mathbf{R}(i,j)$ is the element of \mathbf{R}_B^A matrix at the i -th row and j -th column. $\mathbf{R}(1,1)$ and

$\mathbf{R}(3,3)$ are not equal to zero, since the angular increments during the update interval are normally small angles. If the Euler angles can be treated as small angles, i.e. these

approximations hold

$$\begin{aligned}\sin \theta &\doteq \theta \\ \cos \theta &\doteq 1\end{aligned}\tag{3.10}$$

then Eq. (3.8) can be simplified as

$$\mathbf{R}_b^a = \begin{bmatrix} 1 & \theta_z & -\theta_y \\ -\theta_z & 1 & \theta_x \\ \theta_y & -\theta_x & 1 \end{bmatrix} = \mathbf{I} - \mathbf{E}\tag{3.11}$$

where \mathbf{E} is the skew-symmetric matrix of Euler angles:

$$\mathbf{E} = \begin{bmatrix} 0 & -\theta_z & \theta_y \\ \theta_z & 0 & -\theta_x \\ -\theta_y & \theta_x & 0 \end{bmatrix}\tag{3.12}$$

Two basic operations with skew-symmetric matrix are given as follows:

$$\mathbf{A}^T = -\mathbf{A}\tag{3.13}$$

$$\mathbf{a} \times \mathbf{b} = \mathbf{A}\mathbf{b} = -\mathbf{B}\mathbf{a}\tag{3.14}$$

where \mathbf{A} and \mathbf{B} is the skew-symmetric matrices of vectors \mathbf{a} and \mathbf{b} , respectively.

3.2.2.3 IMU and navigation equations

The sensors or equipments used in an INS are called inertial measurement units (IMU), which are the main components of INSs used in aircraft, spacecraft, and watercraft, as well as guided missiles. An IMU consists of two main parts: three orthogonal accelerometers and three orthogonal gyroscopes, which sense the vehicle accelerations and angular velocities, respectively. Please note that the mechanic gyroscope and accelerometer is introduced in the following part just to show the basic principle of IMU. For MEMS, there are no such mechanic gyroscope and accelerometer components, but it is based on nanoelectromechanical systems and nanotechnology.

A gyroscope is a device for measuring or maintaining orientation, based on the principles of conservation of angular momentum. An accelerometer is a device that measures the acceleration along a specific axis. It should be noted that the IMU

accelerometers sense only specific forces in INS (See Eq. 3.20). This concept is of vital importance in inertial navigation.

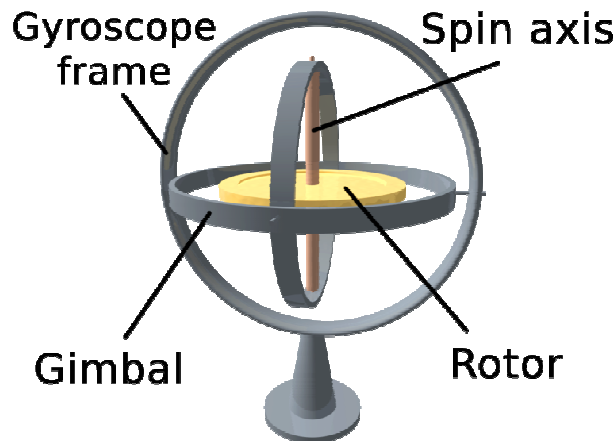


Fig. 3.6 The structure of a TDF gyroscope (Wikipedia, Gyroscope figure)

Gyroscope: A two-degree of freedom (TDF) rigid rotor gyro is shown in Fig. 3.6. The rotor of a gyro is based on Newton's second law

$$F^i = \frac{dp^i}{dt} \quad (3.15)$$

where the notation F is the force, $p^i = mr^i$ is the linear momentum.

If we take the cross product between r^i at each side of Eq. (3.15), then we obtain

$$r^i \times F^i = r^i \times \frac{dp^i}{dt} \quad (3.16)$$

And from Eq. (3.16), we can see that the left side is the moment of the force about the origin of the coordinate system. The right side is the time derivative of the angular momentum L^i . As a result, this equation can also be written as

$$\frac{dL^i}{dt} = r^i \times F^i \quad (3.17)$$

From Eq. (3.17), we can draw the conclusion that the torque applied to a particle is equal to the moment of force acting on it. If the particle is rotating around the origin, then it becomes

$$\frac{dL^i}{dt} = 0 \quad (3.18)$$

which means that L is constant. The physical meaning is that for a particle moving in a central field, the torque is always zero, or the angular momentum is always constant. This also means that the motion of such a particle is in a plane.

That is the principle that a mechanical gyro helps to maintain a specific frame (i.e. navigation frame) during the motion and sense the rotations of the body frame and then align the axis with the gyro axis with the aid of some servo motors.

Accelerometer: The first accelerometer was actually a gravity meter as shown in Fig. 3.7. In fact, all kinds of gravity meters are accelerometers. Based on Newton's Second Law, the equation of motion is

$$\mathbf{f} = m\mathbf{a} \quad (3.19)$$

where \mathbf{f} is the force, m is the mass, $\mathbf{a} = \ddot{\mathbf{x}}$ is the acceleration. Therefore, accelerometers are built on the principle of measuring the force exerted on a test body of a known mass along a given axis.

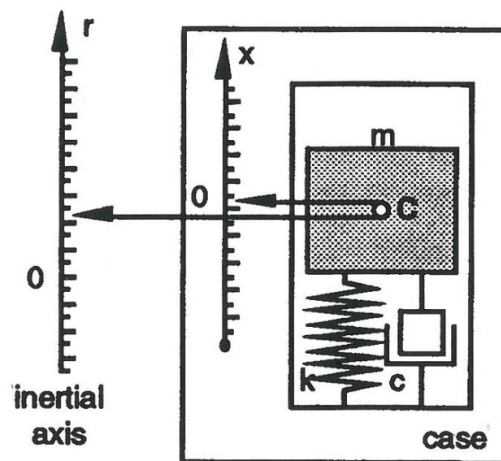


Fig 3.7 Schematic structure of an accelerometer (Horemuz 2006, p. 25)

The compression of the spring is given by Hooke's Law: $f_s = -kX$, where k is the spring coefficient, X is the spring displacement. Therefore the equation of motion can be written as

$$m\ddot{x} = f_s = -kX \quad (3.20)$$

A complete force analysis can be shown as

$$-kx(C) - c\dot{x}(C) + mg = m\ddot{r}(C) \quad (3.21)$$

where C is the mass center, c is the damping coefficient.

Since $r(C) = x(C) + r(o)$, where o is the initial balance point, Eq. (3.21) can be rewritten as

$$\ddot{x}(C) + \frac{k}{m}x(C) + \frac{c}{m}\dot{x}(C) = g - \ddot{r}(o) \quad (3.22)$$

Therefore we can see that the output of an accelerometer in the vicinity of the Earth is: $g - \ddot{r}(o)$, which is called specific force. If the accelerometer is stationary on the Earth surface, i.e. $\ddot{r}(o) = 0$, the output is gravity g .

Navigation equation: The fundamental equation for the motion in the gravitational field of the Earth is

$$\ddot{\mathbf{r}}^i = \mathbf{f}^i + \overline{\mathbf{g}}^i \quad (3.23)$$

where \mathbf{r}^i is the position vector from the origin of the inertial frame to the vehicle;

\mathbf{f}^i is the specific force vector;

$\overline{\mathbf{g}}^i$ is the gravitational acceleration vector.

The raw observations from the strapdown INS (SINS) are two orthogonal triads of specific forces and angular velocities in the b-frame. Both of them must be properly scaled and corrected for the systematic errors. We store the scale factors for the accelerometers in vector $\mathbf{S}_a^T = (S_{ax} \ S_{ay} \ S_{az})$, scale factors for the gyroscopes in $\mathbf{S}_g^T = (S_{gx} \ S_{gy} \ S_{gz})$, gyro drifts in vector \mathbf{d} and accelerometer biases in vector \mathbf{b} . Then the corrected accelerometer observations are computed as

$$\mathbf{f}^b = \begin{pmatrix} f_x^b \\ f_y^b \\ f_z^b \end{pmatrix} = \begin{pmatrix} \tilde{f}_x^b S_{ax} \\ \tilde{f}_y^b S_{ay} \\ \tilde{f}_z^b S_{az} \end{pmatrix} - \mathbf{b} \quad (3.24)$$

and the corrected gyro observations

$$\boldsymbol{\omega}_{ib}^b = \begin{pmatrix} \omega_{ibx}^b \\ \omega_{iby}^b \\ \omega_{ibz}^b \end{pmatrix} = \begin{pmatrix} \tilde{\omega}_{ibx}^b S_{gx} \\ \tilde{\omega}_{iby}^b S_{gy} \\ \tilde{\omega}_{ibz}^b S_{gz} \end{pmatrix} - \mathbf{d} \quad (3.25)$$

where $(\tilde{\mathbf{f}}^b), (\tilde{\boldsymbol{\omega}}_{ib}^b)$ are the raw output of the accelerometers and gyroscopes, respectively.

In the following part, we will derive a simple set of inertial navigation equations: (For more detailed navigation equations in different frames, we refer to Jekeli 2001, pp. 123-138; Horemuz 2006, pp. 33-43). If the angular velocity can be assumed as constant during the time interval Δt , we can get the rotation vector as

$$\boldsymbol{\theta}_{ib}^b = \int_{t_{k-1}}^{t_k} \boldsymbol{\omega}_{ib}^b dt \approx \begin{pmatrix} \omega_{ibx}^b \Delta t & \omega_{iby}^b \Delta t & \omega_{ibz}^b \Delta t \end{pmatrix}^T \quad (3.26)$$

where $\Delta t = t_k - t_{k-1}$.

The rotation matrix in the navigation frame is then

$$\mathbf{R}_{b(k)}^n = \mathbf{R}_{b(k-1)}^n \left(\mathbf{I} - \mathbf{S} \frac{\sin \theta_b}{\theta_b} + \mathbf{S}^2 \frac{1 - \cos \theta_b}{\theta_b^2} \right) \quad (3.27)$$

where $\theta_b = \|\boldsymbol{\theta}_b\|$, $\mathbf{S} = \begin{pmatrix} 0 & -\theta_z & \theta_y \\ \theta_z & 0 & -\theta_x \\ -\theta_y & \theta_x & 0 \end{pmatrix}$ and

$$\boldsymbol{\theta}_b = (\theta_{bx} \quad \theta_{by} \quad \theta_{bz})^T = \left(\frac{\omega_{bx}(t_k) + \omega_{bx}(t_{k-1})}{2} \Delta t \quad \frac{\omega_{by}(t_k) + \omega_{by}(t_{k-1})}{2} \Delta t \quad \frac{\omega_{bz}(t_k) + \omega_{bz}(t_{k-1})}{2} \Delta t \right)^T \quad (3.28)$$

The velocity update is given as follows

$$\mathbf{V}^n(t_k) = \mathbf{V}^n(t_{k-1}) + \Delta \mathbf{V}^n \quad (3.29)$$

where $\Delta \mathbf{V}^n = \mathbf{R}_b^n \Delta \mathbf{V}^b + \mathbf{g}^n \Delta t$ and $\Delta \mathbf{V}^b = \frac{1}{2} (\mathbf{f}^b(t_k) + \mathbf{f}^b(t_{k-1})) \Delta t$ and $\mathbf{g}^n = \mathbf{R}_b^n \mathbf{g}^b$.

Finally, the position update is

$$\mathbf{P}^n(t_k) = \mathbf{P}^n(t_{k-1}) + \mathbf{V}^n(t_{k-1})\Delta t + \frac{1}{2}(\mathbf{g}^n + \mathbf{f}^n)\Delta t^2 \quad (3.30)$$

where $\mathbf{f}^n = \mathbf{R}_b^n (\mathbf{f}^b(t_k) + \mathbf{f}^b(t_{k-1})) / 2$

Please note that these equations are valid in such situations where the sculling error vanishes, i.e. the angular rate and specific force is constant during the integration time interval, and this is no velocity rotation, i.e. the velocity direction is the same. The sculling error is the integrated contribution of the high-frequency content in velocity update (For more details about sculling error, we refer to Jekeli 2001, pp. 124-157).

Inertial navigation is based on the integration of sensed specific forces and angular velocities with respect to time. Therefore it deals with the solution of a series of differential equations. Since the mathematical platform is adopted in SINS, i.e. the attitude matrix is derived by means of real-time computations, the attitude update algorithm is the crucial part in INS navigation, because it is the main factor that influences the navigation accuracy.

There are several traditional attitude update algorithms, namely Euler angle algorithm, direction cosine algorithm and quaternion algorithm. Each method has its advantages and disadvantages. The quaternion algorithm is widely used in practice due to its simplicity (For more detailed information, we refer to Jekeli 2001, pp. 9-19; Horemuz 2006, pp. 12-24).

3.2.3 INS sensors calibration

Computation of sensor biases from static observations: Using data from a static IMU in different positions enables us to estimate some systematic errors of the sensors: accelerometer biases and gyro drifts. The basic observations should be performed in the following way:

- 1- Place the IMU on a table, as precisely leveled as possible. Orient x^b axis towards north, y^b axis towards east and z^b axis downwards (position 1).
- 2- Record the static data for about 15 seconds. Compute mean values of specific forces ($\tilde{\mathbf{f}}_1$) and of angular velocities ($\tilde{\boldsymbol{\omega}}_1$).
- 3- Stop the recording and rotate the IMU by 180° around z^b , so now x^b axis points to the south. Then again record around 15 seconds of data. The mean values are denoted by $\tilde{\mathbf{f}}_2$ and $\tilde{\boldsymbol{\omega}}_2$ (position 2).

In position 1, the accelerometers x and y should sense zero specific force, if the table is perfectly leveled and the accelerometers are bias-free. If these conditions are not fulfilled (which is usually the case), then we can write the following equations:

$$\begin{aligned}\tilde{f}_{1x} &= g_x + b_x \\ \tilde{f}_{1y} &= g_y + b_y\end{aligned}\tag{3.30}$$

Similarly for position 2:

$$\begin{aligned}\tilde{f}_{2x} &= -g_x + b_x \\ \tilde{f}_{2y} &= -g_y + b_y\end{aligned}\tag{3.31}$$

The bias can be computed:

$$\mathbf{b} = \frac{1}{2}(\tilde{\mathbf{f}}_1 + \tilde{\mathbf{f}}_2)\tag{3.32}$$

The corrected measurements can be computed by:

$$\mathbf{f} = \tilde{\mathbf{f}} - \mathbf{b}\tag{3.33}$$

In position 1, the gyroscope should sense just the north, east and down components of the Earth's rotation:

$$\boldsymbol{\omega}_{ie}^n = (\omega_e \cos\varphi \quad 0 \quad -\omega_e \sin\varphi)^T\tag{3.34}$$

So the gyro drifts can be computed as:

$$\begin{aligned}d_x &= \tilde{\omega}_x - \omega_e \cos\varphi \\ d_y &= \tilde{\omega}_y \\ d_z &= \tilde{\omega}_z + \omega_e \sin\varphi\end{aligned}\tag{3.35}$$

and the measurement angular rates are then corrected by:

$$\boldsymbol{\omega} = \tilde{\boldsymbol{\omega}} - \mathbf{d}\tag{3.36}$$

3.2.4 INS Initialization and Alignment

Before an IMU gets ready to use, the b-frame axes must be aligned with the navigation frame axes, or the direction cosines matrix between the b-frame and n-frame must be established, i.e. the rotation matrix \mathbf{R}_b^n .

The initial position and velocity required as integration constants in the primary navigation equations are simply inserted into the system as input data. Leveling a gimbaled platform is accomplished by monitoring the output of two mutually perpendicular accelerometers with their input axes parallel to the platform. If the platform is stationary, then the accelerometers sense only the components of the gravity acceleration vector. The platform is leveled when its vertical axis is aligned with the direction of gravity; then the horizontal accelerometers on the platform sense zero specific force. Thus, an approximate leveling ensures by rotating the platform pitch and roll gimbals using the servo motors, such that the output of each horizontal accelerometer vanishes.

