Precise Velocity and Acceleration Determination Using a Standalone GPS Receiver in Real Time

Jianjun ZHANG

RMIT

School of Mathematical and Geospatial Sciences
Royal Melbourne Institute of Technology
Australia

Declaration by the Candidate

I hereby declare that the work presented in this thesis is, to the best of my knowledge, original except where acknowledged in the context. No part of this document has been submitted for a degree of any kind, at this or any other academic institution. Furthermore, the work presented has been carried out since the official commencement date of the program.

Jianjun Zhang

3/28/2007

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisors Prof. Dr. Kefei Zhang and Dr. Ron Grenfell for their guidance, encouragement and support. A great deal of thanks should also go to Rod Deakin, David Silcock, and Lucas Holden. It is a wonderful experience working as a team member in the Measurement Science Group of RMIT.

Many thanks go to the staff and postgraduate students in the School of Mathematical and Geospatial Sciences, RMIT University. Particularly I would like to thank Denis Ma, Afshin Alizadeh Shabani, and Cecilia Xia for the time spent together with discussions and friendships.

I would also like to thank my family and friends for their continuous support throughout this research. In particular I would like to thank my wife, Ting Yang for her love, sacrifice and patience.

Last but not the least; I thank my sponsor, CRC for microTechnology, for providing me scholarship for the first three years of this research.

TABLE OF CONTENTS

CHAPTER SIX

CHAPTER SEVEN

Chapter One

INTRODUCTION

This chapter states the research background and clarifies the research objectives. It also gives an outline of the thesis and highlights the contributions of the research.

1.1 Research Background: Navigation and GPS

The word "navigation" (Latin navis, "boat"; agire, "guide") has a traditional meaning of being the art and science of manoeuvring safely and efficiently from one point to another [Rousmaniere, 2003], usually in real time. One of the scientific definitions of navigation is to accurately determine position and velocity relative to a known reference [Farrell and Barth, 1999]. The long evolution of navigation techniques has resulted in a state-of-the-art system named NAVigation Satellite Time And Ranging (NAVSTAR) Global Positioning System (GPS).

The NAVSTAR GPS is an all-weather, space-based radionavigation system controlled and operated by the United States Department of Defense [DoD]. With the provision of seven dimensional information (three dimensions of position, three dimensions of velocity and one dimension of precise time) in the World Geodetic System 1984 (WGS-84) [NIMA, 2000], GPS is a total navigation solution of positioning, velocity determination and time transfer (PVT) for users anywhere on or near the Earth's surface.

One of the main problems of GPS, however, is that under the current category of GPS standard positioning service (SPS), a civilian user can only achieve an accuracy level of ±10~15m 3-Dimension (3D) Root Mean Square (RMS) [Stenbit, 2001]. This is because in a typical GPS observation system, the broadcast satellite ephemeris and range measurements from a GPS receiver manifest errors and biases that contaminate the position solution.

The errors and biases, for example, the satellite signal propagation errors in the atmosphere, are characterised as having spatial and temporal correlations. Two GPS receivers not far away from each other will experience almost the same errors and biases from GPS when observing simultaneously. Differential GPS (DGPS) therefore becomes the most prevalent technique to take advantage of the correlation characteristics where the common errors and biases can be effectively cancelled out.

Essentially, DGPS is a kind of relative positioning system where at least two sets of GPS receivers are required to operate simultaneously. With one (or more) receiver(s) functioning as reference station(s) with known coordinates, the relative position of a roving receiver can be determined if the differential corrections can be generated and applied, either in post-processing or in real time.

The real-time code range differential GPS system is known as RTD (Real Time Differential) and has an accuracy range of ±0.5m~2m. The real-time carrier phase differential GPS system is often called RTK (Real Time Kinematic) with centimetre level accuracy in real time. The distinction between RTD and RTK lies in the different types of observables used and the positional accuracy achievable. In both techniques, data links are required for transmission and reception of the differential corrections.

For two decades, research and development of various DGPS systems has been undertaken in the GPS community. Since error correlation degrades with the distance between a roving receiver and a reference station, and degrades rapidly with the

elapse of time, in order to ensure the rover obtains a high level of accuracy, dense spacing of reference stations with expensive radio equipment are needed. A huge amount of money has been invested on the research and development (R&D) in this area, and on the construction of such DGPS infrastructure, which varies from the Local Area DGPS (LADGPS) and Wide Area Differential GPS (WADGPS), to the expensive and sophisticated Wide Area Augmentation System (WAAS) and Space Based Augmentation Systems (SBAS) in general.