3.2.4.1 Coarse alignment and fine alignment

For the strapdown IMUs a coarse alignment can be done in an analytical and numerical way. Assuming a stationary IMU, the accelerometers sense only the components of the gravity vector, and the gyros sense only the angular velocity of Earth rotation

The coarse alignment procedures are shown as follows (for detailed derivation, we refer to Horemuz 2006, pp. 56-59): First compute the mean values of the raw data

$$\begin{aligned} f_x &= \text{mean}(\tilde{f}_x); \quad f_y = \text{mean}(\tilde{f}_y); \quad f_z = \text{mean}(\tilde{f}_z) \\ \omega_x &= \text{mean}(\tilde{\omega}_x); \quad \omega_y = \text{mean}(\tilde{\omega}_y); \quad \omega_z = \text{mean}(\tilde{\omega}_z) \end{aligned} \quad (3.37)$$

With these mean values, we can compute the direction cosines matrix between the b-frame and n- frame $\mathbf{R}_n^b = (\mathbf{R}_n^b)^T$.

$$\mathbf{R}_n^b = \begin{bmatrix} \frac{\omega_x}{\omega_e \cos \varphi} - \frac{f_x \sin \varphi}{g \cos \varphi} & \frac{f_z \omega_y - f_y \omega_z}{g \omega_e \cos \varphi} & -\frac{f_x}{g} \\ \frac{\omega_y}{\omega_e \cos \varphi} - \frac{f_y \sin \varphi}{g \cos \varphi} & \frac{f_x \omega_z - f_z \omega_x}{g \omega_e \cos \varphi} & -\frac{f_y}{g} \\ \frac{\omega_z}{\omega_e \cos \varphi} - \frac{f_z \sin \varphi}{g \cos \varphi} & \frac{f_y \omega_x - f_x \omega_y}{g \omega_e \cos \varphi} & -\frac{f_z}{g} \end{bmatrix} \quad (3.38)$$

Similar with Eq. (3.8) and (3.9), we can obtain the misalignment angles from Eq. (3.37)

$$\text{roll} = \arctan\left(\frac{f_y}{f_z}\right)$$

$$\text{pitch} = \arctan\left(\frac{f_x}{\sqrt{f_y^2 + f_z^2}}\right) \quad (3.39)$$

$$\text{yaw} = \arctan\left(\frac{f_z\omega_y - f_y\omega_z}{g\omega_x - f_x\omega_e \sin\varphi}\right)$$

which are the orientation angles.

Little angle errors would be more or less left between the practical and ideal alignments after a coarse alignment. This is due to the systematic errors in the sensors that cannot be calibrated in the lab, particularly biases that have different values each time the system is turned on. Also there would be different kinds of stochastic errors in each session as shown in Chapter 2. The coarse alignment is refined based on the assumption that external information provides means to estimate the systematic instrument errors. The calibration proceeds on the basis of external observations (both position and velocity) in a Kalman filter.

4. Robust Kalman filtering with constraints

The Kalman filter is named after Rudolf E. Kalman. In 1960 he published the classic paper proposing a recursive solution to the discrete-data linear filtering problem (Kalman, 1960). The Kalman filter is essentially a set of mathematical equations that implement a predictor-corrector type estimator optimally minimizing the estimated error covariance. From that time, the Kalman filter has been the subject of extensive research and application, particularly in the area of autonomous or assisted navigation (Welch and Bishop, 2001).

4.1 Extended Kalman Filter

The Kalman filter deals with the state estimate of a discrete-time process of a linear stochastic difference equation. But in real applications most of the process equations are non-linear. A Kalman filter in this linearized situation is referred to as an Extended Kalman Filter (EKF). The state functions need not to be linear functions of the state but instead be differentiable functions.

For a dynamic system the state transition and observation models can be given as

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{w}_{k-1} \quad (4.1)$$

$$\mathbf{z}_k = h(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1} \quad (4.2)$$

where \mathbf{x} is the state vector, \mathbf{z} is the measurement, the subscript is the discrete time index, f and h are the functions to compute predicted state and predicted measurements, respectively, \mathbf{w}_k and \mathbf{v}_k are the process and observation noises, which are both assumed to be zero-mean Gaussian noises with covariance \mathbf{Q}_k and \mathbf{R}_k , respectively, i.e. $\mathbf{w}_k \sim N(0, \mathbf{Q}_k)$ and $\mathbf{v}_k \sim N(0, \mathbf{R}_k)$.

As the same with Kalman filter, the EKF algorithm has two main steps, i.e. predict and update steps:

$$\text{Predicted state estimate } \hat{\mathbf{x}}_{k|k-1} = f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_{k-1}) \quad (4.3)$$

$$\text{Predicted covariance estimate } \mathbf{P}_{k|k-1} = \mathbf{F}_{k-1} \mathbf{P}_{k-1|k-1} \mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1} \quad (4.4)$$

$$\text{Kalman gain } \mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (4.5)$$

$$\text{Updated state estimate } \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \tilde{\mathbf{y}}_k \quad (4.6)$$

$$\text{Updated covariance estimate } \mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \quad (4.7)$$

where the innovation is $\tilde{\mathbf{y}}_k = \mathbf{z}_k - h(\hat{\mathbf{x}}_{k|k-1})$, and the state transition and observation

matrices are defined by the following two Jacobians: $\mathbf{F}_{k-1} = \frac{\delta f}{\delta \mathbf{x}}|_{\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_{k-1}}$ and $\mathbf{H}_{k-1} = \frac{\delta h}{\delta \mathbf{x}}|_{\hat{\mathbf{x}}_{k-1|k-1}}$.

The first two equations above are the prediction, and the last two are the update, both of which use the Gain matrix \mathbf{K} shown in Eq. (4.5). The circular progress of the EKF is shown in Fig. 4.1.

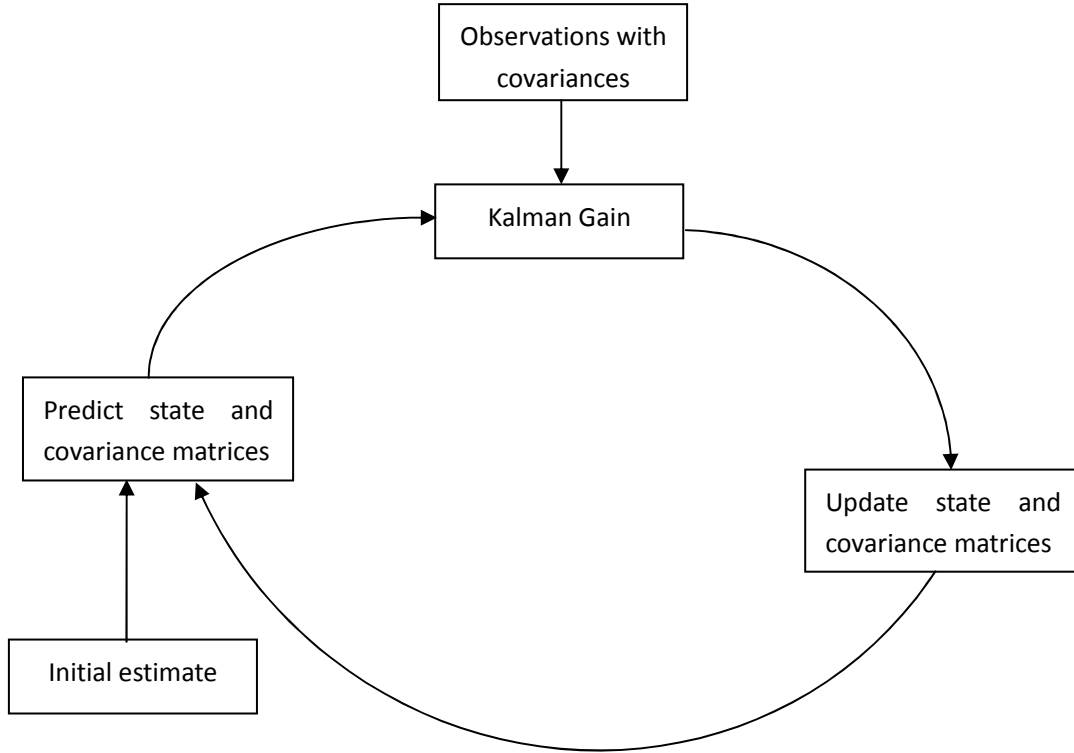


Fig. 4.1 The circular progress of EKF

4.2 Land vehicle model

For LVN we can build up a dynamic model for this specific application, so that we can use the dynamic model to further reduce the navigation errors. Before we build the LVN model, some physical assumptions must be given. We assume that the vehicle does not slip in the lateral direction, which is reasonable for travel in a

constant direction in absence of ice or snow (Li et al., 2010). A second assumption is that the vehicle stays on the ground. Both of these two assumptions are fulfilled in most cases and we can accept them as the typical conditions for LVN.

In this thesis we study the kinematic models of LVN, i.e. velocity and height constraints, including the lever arm correction when turning the direction. A third possible constraint could be obtained by map matching. Both the original nonlinear kinematic equations and the linearized ones for velocity and height constraints will be presented. Then we use these models to constrain the navigation errors and correct the INS error drifts in real time. In this way, better positioning accuracy and longer INS stand-alone period can be reached compared to not taking these constraints into account.

4.2.1 Velocity constraint

With both of the two assumptions above, the velocity of the vehicle in the direction perpendicular to the movement of the vehicle must be zero (i.e. $v_x^b = v_z^b = 0$), as shown in Fig. 4.2. The IMU sensor on the vehicle is responsible for providing the longitudinal acceleration measurements, while the velocity constraints provide the assumed lateral and vertical velocities (i.e. we assume that the lateral and vertical acceleration measurements are just white noises, if the vehicle velocity does not change the direction within a short interval. If this assumption does not hold, see the lever arm effect in Section 4.2.3).

In this way, the three-dimensional (3D) velocity update can be reached. Since this 3D velocity update adds the physical information of the LVN to the Kalman filter, it can improve the practical navigation performance.

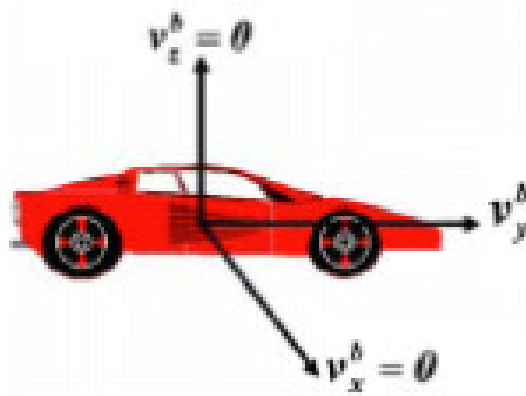


Fig. 4.2 Vehicle velocity constraints (Godha and Cannon, 2007)

The body-frame velocity error equation is derived as follows:

$$\mathbf{v}^b = (\mathbf{R}_b^n)^T \mathbf{v}^n \quad (4.8)$$

Perturbing this equation, we obtain (Shin, 2001)

$$\begin{aligned} \mathbf{v}^b + \delta \mathbf{v}^b &= [(\mathbf{I} - \mathbf{E}^n) \mathbf{R}_b^n]^T (\mathbf{v}^n + \delta \mathbf{v}^n) \\ &= [\mathbf{R}_n^b (\mathbf{I} - \mathbf{E}^n)^T] (\mathbf{v}^n + \delta \mathbf{v}^n) \\ &= \mathbf{R}_n^b (\mathbf{I} + \mathbf{E}^n) (\mathbf{v}^n + \delta \mathbf{v}^n) \end{aligned} \quad (4.9)$$

where \mathbf{E}^n is the skew-symmetric matrix of misalignment angle $\boldsymbol{\epsilon}^n$. Please note that from Eq. (4.8) to Eq. (4.9), Eq. (3.11) is applied since $\boldsymbol{\epsilon}^n$ is quite small; in Eq. (4.9), Eq. (3.13) is used at the end.

Therefore we can get

$$\begin{aligned} \delta \mathbf{v}^b &= \mathbf{R}_n^b (\delta \mathbf{v}^n + \mathbf{E}^n \mathbf{v}^n) \\ &= \mathbf{R}_n^b \delta \mathbf{v}^n + \mathbf{R}_n^b (\boldsymbol{\epsilon}^n \times) \mathbf{v}^n \\ &= \mathbf{R}_n^b \delta \mathbf{v}^n - \mathbf{R}_n^b (\mathbf{v}^n \times) \boldsymbol{\epsilon}^n \end{aligned} \quad (4.10)$$

where Eq. (3.14) is applied at the end.

These equations indicate that the velocity constraints not only help to improve the velocity estimation, but also aids in attitude estimation, particularly the estimation of roll and pitch (Godha and Cannon, 2005).

We can assume the following two constraints as measurement updates to the Kalman filter:

$$v_x^b \approx 0 \quad (4.11)$$

$$v_z^b \approx 0 \quad (4.12)$$

So two misclosure elements

$$\begin{bmatrix} \delta v_x^b \\ \delta v_z^b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} v_x^b \\ v_z^b \end{bmatrix} \quad (4.13)$$

can be added in the state vector of navigation errors.

Due to the misalignment angles, the practical velocities along the lateral and vertical directions are not just zero-mean noises. We can get the velocity attitude from the projection of the forward velocity with respect to the lateral and vertical directions (Shin, 2001). If the misalignment angle is θ , then the projection components along

the X-axis direction are $v_y^b \sin \theta$ as shown in Fig. 4. 3. From this relation we can approximately estimate the magnitude of the measurement noise for velocity constrains.

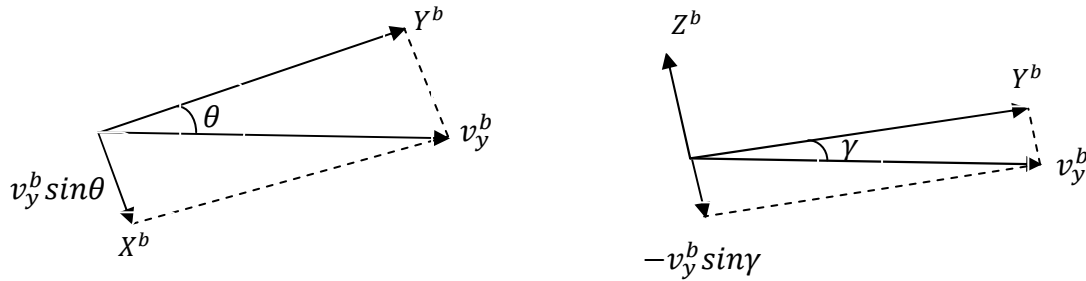


Fig. 4.3 Projection of the forward velocity (Shin, 2001)

4.2.2 Height constraint

Since a vehicle generally stays on a road, the height of the vehicle does not vary more than a few tens of metres within a short period or a small region. Therefore, the height solution computed before a GPS outage can be used as a height constraint during the GPS outage. As we all know, a 3D position solution requires observations from four satellites to solve for the horizontal position, vertical position and the clock errors. If the height is constrained to a known value, then the remaining unknowns can be solved with one less satellite. This can potentially help to improve the overall horizontal solution accuracy, during periods of partial GPS outages.

The relationship between the Cartesian coordinates and the geodetic coordinates is established as shown in Fig. 4.4 (Jekeli 2001, p. 23)

$$\begin{aligned} x &= (N + h) \cos \phi \cos \lambda \\ y &= (N + h) \cos \phi \sin \lambda \\ z &= [N(1 - e^2) + h] \sin \phi \end{aligned} \quad (4.14)$$

where x, y, z are the Cartesian coordinates of the vector $\mathbf{r} = (x, y, z)^T$; $N = \frac{a}{\sqrt{1 - e^2 \sin^2 \phi}}$

is the prime vertical radius of the curvature of the reference ellipsoid perpendicular to the meridian at point P ; ϕ, λ, h are the latitude, longitude and ellipsoidal height, and e is the first eccentricity.

Then we can get the error perturbations:

$$\begin{bmatrix} \delta r_x \\ \delta r_y \\ \delta r_z \end{bmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{bmatrix} \delta \phi \\ \delta \lambda \\ \delta h \end{bmatrix} \quad (4.15)$$

where

$$\begin{aligned} a_{11} &= T \cos \phi \cos \lambda - (N + h) \sin \phi \cos \lambda \\ a_{12} &= -(N + h) \sin \phi \sin \lambda \\ a_{13} &= \cos \phi \cos \lambda \\ a_{21} &= T \cos \phi \sin \lambda - (N + h) \sin \phi \sin \lambda \\ a_{22} &= (N + h) \cos \phi \cos \lambda \\ a_{23} &= \cos \phi \sin \lambda \\ a_{31} &= T(1 - e^2) \sin \phi + N(1 - e^2) \cos \phi + h \cos \phi \\ a_{32} &= 0 \\ a_{33} &= \sin \phi \end{aligned}$$

and the term T is given by

$$T = -\frac{dN}{d\phi} = \frac{-ae^2 \sin \phi \cos \phi}{\sqrt{(1 - e^2 \sin^2 \phi)^3}}$$

where a is the semi-major axis of the reference ellipsoid.

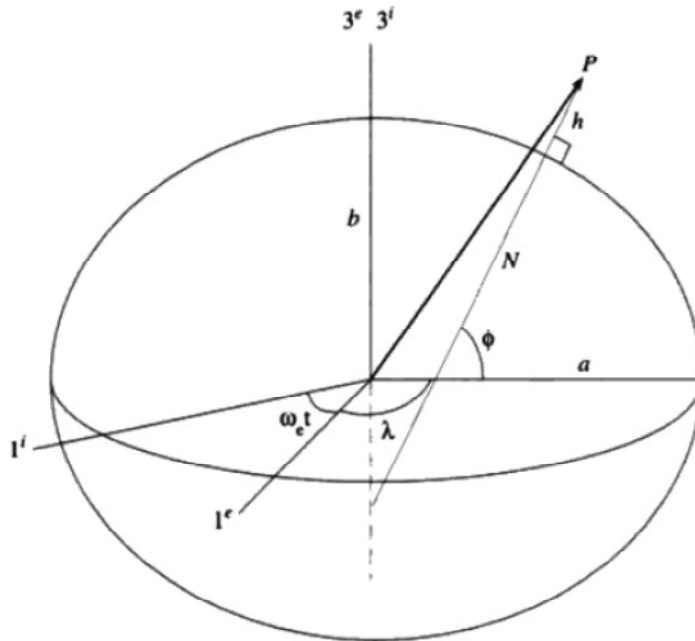


Fig. 4.4 i-frame, e-frame and geodetic coordinates (Jekeli 2001, p. 23)

In some tough environments like navigation in urban areas, it is quite common that the GPS receiver may lose the track of the signals from some satellites for a short time. In such cases, the height constraint can compensate for one satellite signal in relatively flat areas if only 3 satellite signals are tracked at the same time. Even less than 3 tracked satellites can still help to estimate the INS errors and improve the positional and attitude precision, especially the latter (Horemuz and Sjöberg, 2001).

4.2.3 Lever arm effect

Since the IMU can be mounted at different positions on the vehicle, there is always a lever arm between the mass centre of the vehicle and the IMU to the velocity constraints. This lever arm would cause different velocities for the GPS receiver and IMU, especially when the vehicle changes directions as shown in Fig. 4.5, where δ is the front steering angle, Ψ is the yaw angle, β is the side slip angle at the center of gravity (CG), L_{CG}^{IMU} is the lever arm between the IMU and CG of the vehicle. This figure illustrates a simplified 2D vehicle model, also known as the bicycle model, which has been used for a land vehicle model (Ouladsine et al., 2007).

If the velocity constraints hold at the CG point, i.e. zero velocity, the lateral velocity of the INU is not zero but can be given by (Li, 2009)

$$v_{IMU}^y = \dot{\Psi} L_{CG}^{IMU} \quad (4.16)$$

Therefore the lever arm effect on lateral velocity must be compensated before applying velocity constraints when the vehicle changes the direction. A similar analysis can be applied to the vertical velocity constraint.

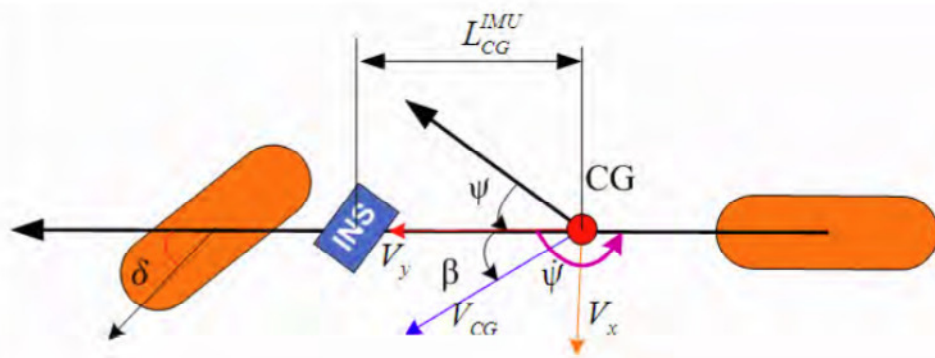


Fig. 4.5 2D vehicle model (Li, 2009)

4.3 Constrained Kalman filter

For the Kalman filter in GPS/INS integrated systems, there are a huge number of related papers. Most of these papers mainly deal with various improvements or applications on the algorithm itself. However, it is verified that Kalman filtering with state constraints can significantly improve the estimation accuracy of the filter (e.g. Simon and Chia, 2002).

For the specific applications like LVN, the more detailed information of the physical system is utilized to the Kalman filter, the better navigation perform it would get. With the velocity constrain and height constrain as well as the lever arm effect correction for LVN discussed above, we can set these constrains to the Kalman filter in the integration system.

There are various ways of implementing a Kalman filter with linear constrains, namely model reduction (Wen and Durrant-Whyte, 1992), perfect measurements (Wang et al. 2002), estimate projection (Simon and Chia, 2002), gain projection (Sircoulomb et al. 2008), probability density function truncation (Simon and Simon, 2010), and system projection (Ko and Bitmead, 2007). Under certain conditions, all these approaches result in the same state estimate (Simon and Chia, 2002), but each approach itself has its own advantages and disadvantages. For example, treating state constraints simply by reducing the system model parameters may lose the physical meaning of the state variables, and the approach of perfect measurements would cause singular noise covariance matrix.

Besides the integrated constrained Kalman filter mentioned above, there is also the two-stage recursive estimator, which considers measurements and constraints separately (Yang et al., 2010). The first step is the Kalman filtering without constraints, and the second step is a projecting process of the constraints.

For a discrete linear system

$$\mathbf{X}_k = \Phi_{k,k-1}\mathbf{X}_{k-1} + \mathbf{w}_k \quad (4.17)$$

$$\mathbf{Z}_k = \mathbf{A}_k\mathbf{X}_k + \mathbf{v}_k \quad (4.18)$$

where $\Phi_{k,k-1}$ is a transition matrix, \mathbf{A}_k is a design matrix, we assume the linearized constraints are given by the general equation

$$\mathbf{B}_k\hat{\mathbf{X}}_k - \mathbf{D}_k = 0 \quad (4.19)$$

where $\hat{\mathbf{X}}_k$ is the estimated state vector with covariance matrix $\Sigma_{\hat{\mathbf{X}}_k}$, \mathbf{B}_k is the design matrix, \mathbf{D}_k is a constant vector. Then the Lagrangian optimization condition

for the recursive Kalman filter is (Yang et al. 2010)

$$\Omega_k = \sum_{i=1}^{n_k} \mathbf{P}_{k_i} \rho(V_{k_i}) + (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k)^T \boldsymbol{\Sigma}_{\bar{\mathbf{X}}_k}^{-1} (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k) + 2\boldsymbol{\beta}_k^T (\mathbf{B}_k \hat{\mathbf{X}}_k - \mathbf{D}_k) = \min \quad (4.20)$$

where $\mathbf{P}_k = \boldsymbol{\Sigma}_k^{-1}$ is the weight matrix of \mathbf{Z}_k , $\bar{\mathbf{X}}_k$ is the predicted state vector with covariance matrix $\boldsymbol{\Sigma}_{\bar{\mathbf{X}}_k} = \boldsymbol{\Phi}_{k,k-1} \boldsymbol{\Sigma}_{\hat{\mathbf{X}}_k} \boldsymbol{\Phi}_{k,k-1}^T + \mathbf{Q}_k$, $\tilde{V}_{k_i} = V_{k_i} / \sigma_{V_{k_i}}$ is the standardized residual of $V_{k_i} = A_{k_i} \hat{X}_{k_i} - Z_{k_i}$, $\boldsymbol{\beta}_k$ is the Lagrangian vector, $\rho(\cdot)$ is a convex and continuous function, satisfying that:

- 1) it is differentiable;
- 2) $\rho(0) = 0$;
- 3) There exists a constant $c > 0$ such that ρ is strictly increasing on $[0, c]$ and constant on $[c, \infty]$.

Taking the partial derivative of Ω_k with respect to $\hat{\mathbf{X}}_k$ and let the derivatives be zero, we can obtain (For detailed derivation, see Appendix A)

$$\hat{\mathbf{X}}_k = (\mathbf{A}_k^T \bar{\mathbf{P}}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\mathbf{A}_k^T \mathbf{P}_k \mathbf{Z}_k + \mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k) - (\mathbf{A}_k^T \bar{\mathbf{P}}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} \mathbf{B}_k^T \boldsymbol{\beta}_k \quad (4.21)$$

where $\mathbf{P}_{\bar{\mathbf{X}}_k} = \boldsymbol{\Sigma}_{\bar{\mathbf{X}}_k}^{-1}$ is the weight matrix of predicted state vector $\bar{\mathbf{X}}_k$, $\bar{\mathbf{P}}_k$ is the equivalent weight matrix of the measurements \mathbf{Z}_k , as the robust variation against outliers, which will be introduced in Eq. (4.32).

Denoting the initial condition

$$\hat{\mathbf{X}}_k^0 = \bar{\mathbf{P}}_{\hat{\mathbf{X}}_k^0}^{-1} (\mathbf{A}_k^T \mathbf{P}_k \mathbf{Z}_k + \mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k) \quad (4.22)$$

with covariance matrix

$$\bar{\mathbf{P}}_{\hat{\mathbf{X}}_k^0} = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{Z}_k + \mathbf{P}_{\bar{\mathbf{X}}_k}) = \bar{\boldsymbol{\Sigma}}_{\hat{\mathbf{X}}_k^0}^{-1} \quad (4.23)$$

we can rewrite Eq. (4.21) as

$$\hat{\mathbf{X}}_k = \hat{\mathbf{X}}_k^0 - \bar{\boldsymbol{\Sigma}}_{\hat{\mathbf{X}}_k^0} \mathbf{B}_k^T \boldsymbol{\beta}_k \quad (4.24)$$

Considering the constraint equation

$$\mathbf{B}_k \hat{\mathbf{X}}_k - \mathbf{D}_k = \mathbf{B}_k \left(\hat{\mathbf{X}}_k^0 - \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} \mathbf{B}_k^T \boldsymbol{\beta}_k \right) - \mathbf{D}_k = 0 \quad (4.25)$$

we get the Lagrangian vector

$$\boldsymbol{\beta}_k = \left(\mathbf{B}_k \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} \mathbf{B}_k^T \right)^{-1} (\mathbf{B}_k \hat{\mathbf{X}}_k^0 - \mathbf{D}_k) \quad (4.26)$$

So the estimation in Eq. (4.21) can be rewritten as

$$\hat{\mathbf{X}}_k = \hat{\mathbf{X}}_k^0 - \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} \mathbf{B}_k^T \left(\mathbf{B}_k \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} \mathbf{B}_k^T \right)^{-1} (\mathbf{B}_k \hat{\mathbf{X}}_k^0 - \mathbf{D}_k) \quad (4.27)$$

Let the discrepancy vector be

$$\mathbf{S}_k = \mathbf{B}_k \hat{\mathbf{X}}_k^0 - \mathbf{D}_k \quad (4.28)$$

Then the covariance matrix becomes

$$\Sigma_{\mathbf{S}_k} = \mathbf{B}_k \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} \mathbf{B}_k^T \quad (4.29)$$

Finally we obtain (see Appendix B)

$$\hat{\mathbf{X}}_k = \hat{\mathbf{X}}_k^0 - \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} \mathbf{B}_k^T \Sigma_{\mathbf{S}_k}^{-1} \mathbf{S}_k \quad (4.30)$$

with covariance matrix

$$\Sigma_{\hat{\mathbf{X}}_k} = \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} - \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} \mathbf{B}_k^T \Sigma_{\mathbf{S}_k}^{-1} \mathbf{B}_k \bar{\Sigma}_{\hat{\mathbf{X}}_k^0} \quad (4.31)$$

It should be mentioned that this two-stage recursive estimator is equivalent to the integrated constrained Kalman filter.

In this thesis the two-stage recursive estimator is preferred as the implementation of constrained Kalman filter. For the loose GPS/INS integration, the physical constraints can be applied to the Kalman filter directly as the additional observation equations, since the filter deals with the navigation solutions from each of the systems in the loose integration structure.

On the other hand, in the tight integration, the situation becomes more complicated.

With the general introduction of tight GPS/INS integration architecture in Chapter 3, the diagrammatical structure of the tight GPS/INS integration with velocity and height constraints is shown in Fig. 4.6. We do not mean to give a detailed theoretical analysis about this section in this thesis. For more detailed information on this part, we refer to Godha (2006); Godha and Cannon (2007) and Li (2009).

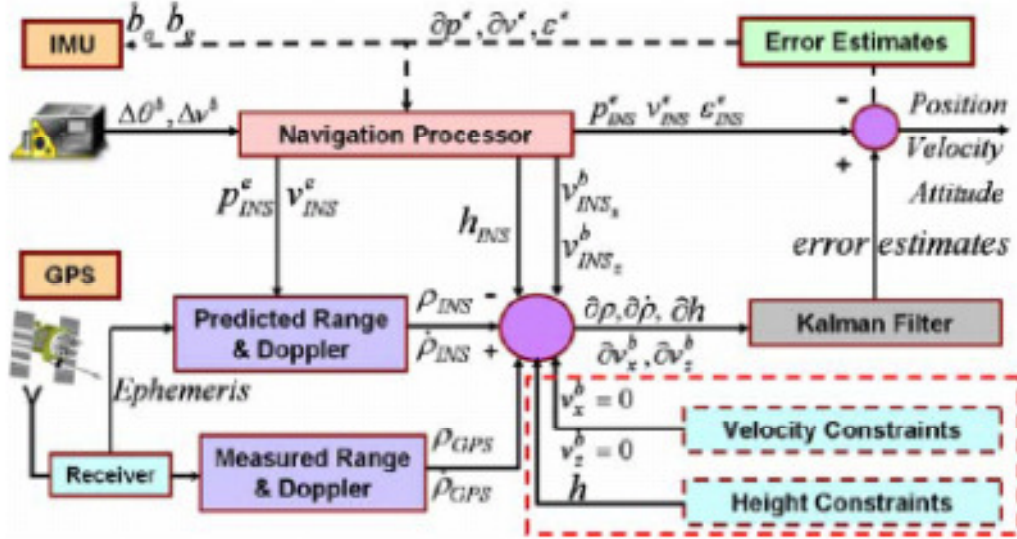


Fig. 4.6 Kalman filter with constraints in a tight integrated system (Godha, 2006)

4.4 Outliers and robust filter

An important hazard in the navigation system is the sensor outliers, which could result in a great increase of the positioning error. In statistics, an outlier is an observation that is numerically distant from the rest of the data. An outlying observation, or outlier, is one that appears to deviate markedly from other members of the sample in which it occurs (Grubbs, 1969).

There are many anomalous causes of outliers, including sensor malfunction, error data transmission and operating mistakes. No matter which kind of the causes, we must have some methods to detect and eliminate or reduce the influence of the outliers once they merge into the measurements. However, the standard Kalman filter is not robust to outliers. Therefore different variations of the Kalman filter have been proposed to deal with this problem.

The most common way to get rid of outliers is setting a rejection threshold. However, simply using a rejection threshold is not a good way to eliminate them, since this method can cause the measurement weights to jump from the original value to zero if the measurement errors are located near both sides of the threshold. Moreover this method requires manual parameter tuning, or heuristics parameter estimation

procedures. Another simple approach is treating the state constraints by reducing the system model parameters or as perfect measurements, but this method would lose the physical meaning of the state variables. A third method is treating the state equality constraints as perfect measurements with zero measurement noise, but this would result in a singular measurement noise covariance matrix (Simon and Chia, 2002).

To avoid these disadvantages, robust maximum likelihood estimation (M- estimation) is applied to the Kalman filter along with kinematic constraints to give smooth weights for both measurement errors and kinematic model errors (Yang et al., 2010). For this robust Kalman filter, the weight of each measurement is adaptively changed according to its standardized residual. In this way the outliers can be rejected and the contribution of measurements with larger errors can be suppressed by decreasing the weight to avoid causing significant rise of the standard errors.

To obtain the robust Kalman filter, the equivalent weight is computed first. There are many equivalent weight functions that we can choose, such as Huber weight function (Huber, 1981), and bifactor reduction factor (Yang et al., 2002; Yang et al., 2010). The latter is given below

$$\bar{P}_{k_i} = \begin{cases} P_{k_i} & |\tilde{V}_{k_i}| \leq c_0 \\ P_{k_i} \frac{c_0}{|\tilde{V}_{k_i}|} \left(\frac{c_1 - |\tilde{V}_{k_i}|}{c_1 - c_0} \right) & c_0 < |\tilde{V}_{k_i}| \leq c_1 \\ 0 & c_1 < |\tilde{V}_{k_i}| \end{cases} \quad (4.32)$$

where $\tilde{V}_{k_i} = V_{k_i}/\sigma_{V_{k_i}}$ is the standardized residual of V_{k_i} ; c_0 and c_1 are two constants with empirical values: $c_0 = 1.0 \sim 1.5$ and $c_1 = 4.5 \sim 8.5$. There are many empirical rules to choose these constants, and Xu (1993) proposed a theoretical method, namely to determine the constants based on the confidence intervals of the parameter estimates.

Based on all the discussion above in this chapter, both the constrained physical equations and the equivalent weight matrix are adopted in the EKF. As a result the robust constrained Kalman filter is obtained and it will, hopefully, improve the navigation performance for LVN applications.

5. Tests and analyses

In this chapter, we give out the detailed description of the integration system model and the integration structure. Based on these models, we have performed a series of simulations and tests.

5.1 Implementation of the integrated system with EKF

To implement the Kalman filter for the integrated navigation system, first we should build up the system model, i.e. to determine the state vector and derive the system dynamic equations as well as observation equations. Based on the system model, the EKF introduced in Section 4.1 can be implemented.

5.1.1 State vector and system dynamic equations

The state vector contains the unknown parameters to be estimated in the Kalman filter. In this thesis we choose 15 parameters for the state vector of the GPS/INS integrated navigation system. The parameters can be divided into two groups as:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1^T & \mathbf{x}_2^T \end{bmatrix}^T \quad (5.1)$$

where the first group contains three dimensional errors in position, velocity and orientation of the IMU:

$$\mathbf{x}_1 = \begin{bmatrix} (\delta \mathbf{r}^e)^T & (\delta \mathbf{v}^e)^T & (\boldsymbol{\varepsilon}^e)^T \end{bmatrix}^T = \begin{bmatrix} \delta x_i, \delta y_i, \delta z_i, \delta v_x, \delta v_y, \delta v_z, \varepsilon_x, \varepsilon_y, \varepsilon_z \end{bmatrix}^T \quad (5.2)$$

and the second group is the sensors errors:

$$\mathbf{x}_2 = \begin{bmatrix} \mathbf{d}^T & \mathbf{b}^T \end{bmatrix}^T \quad (5.3)$$

The position and velocity of the vehicle can be computed:

$$\begin{aligned} x_i &= x_{i0} + \delta x_i, & v_x &= v_{x0} + \delta v_x \\ y_i &= y_{i0} + \delta y_i, & v_y &= v_{y0} + \delta v_y \\ z_i &= z_{i0} + \delta z_i, & v_z &= v_{z0} + \delta v_z \end{aligned} \quad (5.4)$$

where x_{i0}, y_{i0}, z_{i0} are coordinates, v_{x0}, v_{y0}, v_{z0} are velocities as delivered by IMU.

The orientation can be computed from the transformation matrix \mathbf{R}_b^e :

$$\hat{\mathbf{R}}_b^e = (\mathbf{I} - \mathbf{E}^e) \tilde{\mathbf{R}}_b^e \quad (5.5)$$

$$\mathbf{E}^e = \begin{bmatrix} 0 & -\epsilon_z & \epsilon_y \\ \epsilon_z & 0 & -\epsilon_x \\ -\epsilon_y & \epsilon_x & 0 \end{bmatrix} \quad (5.6)$$

and the orientation relative to the navigation frame becomes

$$\mathbf{R}_n^b = \mathbf{R}_e^b \cdot \mathbf{R}_n^e \quad (5.7)$$

The IMU sensors errors in the second group contain two parts:

The accelerometer's error $\delta a(t)$ consists of the constant (b) and random (w_A) parts:

$$\delta a(t) = b + w_A(t) \quad (5.8)$$

where

- b is the accelerometer bias, which is modelled as random constant. This bias gets a random value within certain limit, usually specified as bias repeatability. This bias must be determined after each start of IMU.
- $w_A(t)$ is random walk noise of the accelerometer

Similarly, the error of the gyro can be described as

$$\delta \omega(t) = d + w_G(t) \quad (5.9)$$

where

- d is the gyro drift. The same remark holds as for b.
- $w_G(t)$ is random walk noise of the gyro

The sensor errors are defined as

$$\begin{aligned} \hat{\omega} &= \tilde{\omega} + \delta \omega \\ \hat{a} &= \tilde{a} + \delta a \end{aligned} \quad (5.10)$$

The INS errors are modelled by a linear system described generally by the following dynamic state equation

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t) \mathbf{x}(t) + \mathbf{G}(t) \mathbf{u}(t) \quad (5.11)$$

and the measurement model

$$\mathbf{z}(t) = \mathbf{H} \mathbf{x}(t) + \mathbf{v}(t) \quad (5.13)$$

where \mathbf{x} is the state vector, \mathbf{F} is the system dynamic matrix, \mathbf{G} is input matrix, t is time and \mathbf{u} is a vector forcing function, whose elements are white noise. The discrete solution of Eq. (5.11) is

$$\mathbf{x}_k = \Phi_{k,k-1} \mathbf{x}_{k-1} + \mathbf{w}_k \quad (5.12)$$

where $\Phi_{k,k-1}$ is the transition matrix from epoch $k-1$ to k and \mathbf{w}_k is the noise in the state vector.

The Kalman filter can be applied iteratively based on this discrete solution. In the following part, we present the solutions of each key variables and steps in the Kalman filter. (For more information, we refer to Farrell 1999, pp. 84-86; Horemut 2006, pp. 68-70):

The relationship between the transition matrix Φ and dynamic matrix \mathbf{F} can be expressed by

$$\Phi_{k+1,k} = e^{\mathbf{F} \cdot \Delta t} \approx \mathbf{I} + \mathbf{F}_k \Delta t + \frac{1}{2} \mathbf{F}_k^2 \Delta t^2 + \frac{1}{3!} \mathbf{F}_k^3 \Delta t^3 + \frac{1}{4!} \mathbf{F}_k^4 \Delta t^4 + \dots \quad (5.14)$$

is the transition matrix between epochs k and $k+1$, where $\Delta t = t_{k+1} - t_k$ and

$$\mathbf{w}_k = \int_k^{k+1} \Phi_{k+1,\tau} \mathbf{G}_\tau \mathbf{u}_\tau d\tau \quad (5.15)$$

is the driven response at epoch $k+1$ due to the presence of the white noise input during interval Δt .

The covariance of \mathbf{w}_k , can be expressed as:

$$\begin{aligned} \mathbf{Q}_k &= E[\mathbf{w}_k \mathbf{w}_k^T] = \\ &E \left[\int_k^{k+1} \int_k^{k+1} \mathbf{T}(k+1,s) \mathbf{G}(s) \mathbf{u}(s) \mathbf{u}^T(t) \mathbf{G}^T(t) \mathbf{T}^T(k+1,t) dt ds \right] = \\ &\int_k^{k+1} \int_k^{k+1} \mathbf{T}(k+1,s) \mathbf{G}(s) E[\mathbf{u}(s) \mathbf{u}^T(t)] \mathbf{G}^T(t) \mathbf{T}^T(k+1,t) dt ds \end{aligned} \quad (5.16)$$

$$E[\mathbf{u}(s) \mathbf{u}^T(t)] = \mathbf{Q} = \text{diag}(\mathbf{q}_d \quad \mathbf{q}_b) \quad (5.17)$$

\mathbf{q}_d and \mathbf{q}_b are 3x3 diagonal matrices, whose elements are the noise PSD of gyros and accelerometers in the IMU, respectively.

Using equation (5.14), the solution of the integral (5.16) can be approximated by

following expansion:

$$\begin{aligned} \mathbf{Q}_k = E[\mathbf{w}_k \mathbf{w}_k^T] &\approx \mathbf{Q}_G \Delta t + (\mathbf{F} \mathbf{Q}_G + \mathbf{Q}_G \mathbf{F}^T) \frac{\Delta t^2}{2} + \\ &+ \left[\mathbf{F}^2 \mathbf{Q}_G + 2\mathbf{F} \mathbf{Q}_G \mathbf{F}^T + \mathbf{Q}_G (\mathbf{F}^T)^2 \right] \frac{\Delta t^3}{6} + \\ &+ \left[\mathbf{F}^3 \mathbf{Q}_G + 3\mathbf{F} \mathbf{Q}_G (\mathbf{F}^T)^2 + 3\mathbf{F}^2 \mathbf{Q}_G \mathbf{F}^T + \mathbf{Q}_G (\mathbf{F}^T)^3 \right] \frac{\Delta t^4}{24} + \dots \end{aligned} \quad (5.18)$$

where

$$\mathbf{Q}_G = \mathbf{G} \mathbf{Q} \mathbf{G}^T \quad (5.19)$$

In the similar way, the measurement model Eq. (5.12) can be converted into the discrete form, i.e.

$$\mathbf{z}_{k+1} = \mathbf{H}_{k+1} \mathbf{x}_{k+1} + \mathbf{v}_{k+1} \quad (5.20)$$

where \mathbf{H}_{k+1} is the design matrix of the system at time t_{k+1} , and \mathbf{v}_{k+1} is the measurement noise at time t_{k+1} , with a covariance matrix \mathbf{R}_{k+1} . Here we use the double-differenced phase and code pseudoranges in the simulations and we assume that the baseline is short so that the atmospheric errors are negligible.

5.1.2 Differenced GPS observation equations

In the case of differential or relative positioning, at least two GPS receivers measure pseudoranges to a set of common satellites simultaneously. Let two receivers on points A and B measure a satellite s and the point A is a known, reference point. The code and phase observation equations can be written as:

$$\begin{aligned} \lambda \phi_A^s(\tilde{t}_A) - \dot{\rho}_A^s(\tilde{t}_A) \tilde{\delta}_A + c \tilde{\delta}^s &= \lambda \phi_A^s(t) + c \tilde{\delta}^s = \rho_A^s(t) + c \tilde{\delta}_A + \lambda N_A^s \\ \lambda \phi_B^s(\tilde{t}_B) - \dot{\rho}_B^s(\tilde{t}_B) \tilde{\delta}_B + c \tilde{\delta}^s &= \lambda \phi_B^s(t) + c \tilde{\delta}^s = \rho_B^s(t) + c \tilde{\delta}_B + \lambda N_B^s \end{aligned} \quad (5.21)$$

$$\begin{aligned} P_A^s(\tilde{t}_A) - \dot{\rho}_A^s(\tilde{t}_A) \tilde{\delta}_A + c \tilde{\delta}^s &= P_A^s(t) + c \tilde{\delta}^s = \rho_A^s(t) + c \tilde{\delta}_A \\ P_B^s(\tilde{t}_B) - \dot{\rho}_B^s(\tilde{t}_B) \tilde{\delta}_B + c \tilde{\delta}^s &= P_B^s(t) + c \tilde{\delta}^s = \rho_B^s(t) + c \tilde{\delta}_B \end{aligned}$$

where $\phi_A^s(\tilde{t}_A)$, $P_A^s(\tilde{t}_A)$ is the phase difference and pseudorange observed at time \tilde{t}_A by receiver A to satellite s ; \tilde{t}_A is the nominal time of the signal reception measured by the

clock of receiver A; \tilde{t}^s is the nominal time of signal transmission measured by the clock of satellite s . We can omit the time t for concision, since all observables are correlated to the same time instance, i.e.:

$$\begin{aligned} P_A^s &= P_A^s(t_A) - \dot{\rho}_A^s(t_A) \delta \tilde{t}_{tA} \\ \lambda \varphi_A^s &= \lambda \varphi_A^s(t_A) - \dot{\rho}_A^s(t_A) \delta \tilde{t}_{tA} \end{aligned} \quad (5.22)$$

To reduce the atmospheric effects that are common to both stations, it is useful to form differenced equations. The single differences can be obtained by the difference equations of the two receivers towards one satellite:

$$\begin{aligned} \lambda \varphi_{AB}^s &= \lambda \varphi_B^s - \lambda \varphi_A^s = \rho_{AB}^s + c \delta \tilde{t}_{AB} + \lambda N_{AB}^s \\ P_{AB}^s &= P_B^s - P_A^s = \rho_{AB}^s + c \delta \tilde{t}_{AB} \end{aligned} \quad (5.23)$$

where

$$\begin{aligned} \rho_{AB}^s &= \rho_B^s - \rho_A^s = \rho_{B0}^s + a_{X,B}^s \Delta x + a_{Y,B}^s \Delta y + a_{Z,B}^s \Delta z - \rho_A^s \\ \delta \tilde{t}_{AB} &= \delta \tilde{t}_B - \delta \tilde{t}_A \\ N_{AB}^s &= N_B^s - N_A^s \end{aligned} \quad (5.24)$$

In the similar way, it is possible to form double differences between two single differences. Let two receivers A, B observing simultaneously two satellites s and t . For each of the satellites a pair of single difference equation similar to (5.23) can be formed:

$$\begin{aligned} \lambda \varphi_{AB}^s &= \rho_{AB}^s + c \delta \tilde{t}_{AB} + \lambda N_{AB}^s \\ \lambda \varphi_{AB}^t &= \rho_{AB}^t + c \delta \tilde{t}_{AB} + \lambda N_{AB}^t \\ P_{AB}^s &= \rho_{AB}^s + c \delta \tilde{t}_{AB} \\ P_{AB}^t &= \rho_{AB}^t + c \delta \tilde{t}_{AB} \end{aligned} \quad (5.25)$$

By subtracting code and phase equations, respectively, we get the following double difference equations:

$$\begin{aligned}\lambda\varphi_{AB}^{st} &= \lambda\varphi_{AB}^t - \lambda\varphi_{AB}^s = \rho_{AB}^{st} + \lambda N_{AB}^{st} \\ P_{AB}^{st} &= P_{AB}^t - P_{AB}^s = \rho_{AB}^{st}\end{aligned}\quad (5.26)$$

where

$$\begin{aligned}\rho_{AB}^{st} &= \rho_{AB}^t - \rho_{AB}^s \\ N_{AB}^{st} &= N_{AB}^t - N_{AB}^s\end{aligned}\quad (5.27)$$

From Eq. (5.26), we can see the receiver clock errors δt^s are eliminated, which is the most important feature of the double differences. Taking into account linearised topocentric distance, the observation equations (5.26) can be written in linear form as:

$$\begin{aligned}\lambda\varphi_{AB}^{st} - \rho_{AB,0}^{st} &= a_{XB}^{st}\Delta X + a_{YB}^{st}\Delta Y + a_{ZB}^{st}\Delta Z + \lambda N_{AB}^{st} \\ P_{AB}^{st} - \rho_{AB,0}^{st} &= a_{XB}^{st}\Delta X + a_{YB}^{st}\Delta Y + a_{ZB}^{st}\Delta Z\end{aligned}\quad (5.28)$$

where the pseudorange is linearised to:

$$\rho_A^s(\tilde{t}_A) = \rho_{A0}^s(\tilde{t}_A) + \frac{\partial \rho_{A0}^s(\tilde{t}_A)}{\partial X_{A0}}\Delta X + \frac{\partial \rho_{A0}^s(\tilde{t}_A)}{\partial Y_{A0}}\Delta Y + \frac{\partial \rho_{A0}^s(\tilde{t}_A)}{\partial Z_{A0}}\Delta Z \quad (5.29)$$

with

$$\begin{aligned}\frac{\partial \rho_{A0}^s(\tilde{t}_A)}{\partial X_{A0}} &= a_X^s = -\frac{X^s - X_{A0}}{\rho_{A0}^s} \\ \frac{\partial \rho_{A0}^s(\tilde{t}_A)}{\partial Y_{A0}} &= a_Y^s = -\frac{Y^s - Y_{A0}}{\rho_{A0}^s} \\ \frac{\partial \rho_{A0}^s(\tilde{t}_A)}{\partial Z_{A0}} &= a_Z^s = -\frac{Z^s - Z_{A0}}{\rho_{A0}^s}\end{aligned}\quad (5.30)$$

5.1.3 System observation equations

In practice, there is always an offset between the GPS antenna and the IMU as shown in Fig. 5.1, which therefore is required to be compensated.

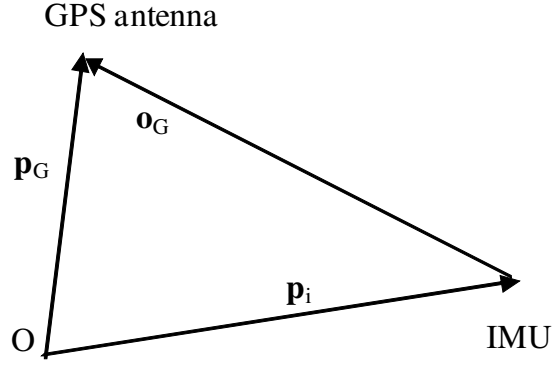


Fig. 5.1 GPS antenna offset.

Let us denote the vector from IMU to the GPS expressed in the body coordinate system as \mathbf{o}_G^b , the positional vector of the antenna as \mathbf{p}_G and the positional vector of the IMU as \mathbf{p}_i . The offset \mathbf{o}_G^b can be computed by:

$$\mathbf{o}_G^b = \mathbf{R}_e^b \mathbf{o}_G^e = \mathbf{R}_e^b (\mathbf{p}_G^e - \mathbf{p}_i^e) \quad (5.31)$$

If vector \mathbf{y} contains coordinates of the GPS antenna determined by GPS, then the GPS observation equation will be as follows:

$$\tilde{\mathbf{y}} = \hat{\mathbf{p}}_G^e - \boldsymbol{\delta} = \hat{\mathbf{p}}_i^e + \hat{\mathbf{R}}_b^e \mathbf{o}_G^b - \boldsymbol{\delta} \quad (5.32)$$

where $\boldsymbol{\delta}$ denotes the residual vector. Taking into account the approximation $\hat{\mathbf{R}}_b^e = (\mathbf{I} - \mathbf{E}^e) \tilde{\mathbf{R}}_b^e$, the observation vector which is used in the Kalman filter is:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{p}}_i^e + \boldsymbol{\delta} \mathbf{p}_i^e + (\mathbf{I} - \mathbf{E}^e) \tilde{\mathbf{R}}_b^e \mathbf{o}_G^b - \boldsymbol{\delta} \quad (5.33)$$

$$\mathbf{L} = \tilde{\mathbf{y}} - \tilde{\mathbf{p}}_i^e - \tilde{\mathbf{R}}_b^e \mathbf{o}_G^b = \boldsymbol{\delta} \mathbf{p}_i^e + (\mathbf{o}_G^e \times) \boldsymbol{\epsilon}^e \quad (5.34)$$

where the misalignment angles in the e-frame $\boldsymbol{\epsilon}^e = [\boldsymbol{\epsilon}_x^e \quad \boldsymbol{\epsilon}_y^e \quad \boldsymbol{\epsilon}_z^e]^T$, the vector product can be transferred to the skew-symmetric operation by Eq. (3.14), i.e.

$$(\mathbf{o}_G^b \times) = \begin{bmatrix} 0 & -o_z^e & o_y^e \\ o_z^e & 0 & -o_x^e \\ -o_y^e & o_x^e & 0 \end{bmatrix}$$

With the observation equation in the Kalman filter, i.e. Eq. (5.34), we can get the design matrix for the loose integration in Eq. (5.13):

$$\mathbf{H}_{\text{Loose}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -\mathbf{o}_z^e & \mathbf{o}_y^e \\ 0 & 1 & 0 & 0 & 0 & 0 & \mathbf{o}_z^e & 0 & -\mathbf{o}_x^e \\ 0 & 0 & 1 & 0 & 0 & 0 & -\mathbf{o}_y^e & \mathbf{o}_x^e & 0 \end{bmatrix} \quad (5.35)$$

To obtain the design matrix for the tight integration, rewrite Eq. (5.29) in matrix form

$$\rho_A^s(\tilde{t}_A) - \rho_{A0}^s(\tilde{t}_A) = \mathbf{A}\mathbf{X} \quad (5.36)$$

where $\mathbf{A} = [a_x^s \ a_y^s \ a_z^s]$ and $\mathbf{X} = [\Delta X \ \Delta Y \ \Delta Z]$.

The left side of Eq. (5.36) is the observation vector for the tight integration as that in Eq. (5.34), i.e. $\mathbf{L} = \rho_A^s(\tilde{t}_A) - \rho_{A0}^s(\tilde{t}_A)$. Multiplying \mathbf{A} matrix at both sides of Eq. (5.34), we get

$$\mathbf{A}\mathbf{L} = \mathbf{A}\delta\mathbf{p}_i^e + \mathbf{A}(\mathbf{o}_G^e \times) \boldsymbol{\varepsilon}^e = \mathbf{A}\mathbf{X} + \mathbf{A}(\mathbf{o}_G^e \times) \boldsymbol{\varepsilon}^e \quad (5.37)$$

From Eq. (5.37), we obtain the design matrix for the tight integration

$$\mathbf{H}_{\text{Tight}} = [\mathbf{A} \ \mathbf{0} \ \mathbf{A}(\mathbf{o}_G^e \times)] \quad (5.38)$$

Please note that each of the 3 sub matrices in Eq. (5.38) is a matrix with 3×3 elements.

5.1.4 Loose integration implementation

In the real positioning experiments, we use the same MEMS-based IMU as introduced in Chapter 2 with the update frequency 100 Hz. On the other hand we use the double-difference observables from a GPS receiver with update frequency 1 Hz, which means that during the interval of two consecutive GPS updates there are 100 INS updates. This provides a convenient approach to interpolate the navigation solutions in high dynamic situations.

The integration algorithm flowchart is shown in Figs. 5.2 and 5.3, including ZVUP (Zero Velocity Update) initialization, INS navigation updates and integrated navigation update. During the ZVUP initialization stage, we know that the IMU is

static in the e-frame, i.e. the velocity and rotation rate should be zero. Therefore the 3D velocities in the horizontal plane and along the vertical direction should be just measurement noise with small standard errors, i.e.

$$\begin{aligned} v_x &\approx 0 \\ v_y &\approx 0 \\ v_z &\approx 0 \end{aligned} \tag{5.35}$$

The static angular velocity estimate due to Earth rotation and misalignment can be given as

$$\hat{\omega}_{eb}^b = \hat{\omega}_{ib}^b - \hat{\mathbf{R}}_e^b \omega_{ie}^e; \quad \hat{\mathbf{R}}_e^b = \tilde{\mathbf{R}}_e^b (\mathbf{I} + \mathbf{E}^e) \tag{5.36}$$

$$\hat{\omega}_{eb}^b = \tilde{\omega}_{ib}^b + \mathbf{d} - \tilde{\mathbf{R}}_e^b (\mathbf{I} + \mathbf{E}^e) \omega_{ie}^e \tag{5.37}$$

With this fact, we can estimate the initial position as well as the attitude and constraint the influences of the IMU sensor noises.

The whole GPS/INS integration can be divided into two main parts: the initial part and the dynamic process part, which are shown in Figs. 5.2 and 5.3, respectively. During the initial stage the vehicle stays still for a period of time to get enough information to determine the initial orientation and sensors noise biases before the moving session. We record all the static data from both IMU and GPS receiver until the vehicle start to move. By means of these static data, ZVUP is carried out and the initial position and orientation are computed as the starting state for the following moving session (see Sections 3.2.3 and 3.2.4).

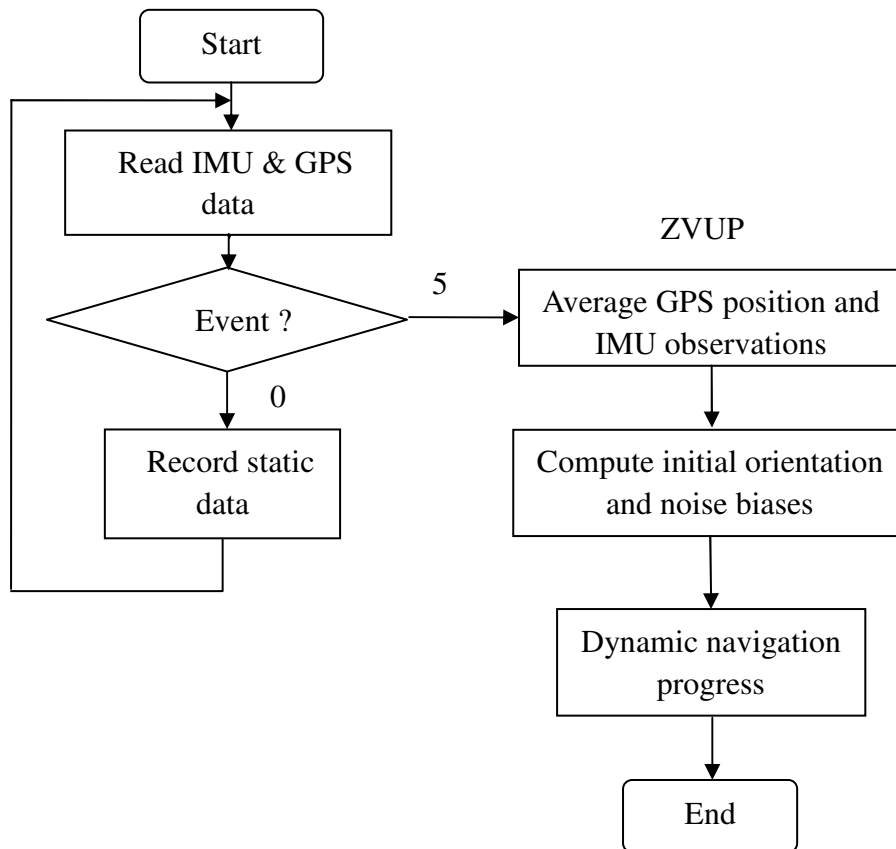


Fig. 5.2 Flowchart of GPS/INS integration initialization algorithm

The events in the flowcharts are defined as (here we only care about these three events related with the LVN):

EVENT 0 – start of static data (initialization)

EVENT 5 – start of dynamic progress data (motion progress of different missions)

EVENT 4 – end of session

During the dynamic progress, each time we read the IMU observations we perform INS navigation update in the e-frame. After each 100 times of INS navigation update, we obtain one GPS difference observation. On arrival of each GPS observation we perform the integrated navigation update with EKF until the end of the session. If there is no GPS observation available, this means a GPS signal gap. In such cases, only INS navigation update is performed as shown in Fig. 5.3.

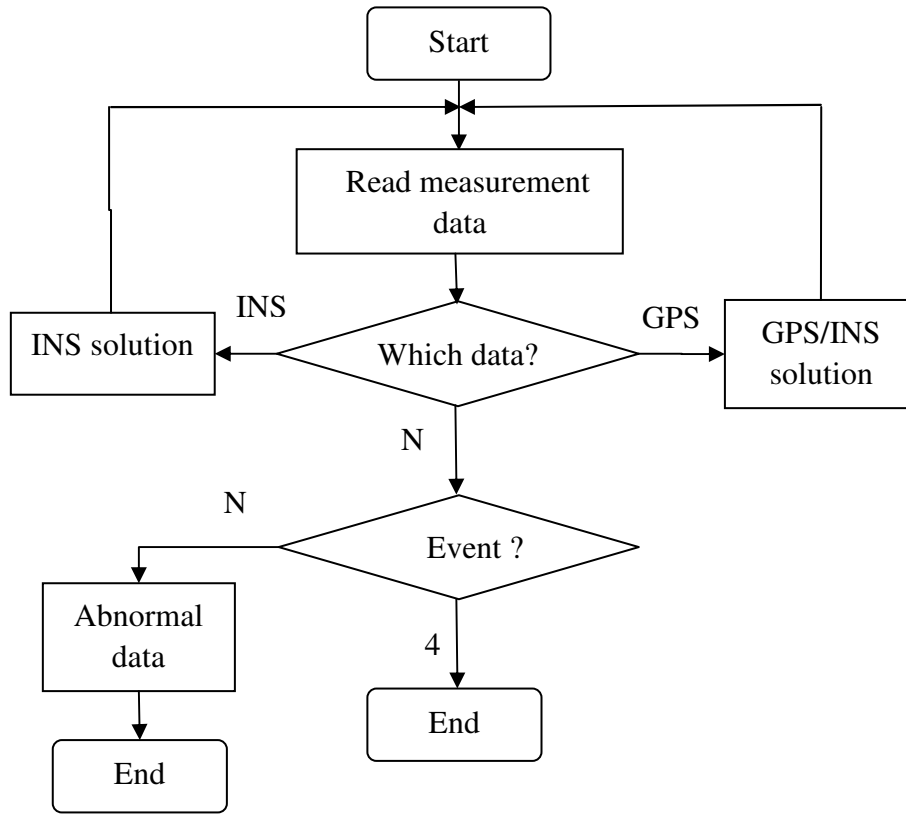


Fig. 5.3 Flowchart of GPS/INS integration dynamic process
(The letter “N” denotes No.)

5.2 Error of tight integration on simulated observations

Using GPS double-difference phase and code pseudorange observations and IMU observations at the raw data level, we simulated and analyzed the standard errors of both position and orientation.

To show the tight integration with realistic conditions, we simulate a flight path as shown in Fig. 5.4. The GPS/INS system is mounted on an aeroplane, which first accelerates and climbs up to 500 m height, then it makes a horizontal loop. After three 90° turns, it starts to land and finally stops close to the starting point. This whole flight takes 680 seconds, in which we simulate a GPS signal outage of 40 seconds to all satellites. Therefore, only INS updates the position and attitude during this interval.

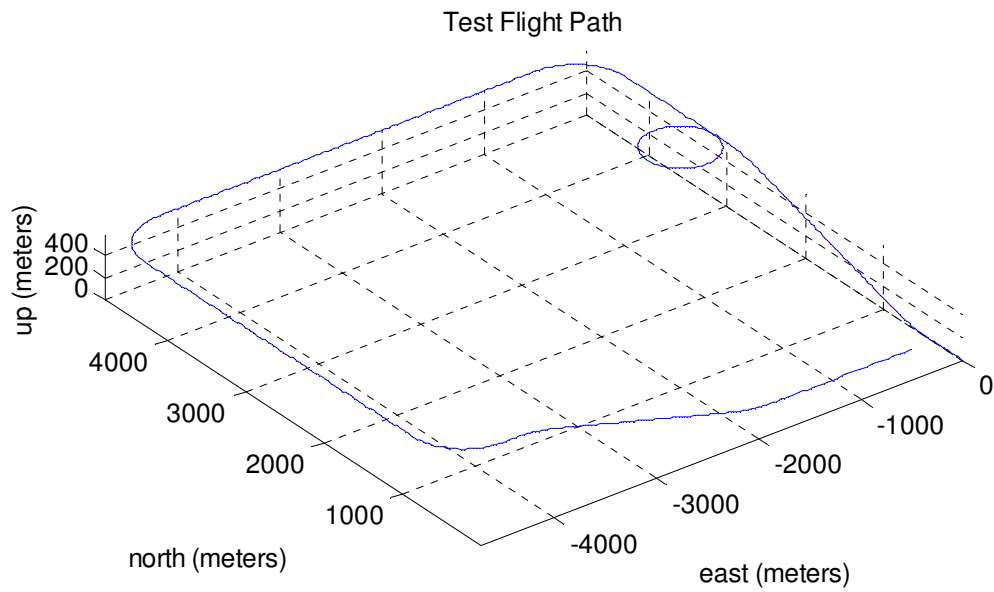


Fig. 5.4 Test flight path

Here we only present the navigation performance in the n-frame during two typical situations, the position and orientation errors of INS with 4 observable satellites and 2 satellites, i.e. INS + 4 satellites and INS + 2 satellites. For other situations and the detailed discussion, we refer to Horemuz and Sjöberg 2001.

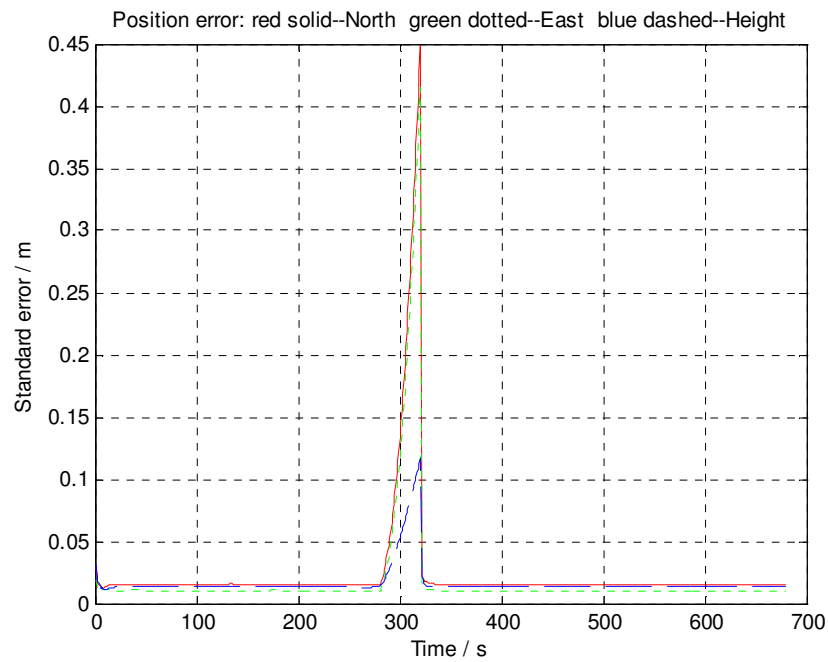


Fig. 5.5 Position errors of INS + 4 satellites with 40s GPS outage

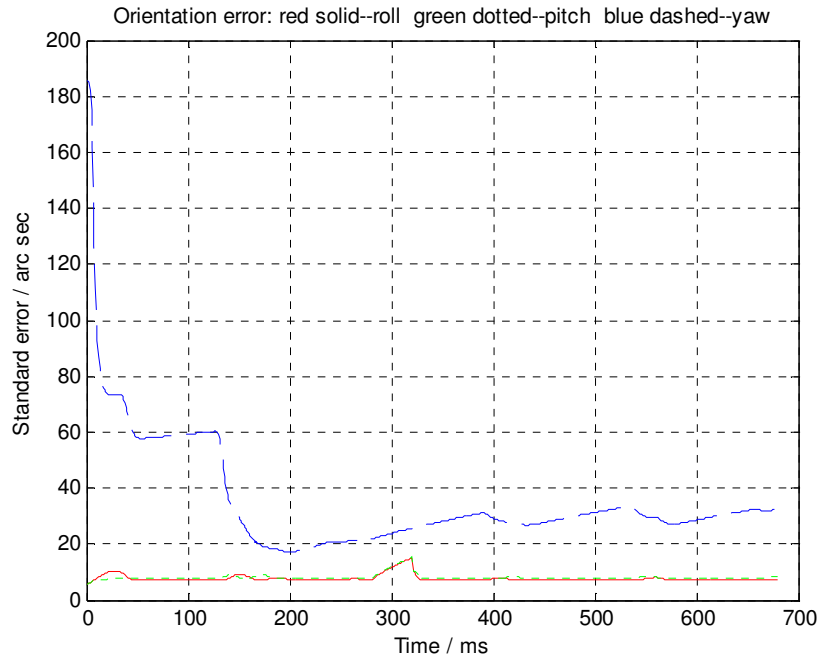


Fig. 5.6 Orientation errors of INS + 4 satellites with 40s GPS outage

From Fig. 5.5 and 5.6, we can see that different maneuvers during the whole progress only slightly affect the position errors, but they are important for yaw determination, since the azimuth becomes observable only in the presence of accelerations other than gravity, i.e. when the vehicle accelerates or turns the direction (Horemuz and Sjöberg, 2001).

Around time 300 seconds there is the period of GPS signal outage, from Fig. 5.4 we can draw the conclusion that the position errors grow quite rapidly during that period. Once the satellites are tracked again later, the position errors drop down to the normal values as before. However, this gap does not make significant influences on the orientation errors at the same time. The orientation errors maintain within a low level all through the whole progress (see Fig. 5.6).

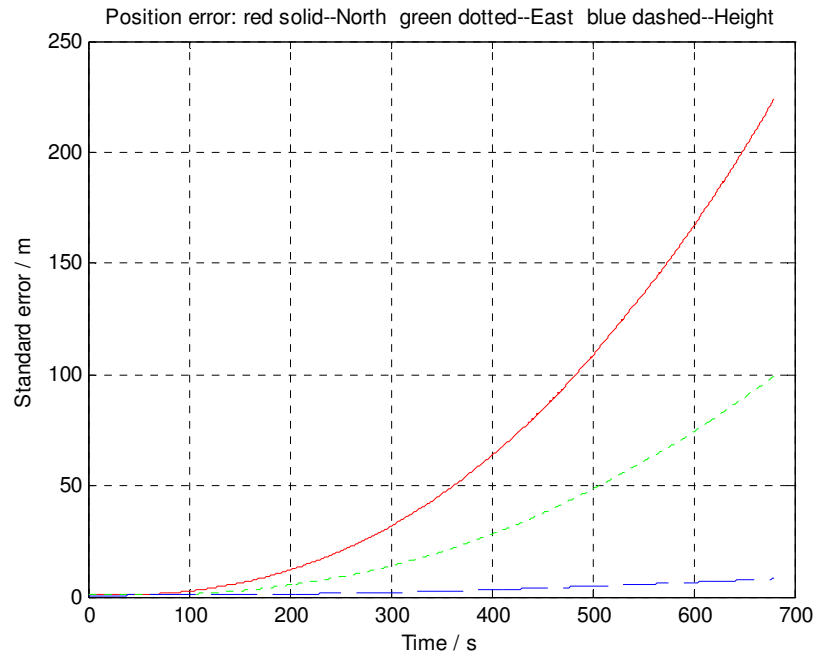


Fig. 5.7 Position errors of INS + 2 satellites with 40s GPS outage

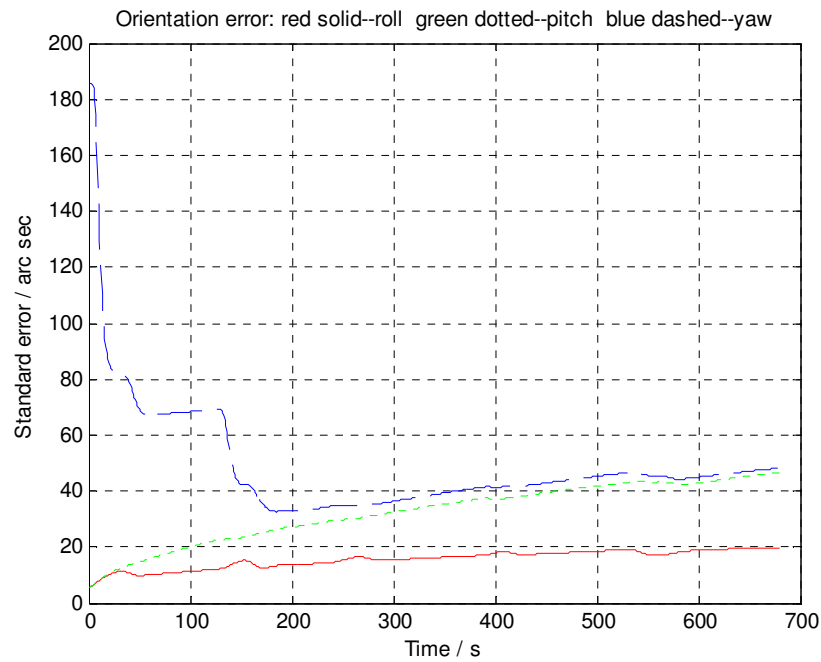


Fig. 5.8 Orientation errors of INS + 2 satellites with 40s GPS outage

As we can see from the Figs. 5.7 and 5.8, there is no remarkable difference within the GPS signal gap, i.e. around 300 second; the position errors grow significantly fast when the number of observed satellites decreases to 2, but the orientation errors can still be kept in a relatively low level. Comparing Figs. 5.8 and 5.6, the differences of orientation errors only increase by 10-30 arc seconds under the situation of 4 and 2

satellites. This is one representation of the benefits with tight integration: even with less than 4 satellites, the GPS information can still contribute to the integration system, especially on keeping the orientation precision.

5.3 Positioning of loose integration on real observations

After the computational simulations, we carried out the following data processing based on real experiments in the scenario of urban area navigation. We get the data set from Visimind AB, which is the real time data for about 10.5 minutes in total with this setup on the vehicle. Among the data, there are several small GPS signal gaps from time to time and there is a large GPS signal gap for about 2 minutes.

The GPS/INS setup in the experiments is shown in Fig. 5.9. The IMU used in the experiment is IMU-FSAS from NovAtel Inc., which is a tactical grade (lowest accuracy grade) IMU.



Fig. 5.9 Setup of GPS antenna and IMU
(Courtesy Visimind AB)

First of all we carry out EKF loose integration in the e-frame with the IMU observations and GPS differenced positioning coordinates. The navigation performances are shown in Figs. 5.10-5.12 (all final results in the following figures are transferred back to the n-frame for better illustrations of their physical meanings).

Based on the same raw data and this loose integration, we analyze the physical constraints (i.e. velocity constraint and height constraint) to the EKF as well as the

influence of simulated gross errors on GPS coordinates in order to test the actual navigation performance of the methods proposed in Chapter 4.

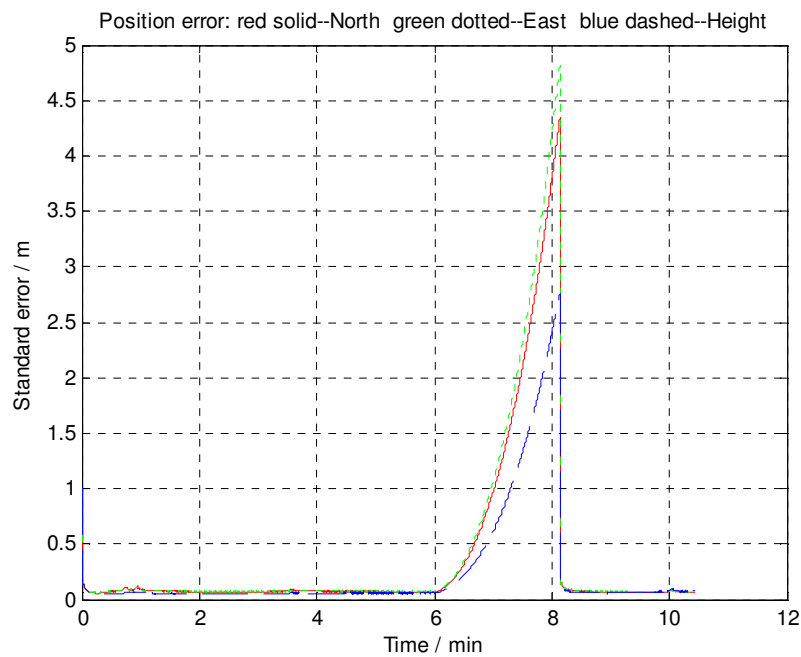


Fig. 5.10 Position errors with EKF
(The main GPS signal gap occurs roughly between 6.0 and 8.1 in the time legend.)

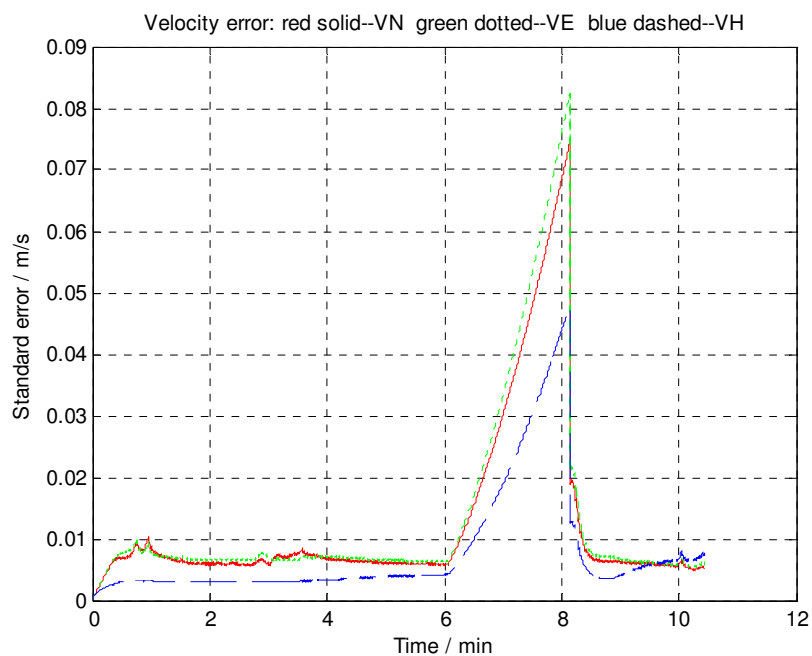


Fig. 5.11 Velocity errors with EKF

Since the IMU update frequency is 100 Hz, the time interval between two continuous data are 10 ms in Figs. 5.10-5.12 and the whole moving session takes time 632.46

seconds. From the beginning, the velocity standard error starts from zero due to the fact of ZVUP as shown in Fig. 5.11. There are several small perturbations of the velocity and position standard errors because of the small GPS signal gaps caused by occasional obstacles from the environment. As we can see from Figs. 5.10 and 5.11, the position and velocity errors grow up significantly when there is a long GPS signal gap. After the gap they drop down to the normal values as the same situation with the tight integration.

From Fig. 5.12, we can see that the orientation errors go down from the initial values with time goes by during the progress, and just grow up about 5-10 arc seconds during the long GPS gap. Even a large GPS gap does not affect the orientation error much in loose integration.

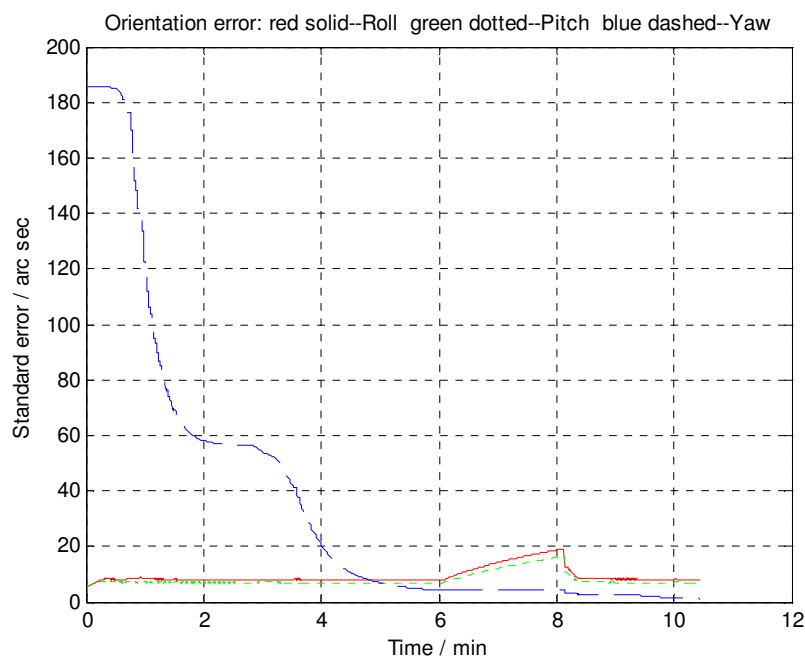


Fig. 5.12 Orientation errors with EKF

From the three figures above, we can see that the error curves are quite smooth from the beginning, since the GPS signal outage periods are quite small at the beginning, except the main one. Then during the large GPS signal outage, the position errors and velocity errors go up significantly, while the orientation errors go up relatively small.

5.4 Analysis on Kalman filter performance and the physical constraints

An important phenomenon with the Kalman filter is the problem of divergence. As is known, the EKF is just the first order approximation of the practical nonlinear system model. Moreover, we only assume that there is a random constant bias and drift in

each of the IMU sensors. From Chapter 2, we can see that this is far from the truth and this assumption is somewhat too simple. Another point is that the IMU sensor noises are changing with time due to its stochastic characters and the change of environment and temperature.

Because of these reasons the EKF navigation may result in divergence if there are frequent and long periods of GPS signal outages. In such cases we could only rely on the simple dynamic model and the inaccurate noise information, the EKF may not adjust the parameters fast and accurate enough. As a result, the errors can be accumulated. If the error grows up to some limits, it may cause numerical problems and lead to the divergence.

For the real time data set introduced in Section 5.3, we only get the recorded data after the experiment without accurate sensor noise parameters. Therefore the EKF loose integration is carried out in the e-frame with empirical parameters. The first minute and 150 seconds position errors are shown in Fig. 5.13 and 5.14, respectively.

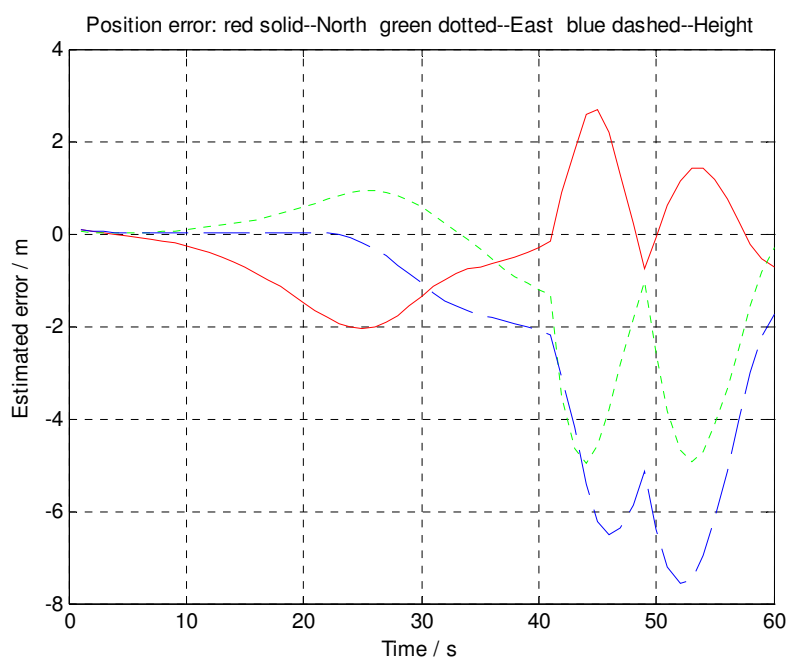


Fig. 5.13 Estimated positioning errors within 1 minute

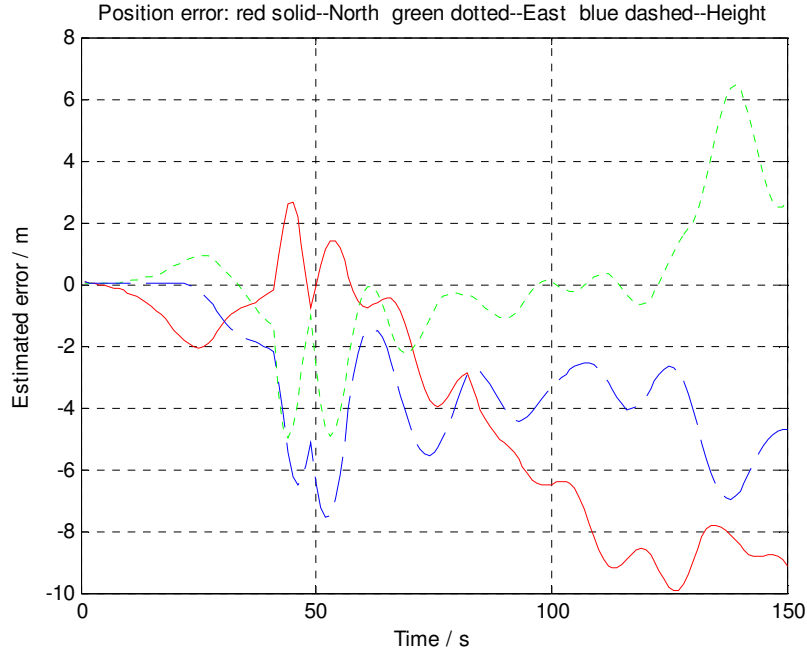


Fig. 5.14 Estimated positioning errors within 150 seconds

The three components of the position error in Fig. 5.13 and 5.14 refer to the three dimensional position errors in Eq. (5.2). Since we do not know the true position of the trajectory and the GPS solution is based on differenced phase observables, the three dimensional position errors, i.e. $\delta \mathbf{r} = [\delta x_i, \delta y_i, \delta z_i]^T$ in Eq. (5.2), can be considered as the estimated position errors of the EKF loose integration. Please note that all position and orientation errors in Fig. 5.10-5.14 are transferred into the n-frame for conciseness and direct viewing with their physical meanings.

From Fig. 5.13 we can see that within one minute the EKF successfully manage to give out reasonable results. There is a gap of 4 seconds followed by 2 seconds with only code observations from the 41 epochs. Then there is another 4 seconds of gap from the 49 epochs. During the gap, the position errors grow up quickly, even reaching up to 4 metres within 4 seconds. We can also notice that the position errors during the following continuous gap accumulate based on the former one, since the two continuous gaps are quite close to each other in time (there is no Kalman filter update during the gaps).

From Fig. 5.14, we can see that due to many GPS gaps the position errors grow up quickly. The position errors along east and down directions can still remain reasonable, but the north direction errors degrade faster than the other two. During the following period of time, the EKF continues to degrade due to some numerical problems. The results are not given here. But the Kalman filter divergence analysis and possible improvement on this key problem can be one of the further research topics.

As a result, during the following tests we omitted the long period of GPS signal outage and only test on the real data with short GPS signal outages, i.e. the first minute of the data. In practice, this does not influence the real navigation significantly, because once there is no GPS signal we can only trust the navigation solution of MEMS based IMU within dozens of seconds and once the GPS signal acquisition is obtained we can trust the positioning result. As for the gap during the navigation route, some backward smoothing algorithm can be used (e.g. RTS smoother; Godha and Cannon, 2005 & 2007).

For the real data, the velocity constraint during the whole moving session can only be applied when the vehicle is driving along the straight direction. In practical tests, the vehicle would not always keep a constant direction, but it changes the direction quite frequently. We did the detailed theoretical analysis in Section 4.2. The physical model itself is quite straight forward in real applications. Since we do not have suitable real data right now, the test of velocity constraint would be combined with the real map information and would be carried out during the following research. We may simulate the GPS and INS observations for the vehicle with map matching, so that the map information can tell us whether or not it goes along a straight line or turn around. We can also use the map as another practical constraint to the real navigation.

The height constraint within the GPS signal gap simply means that we assume the height did not change during the GPS signal gap. It should be noted that the height constraint typically do not provide much improvement during periods of complete outages (e.g. Godha, 2006). It helps especially when there is a partial GPS signal outage as explained in Section 4.2.2, i.e. there are less than 4 observable satellites. It is quite obvious that if the height is fixed the number of observable GPS satellites can be reduced by one. This situation could be helpful with the orientation errors in the tight integration as we can see from Fig. 5.8 (e.g. Horemuz and Sjöberg, 2001).

5.5 Tests on the equivalent weight matrix

In Section 4.4, we proposed to improve the robustness of the Kalman filter by means of the equivalent weight matrix (see Eq. 4.32). To test the performance of this robust Kalman filter, we take the X component of the position errors as an example. First we manually added several random errors with magnitudes from $3-8\sigma$ every 10 seconds to the X component of the position errors, then we compare the differences of the position errors with and without the equivalent weight matrix.

For conciseness, the original X component position error from Fig. (5.13) is re-drawn in Fig. 5.15 (i.e. without outliers). The position errors with random outliers and that with the equivalent weight matrix are shown in Fig. 5.16 (a) and (b), respectively. To compare the differences on position errors due to the equivalent weight matrix, the numerical results before and after equivalent weight matrix on each epoch with random outliers are listed in Table 5.1.

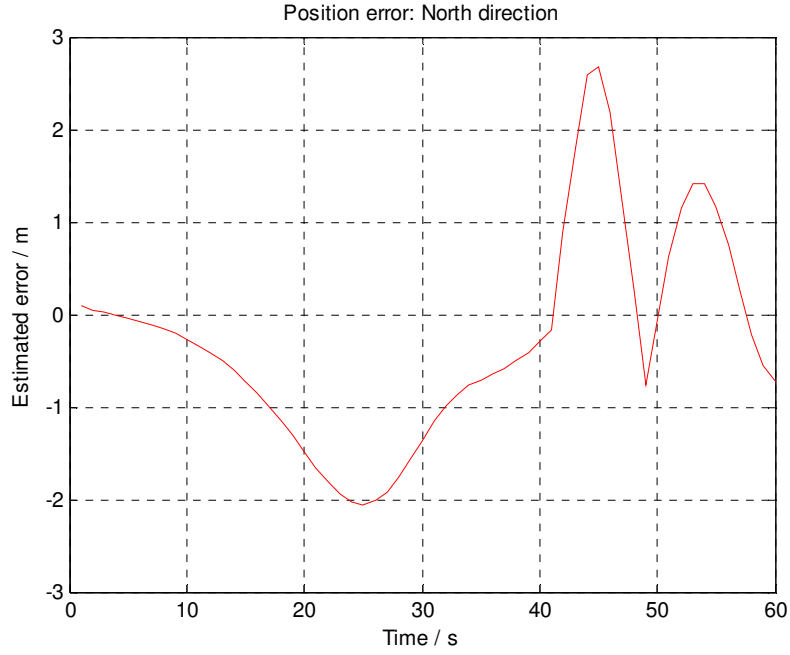


Fig. 5.15 Estimated position errors along north direction

Table 5.1 Proportions of the position errors with random outliers.

δ_0 denotes the original position errors without outliers at each epoch. δ_1 denotes those with random outliers, and δ_2 denotes those with random outliers after equivalent weights.

	10 s	20 s	30 s	40 s	50 s	60s
$(\delta_1 - \delta_0) / \delta_0$	2.2	4.1	1.7	4.8	3.9	6.3
$(\delta_2 - \delta_0) / \delta_0$	0.13	0.01	0.21	0.00	0.01	0.00

In the table, δ_0 is actually the same values at every 10 seconds shown in Fig. 5.15. δ_1 is the values at the same epochs in Fig. 5.16 (a), and δ_2 is that after equivalent weights in Fig. 5.16 (b). Therefore, from the value $(\delta_i - \delta_0) / \delta_0$, $i=1$ and 2 we can evaluate the proportions of position errors caused by the random outliers and the improvement of the equivalent weight matrix.

From Fig. 5.16 (a), we can see that the influence of the outliers on the original position error is remarkable since the outliers get into the system directly, while the curve in Fig. 5.16 (b) is quite smooth, i.e. the position error after equivalent weights can successfully reduce the influence of the added outliers.

From both Table 5.1 and Fig. 5.16, it is quite straight forward to see the differences with and without equivalent weights. For the outliers with large magnitude, the equivalent weights could even be zero (i.e. these outliers are directly picked out) due to the fact that the position errors do not jump to a great extent compared to the previous ones.

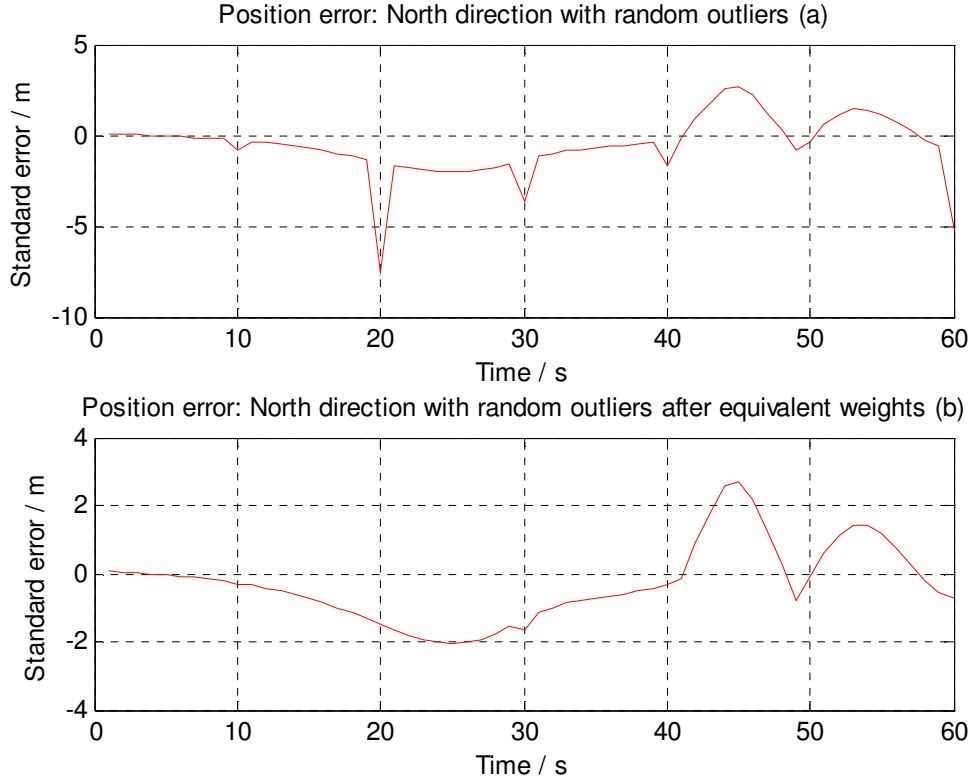


Fig. 5.16 Estimated positioning errors with the equivalent weight matrix

Here we apply the equivalent weights at the final position result level to show its effectiveness for the robust Kalman filter. If we want to apply it at the raw data level, the explicit relation between each sensor observation and the measurement residual should be derived first, so that from the residual variance we can change the weight for each observation at the raw data level.

Of course there are many outlier detection algorithms, which could pick out the discontinuous or continuous outliers in the raw data. These outlier detection algorithms can cooperate with the robust Kalman filter if there are still some outliers that would not be easily picked out and get into the system as one of the normal observations.

Last but not least, an important property of the Kalman filter should be mentioned: the error covariance matrix and the Kalman gain are independent of the actual measurements. They only depend on the noise models. Therefore we know that the outliers only influence the navigation errors, but have no influence on the error covariance matrix. If we know the true positions and compute the posterior error covariance matrix, the outliers also show great influence on it in such cases.

6. Summary, conclusions and recommendations

6.1 Summary and conclusions

The GPS/INS integrated navigation system is no doubt a hot research topic due to the complementary error characteristics of each navigation system. The INS position errors are small in a short term, but they grow without bound over time (see Fig. 2.3). On the other hand, GPS position errors are not so small within a short term, but the accuracy is bounded over time (see Fig. 3.3 and 3.4). However, the GPS signal outage is always a severe problem for GPS positioning (e.g. a 3-minute GPS signal outage can cause positioning error up to 1 km for a vehicle with speed 20 km/h). The normal GPS positioning update frequency is 1 Hz, and therefore it is not suitable for high dynamic applications individually. With INS the integrated system can provide at least 100 updates during each GPS update interval, which is quite suitable for such high dynamic situations.

The performance of the GPS/IMU integrated navigation system is greatly determined by the bridging ability of the stand-alone IMU during GPS signal outage. With better knowledge of the sensor stochastic errors, we can get better estimates of the systematic errors, i.e. the bias and drift of IMU, so that better navigation accuracy and longer bridging time can be reached. To analyze different types of stochastic errors, we tried to build up different stochastic models of the IMU sensors and the practical tests on a MEMS based IMU in Section 2.3 are carried out with these methods, including the AR model, GM process, PSD and AV analysis.

Although different methods of estimating stochastic errors lead to different error models with different coefficients, some stochastic errors of a specific sensor can be verified by comparing and analyzing these methods and their results. The stochastic error models can be further used in the Kalman filter and applied in the GPS/IMU integrated system, which is helpful for bounding the error drift during GPS outage and the faster GPS signal reacquisition.

Different levels of GPS/IMU integration are analyzed and compared in Chapter 3, i.e. loose, tight and ultra tight integration. In this thesis, the simulations are carried out with tight integration and practical tests are carried out with loose integration. Therefore the background of GPS observations and basic IMU navigation principles are introduced.

The Kalman filter is a powerful tool for estimating the errors and data fusion in navigation systems. It can take advantage of different characteristics of both individual systems to provide an integrated navigation solution, which performance is superior to that of either subsystem. A key procedure of the Kalman filter is estimating the drifting parameters of the IMU sensors with the statistical information

of GNSS and IMU. So that the IMU can provide inertial navigation solution with better accuracy during GPS signals outage periods, and the improved position and velocity estimates from the IMU can then be used to reacquire GPS signal much faster when the GPS signal becomes available again.

In Chapter 4, the EKF is briefly introduced for the state estimate with presence of the nonlinearities in practical navigation systems. Within the integrated navigation system, the data fusion algorithm (e.g. EKF in this thesis) is in the vital place for the whole system performance. The EKF can be regarded as the basic implementation of the GPS/IMU navigation system. Aiming at the practical problem of LVN, the physical model of land vehicles can be applied as additional constraints of the positioning errors for the navigation filter (see Fig. 4.5). As a result, the accuracy of the navigation system can be improved.

Under different situations the kinematic models of LVN, i.e. velocity and height constraints as well as the lever arm correction, are proposed. The velocity constraint is suitable when the vehicle moves in constant direction, so the lever arm correction should be applied when turning the direction. When only 3 satellites are observed or the GPS signal outage is relatively long, the height constraint can be applied.

Another practical hazards for the navigation system are sensor outliers. They can greatly influence the position accuracy if it is treated as a normal observation. The robust maximum likelihood estimation with equivalent weight function is applied to adaptively decrease the influence of the outliers.

In Chapter 5, the detailed implementation of the EKF GPS/IMU integrated system in loose structure is proposed. The variance analysis of tight integration is briefly presented. With the GPS receiver and IMU setup on the top of a vehicle, we did the practical data processing based on real time data in an urban area. With these data we first presented the basic EKF navigation performance. Detailed filter performance analysis and possible reasons on the EKF divergence in the long run with this real test are given. The velocity constraint helps especially in urban area when there are frequent GPS signal outages. Under the present experiment limitations, we proposed to apply velocity constraint with real map information as a future research test. For height constraint, we show that it helps to the positioning even when there are only 3 satellite observables. From the tight integration simulations, we can see that even 2 satellites can provide enough information to keep the misalignment errors in a reasonable level when there is partial GPS signal outage.

To test the effectiveness of the equivalent weights proposed in Chapter 4, we added some random outliers at the position error level, and compared the position errors under different magnitude at different epochs. The final results show that it can effectively reduce the influence on the sensor errors and may directly pick out some outliers which are far deviated from the other normal observations. Therefore it can

prevent the outliers from degrading the accuracy and improve the navigation performance.

An important point that must be noted is that the GPS/IMU integrated navigation system is quite dependent on the accuracy of GPS, since we need GPS to correct the IMU sensor errors all the time. However, no GPS evaluation procedure is carried out in this thesis. As we all know, the GPS positioning accuracy is not high at the first epochs but would keep on improving during the beginning stage, and then it would be kept within a bounded level during the following time if not losing the track of the satellites (see Fig. 3.3). The accuracy of GPS can be evaluated in real time situations, including the satellite numbers, PDOP, multipath, etc. In addition, the GPS accuracy would vary from place to place. For example there are totally 11 observable satellites in southern Sweden but only 7 at the northern part.

6.2 Recommendations for further research

This thesis briefly introduces GPS/IMU integrated navigation structures, verifies some stochastic errors in MEMS based IMU and carries out the simulation of tight integration as well as the practical implementation of loose integration. However, both the theoretical and practical research on this topic can be improved in many aspects. Here we outline briefly some of them:

1. The stochastic errors of the IMU sensors, especially the coloured noises. In this thesis we did the tests on the MEMS sensors and got different error models, which can be applied to augment the Kalman filter. However, the coloured noises should be first modeled by differential equations driven by white noise.
2. Tight and ultra tight architecture issues. The tight integration structure is expected to be built based on simulations. Some key issues on tight and ultra tight integration systems can also be the further theoretical research topic, such as IMU aided GPS signal reacquisition. Due to the lab limitations, we do not plan to implement the whole tight system.
3. The Kalman filter. In this thesis we just used EKF as the data fusion algorithm, which is a kind of first-order nonlinear approximation algorithm. EKF improvement against inaccurate system model and noise parameters can be continued as the research topic. There are other filters such as unscented Kalman filter, which can give an approximate accuracy to the second order.
4. Another practical constraint: map match. Now it is quite common that GPS position works with a digital map, so map match technology can be another common practical constraint for LVN. The physical constraint of LVN can be combined with the map information in the future tests.
5. System integrity: gross error and slowly growing error detection. It is better if we can detect these errors and pick them out before they get into the system.

References

Alban S., Akos D. M., Rock, S. M. and Gebre-Egziabher D. (2003). Performance Analysis and Architectures for INS-Aided GPS Tracking Loops. The Institute of Navigation's National Technical Meeting, Anaheim, California, USA.

Allan D. W. (1966). Statistics of atomic frequency standards. *Proceedings of the IEEE*, Vol. 54, No. 2, pp. 221–230.

Babu R. and Wang J. (2004). Improving the quality of IMU-derived Doppler estimates for ultra-tight GPS/INS integration. In *Proceedings of the European Navigation conference*, Rotterdam, pp. 1–8.

Babu R. and Wang J. (2005). Analysis of INS Derived Doppler Effects on Carrier Tracking Loop. *THE JOURNAL OF NAVIGATION* (2005), Vol. 58, pp. 493–507.

Babu R., Wang J. and Rao G. (2008). Algorithms for prediction of INS estimated Doppler in ultra-tight integration. *International Symposium on GPS/GNSS*, Yokohama, Japan, pp. 300-307.

Bendat J. S. and Piersol A. G. (1966). *Measurement and Analysis of Random Data*. New York: John Wiley and Sons.

Bhatti U., Ochieng W. and Feng S. (2007 a). Integrity of an integrated GPS/INS system in the presence of slowly growing errors. Part I: A critical review *Export. GPS Solutions*, Vol. 11, No. 3. pp. 173-181

Bhatti U., Ochieng W. and Feng S. (2007 b). Integrity of an integrated GPS/INS system in the presence of slowly growing errors. Part II: Analysis. *GPS Solutions*, Vol. 11, No. 3. pp. 183-192

Brown A. and Lu Y. (2004). Performance Test Results of an Integrated GPS/MEMS Inertial Navigation Package. *Proceeding of ION GNSS 2004*, Long Beach, California.

Chen W., Chen Y. Q., Ding X.L. (2006). GNSS Applications in Deformation Monitoring, Intelligent Transport Systems, Precise Point Positioning and Atmosphere Study, *GIS De-velopment Asia Pacific*, Vol. 10, Issue 6, pp. 38-43.

El-Diasty M. and Pagiatakis S. (2008). Calibration and stochastic modeling of inertial navigation sensor errors. *Journal of Global Positioning Systems*, Vol. 7, No. 2, pp. 170-182

El-Diasty M. and Pagiatakis S. (2009). A Rigorous Temperature-Dependent

Stochastic Modelling and Testing for MEMS-Based Inertial Sensor Errors. *Sensors*. Vol. 9, No. 11, pp. 8473-8489.

El-Sheimy N. and Niu X. (2007). The Promise of MEMS to the Navigation Community. *InsideGNSS*, March/April 2007: 46-56.

El-sheimy N. (2008). The potential of partial IMUs for land vehicle navigation. *InsideGNSS*, Spring 2008, pp. 16-25.

El-Sheimy N., Hou H. and Niu X. (2008). Analysis and Modeling of Inertial Sensors Using AV, *IEEE Transaction on instrumentation and measurement*, Vol. 57, No.1, pp. 140-149.

Eshel G. (2010). The Yule Walker Equations for the AR Coefficients. http://www.stat.sc.edu/~vesselin/STAT520_YW.pdf. University of South Carolina.

Farrell J.A. and Barth M. (1999). *The Global Positioning System & Inertial Navigation*. McGraw-Hill, New York

Gao G. and Lachapelle G. (2008). A Novel Architecture for Ultra-Tight HSGPS-INS Integration, *Journal of Global Positioning Systems*, Vol. 7, No. 1, pp. 46-61.

Gautier J. (2003). GPS/INS generalized evaluation tool for the design and testing of integrated navigation systems. Ph.D. thesis, Stanford University, USA.

Gebre-Egziabher (2007). What is the difference between 'loose', 'tight', 'ultra-tight' and 'deep' integration strategies for INS and GNSS. *InsideGNSS*, Jan-Feb 2007, pp. 28-33.

Godha S. and Cannon M. E. (2005). Intergration of DGPS with a MEMS-based inertial measurement unit (IMU) for land vehicle navigation application. *Porceeding of ION GPS-05*, Institute of navigation, Long Beach, pp. 333-345.

Godha S. and Cannon M. E. (2007). GPS/MEMS INS integrated system for navigation in urban areas. *GPS Solutions* 2007, Vol. 11, No. 3, pp. 193–203.

Gold K. L. and Brown A. K. (2004). A hybrid integrity solution for precision landing and guidance. *Proceedings of IEEE position, location and navigation symposium*, IEEE, California, pp 165–174

Grewal, M.S., Weill, L.R., and Andrews, A.P (2007). *Global Positioning Systems, Inertial Navigation, and Integration*. 2nd edition. Wiley-Interscience.

Grubbs F. E. (1969). Procedures for detecting outlying observations in samples.

Technometrics 1969, Vol. 11, No. 1, pp. 1–21.

Guerrier S. (2008). Integration of Skew-Redundant MEMS-IMU with GPS for Improved Navigation Performance. M.S. thesis, Dept. Geomatics Engineering, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

Hofmann-Wellenhof B., Lichtenegger H. and Collins J. (2001). GPS, Theory and Practice (Fifth revised edition), Springer Verlag.

Hofmann-Wellenhof B., Legat K. and Wieser M. (2003). Navigation : principles of positioning and guidance. Springer-verlage. Wien, New York.

Horemuz M. and Sjöberg L. E. (2001). Integration of GPS and INS at raw observable level. 3rd Int. Symp. On Mobile Mapping Technology, Cairo, Egypt.

Horemuz M. (2006). Integrated navigation. Lecture notes, Royal Institute of Technology (KTH), Division of Geodesy, Stockholm, Sweden.

Horemuz M. (2007). Algorithm for processing of IMU and camera measurements. Report for project “Dynamic positioning by Integration of Sensor Data”.

Hou H. (2004). Modeling inertial sensors errors using AV, M.S. thesis, MMSS Res. Group, Department of Geomatics Engineering, The University of Calgary, Canada.

IEEE Std 1293-1998. IEEE Standard Specification Format Guide and Test Procedure for Linear, Single-Axis, Non-gyroscopic Accelerometers.

IEEE Std.647-2006. IEEE Standard Specification Format Guide and Test Procedure for Single-Axis Laser Gyros.

IEEE Std.952-1997. IEEE Standard Specification Format Guide and Test Procedure for Single –Axis Interferometric Fiber Optic Gyros.

Jekeli C. (2001). Inertial navigation systems with geodetic applications. Walter de Gruyter GmbH.

Kalman R.E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, Vol. 82, No.1, pp. 35–45.

Kim H., Lee J. G. and Park C. G. (2004), Performance Improvement of GPS/INS Integrated System Using Allan Variance Analysis, The 2004 International Symposium on GNSS/GPS. Sydney, Australia. 6–8 December 2004.

- Ko S. and Bitmead R. (2007). State estimation for linear systems with state equality constraints. *Automatica*, Vol. 43, No.8, pp. 1363-1368.
- Lee Y. C. and O’Laughlin D. G. (1999). A performance analysis of a tightly coupled GPS/inertial system for two integrity monitoring methods. Proceedings of the 12th international technical meeting of the satellite division of the Institute of Navigation, Institute of Navigation, Nashville, pp 1187–1200
- Lee Y. C. and O’Laughlin D. G. (2000). A further analysis of integrity methods for tightly coupled GPS/IRS systems. National technical meetings proceedings “navigating into the new millennium”, Institute of Navigation, Anaheim, pp 166–174
- Leick A. (1995). GPS satellite surveying. Wiley & Sons, New York, Chichester, Toronto.
- Li T. et al. (2010). Ultra-tight Coupled GPS/Vehicle sensor Integration for Land vehicle Navigation. *NAVIGATION*, Vol. 57, No. 4, pp. 248-259.
- Misra P. and Enge P. (2006). Global Positioning System: Signals, Measurements, and Performance. 2nd edition. Ganga-Jamuna Press., Lincoln, Massachusetts, USA.
- Nassar, S. (2005). Accurate INS/DGPS positioning using INS data de-noising and autoregressive (AR) modeling of inertial sensor errors. *Geomatica*, Vol. 59, No. 3, pp. 283-294.
- Neumaier A. and Schneider T. (2001). Estimation of Parameters and Eigenmodes of Multivariate Autoregressive Models. *ACM Transactions on Mathematical Software*, Vol. 27, No. 1, pp. 27- 57.
- Ouladsine M., Shraim H., Fridman L., etc. (2008). Vehicle Parameter Estimation and Stability Enhancement using the Principles of Sliding Mode. Proceedings of 2007 American Control Conference, Newyork, USA.
- Papoulis A. (1991). Probability, Random Variables, and Stochastic Process, Third Edition, McGraw-Hill, Inc.
- Patrich Y. H. and Brown R. G. (2008): A new integrity approach to integrated multi-GNSS systems. *InsideGNSS*, Spring 2008, pp. 24-33.
- Schmidt G. (2010). INS/GPS Technology Trends. NATO RTO Lecture Series. RTO-EN-SET-116, Low-Cost Navigation Sensors and Integration Technology. MIT.
- Schwarz G. (1978). Estimating the dimension of a model. *Ann. Statist.* Vol. 6, No. 2, pp. 461-464.

Shin E. (2001). Accuracy Improvement of Low Cost INS/GPS for Land Application, MSc Thesis, Department of Geomatics Engineering, University of Calgary, Canada, UCGE Report No. 20156.

Simon D. and Simon D.L. (2010). Constrained Kalman filtering via density function truncation for turbofan engine health estimation. *International Journal of Systems Science*. Vol. 41, No. 2, February 2010, pp. 159 – 171

Simon D. and Chia T. L. (2002), Kalman Filtering with State Equality Constraints, *IEEE Trans. IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No. 1, pp. 128-136.

Sinha N. K. and Kuszta B. (1983). *Modeling and Identification of Dynamic Systems*. Springer. New York: Van Nostrand Reinhold.

Sircoulomb V., Israel J., Hoblos G., etc. (2008). State estimation under nonlinear state inequality constraints. A tracking application. 16th Mediterranean Conference on Control and Automation, Ajaccio, France, pp. 1669-1674.

Sjöberg L. E. (2007). *Theory of satellite geodesy*. Lecture notes, Royal Institute of Technology (KTH), Division of Geodesy, Stockholm, Sweden.

Sun D. (2010). Ultra-Tight GPS/Reduced IMU for Land Vehicle Navigation. PhD Thesis, published as Report No. 20305, Department of Geomatics Engineering, The University of Calgary, Canada.

Tehrani M. M. (1983). Ring laser gyro data analysis with cluster sampling technique, *Proceedings of SPIE*, Vol. 412, pp. 207-220.

Tiberius C. C. J. M. (1998). *Recursive data processing for kinematic GPS surveying*. Publications on Geodesy. New series: ISSN 0165 1706, No. 45. Netherlands Geodetic Commission, Delft.

Tsui J. B. (2004). *Fundamentals of Global Positioning System Receivers: A Software Approach*, 2nd Edition. Wiley-Interscience.

Wang L., Chiang Y., and Chang F. (2002). Filtering method for nonlinear systems with constraints, *IEE Proceedings - Control Theory and Applications*, 2002, Vol. 149, No. 6, pp. 525-531.

Welch G. and Bishop G. (2001). *An Introduction to the Kalman Filter*. SIGGRAPH 2001, Course 8, August 12-17.

Wen W. and Durrant-Whyte H. (1992). Model-based multi-sensor data fusion. IEEE International Conference on Robotics and Automation, Nice, France, pp. 1720 -1726.

Xu P. (1993). Consequences of constant parameters and confidence intervals of robust estimation. Boll Geod Sci Affini, Vol.52, pp. 231 - 249.

Yang Y., Song L. and Xu T. (2002). Robust estimator for correlated observations based on bifactor equivalent weights. Journal of Geodesy. Vol. 76, pp. 353-358.

Appendix A: Derivation of constrained Kalman filter

Start from the Lagrangian optimization condition for the Kalman filter :

$$\Omega_k = \sum_{i=1}^{n_k} \mathbf{P}_{k_i} \rho(V_{k_i}) + (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k)^T \boldsymbol{\Sigma}_{\bar{\mathbf{X}}_k}^{-1} (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k) + 2\boldsymbol{\beta}_k^T (\mathbf{B}_k \hat{\mathbf{X}}_k - \mathbf{D}_k) = \min \quad (\text{A.1})$$

Let the derivative of function $\rho(\cdot)$ be denoted as $\psi(\cdot)$, and let the equivalent weight be expressed as

$$\bar{P}_i = P_i \frac{\psi(\bar{V}_{k_i})}{\bar{V}_i} \quad (\text{A.2})$$

Take the partial derivative of Ω_k with respect to $\hat{\mathbf{X}}_k$, and equate to zero. Then we obtain

$$\begin{aligned} \frac{\partial \Omega_k}{\partial \hat{\mathbf{X}}_k} &= \sum_{i=1}^{n_k} \mathbf{P}_{k_i} \psi(V_{k_i}) \mathbf{A}_k + 2(\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k)^T \boldsymbol{\Sigma}_{\bar{\mathbf{X}}_k}^{-1} + 2\boldsymbol{\beta}_k^T \mathbf{B}_k \\ &= 2\mathbf{V}_k^T \bar{\mathbf{P}}_k \mathbf{A}_k + 2(\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k)^T \boldsymbol{\Sigma}_{\bar{\mathbf{X}}_k}^{-1} + 2\boldsymbol{\beta}_k^T \mathbf{B}_k \end{aligned} \quad (\text{A.3})$$

where the derivative of quadratic function

$$\frac{\partial \mathbf{X}^T \mathbf{A} \mathbf{X}}{\partial \mathbf{X}} = \mathbf{A} \mathbf{X} + \mathbf{A}^T \mathbf{X} \quad (\text{A.4})$$

is applied.

Considering the residual

$$\mathbf{V}_k = \mathbf{A}_k \hat{\mathbf{X}}_k - \mathbf{Z}_k \quad (\text{A.5})$$

we obtain

$$(\mathbf{A}_k \hat{\mathbf{X}}_k - \mathbf{Z}_k)^T \bar{\mathbf{P}}_k \mathbf{A}_k + (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k)^T \boldsymbol{\Sigma}_{\bar{\mathbf{X}}_k}^{-1} + \boldsymbol{\beta}_k^T \mathbf{B}_k = 0 \quad (\text{A.6})$$

Transposing Eq. (A.6), we get

$$\mathbf{A}_k^T \bar{\mathbf{P}}_k (\mathbf{A}_k \hat{\mathbf{X}}_k - \mathbf{Z}_k) + \mathbf{P}_{\bar{\mathbf{X}}_k} (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k) + \mathbf{B}_k^T \boldsymbol{\beta}_k = 0 \quad (\text{A.7})$$

From Eq. (A.7), we obtain the explicit solution of state estimate

$$\hat{\mathbf{X}}_k = (\mathbf{A}_k^T \bar{\mathbf{P}}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\mathbf{A}_k^T \mathbf{P}_k \mathbf{Z}_k + \mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k) - (\mathbf{A}_k^T \bar{\mathbf{P}}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} \mathbf{B}_k^T \boldsymbol{\beta}_k \quad (\text{A.8})$$

Eq. (A.8) is the same with Eq. (4.21) in Chapter 4.

Q.E.D.