In short, differential GPS techniques have drawn much of the attention and led to most of the developments in the area of high accuracy GPS positioning. One of the problems with DGPS is that it requires at least two sets of GPS receivers working simultaneously. Another drawback is that it relies on inter-receiver data links. Thirdly, the accuracy of the rover depends on its distance(s) from the reference station(s) and thus DGPS is spatially limited within a local area. Finally, much useful information contained in one-way measurements, such as the total electron content along the signal profile, the water vapour information in the troposphere, is eliminated or discarded in the differential process.

Not as popular as the differential GPS study, though, pioneer research on the un-differenced carrier phase data processing has been carried out in Australia at the University of New South Wales in early 1990s, showing the mathematical equivalence to double-differencing, see Grant [1990] and Grant et al. [1990]. There has been an increased interest as more and more people have started their research and analyses on Precise Point Positioning (PPP) as a result of the significant improvement of the precise ephemeris service from the International GNSS Services (IGS) [Beutler et al., 1994]. This is achieved through hundreds of globally distributed tracking stations, and IGS' commitment to provide a precise ephemeris service in near real time [Springer and Hugentobler, 2001],

PPP is referred to as single GPS receiver positioning, but with centimetre to decimetre level of positional accuracy. The concept of PPP can be dated back to the early 1970s by P. P. Anderle, who first named the method as “precise point positioning” [Kouba and Heroux, 2001]. To achieve a higher accuracy, PPP uses the IGS precise orbital data, accurate satellite clock data, a dual-frequency geodetic receiver, and observes in-view GPS satellites over a long period in static mode. In PPP, since there is no other receiver available for generating and applying differential corrections, all the error sources in GPS positioning need to be carefully modelled and accounted for, and the one-way carrier phase integer ambiguities are estimated together with the user position unknowns. The observation time required depends on the time that the system needs to converge to a position fix with the desired accuracy.

As PPP relies heavily on the IGS precise ephemeris whose full accuracy is only available in post-processing mode, applications of the PPP method in real time and especially in kinematic mode at centimetres accuracy are still an area of challenge.

Time-relative positioning is among many approaches attempting to improve the positional accuracy in standalone mode. The time-relative positioning method was first presented by Ulmer et al. [1995], and evaluated by Michaud and Santerre [2001] after the removal of Selective Availability. Recently, Balard et al. [2006] proposed to use the loop misclosure corrections to improve the accuracy of time-relative positioning in quasi real time. This method suffers from the following problems. Firstly, it requires a user to occupy a surveyed position at least twice. Secondly, the temporal and spatial error correlations degrade with the elapse of time. Finally and most fatally, it is just an approximation since there is no one-to-one mapping from the accumulated carrier phase and the satellite baseline to the change of receiver positions between the two measurement epochs.

Methods to model the change of receiver position with the carrier phase differences between two consecutive epochs to attempt higher positional accuracy face the same

problem as with the time-relative positioning. The method proposed by Ford and Hamilton [2003] is typical of this genre. The relative position with respect to a starting point through accumulated positional changes corrupts over a short period.

In navigation, state is a terminology used to describe the motion of a system, which includes the position and velocity vectors of the system, and the acceleration and jerk if available or needed. The state is expressed using a differential equation if the system is continuous or a difference equation if the system is discrete. The state equation describes how a system state changes from epoch to epoch and what kind of dynamics that the system is experiencing. The more accurate the state equation is, and the more accurate measurements can be made, the better accuracy with which the navigation solution can be achieved.

Inertial Navigation System (INS) solves the navigation problem by integration. An INS usually consists of gyroscopes to sense the angular rates of a system to obtain the system orientations, and 3-axis accelerometers to sense the linear inertial accelerations. Kinematic accelerations of the system are obtained by subtracting the gravitational accelerations from the inertial accelerations. A single integration of kinematic accelerations over time gives the system velocity and a double integration over time yields the system position.

INS is an autonomous navigation system because it is self-contained and not dependent on external sources. Since an integration system is characterised by its unbounded error propagation, particularly in the case of INS progressing, from accelerations to a position is a double integration process where the error accumulation is much quicker than that in just a single integration. As such, INS only has good short-term stability. The three main types of errors in the acceleration measurements in INS, i.e. the scale factor error, drift, and random noise, contribute together to corrupt the state estimates over a period. Thus, periodic system calibrations are required.

In contrast to INS, GPS is recognised as having a superb consistent accuracy in PVT (Position, Velocity and Time) determination. Therefore, an integration of INS and GPS can outperform either single system by combining advantages of the consistency of GPS accuracy and the short-term stability of INS.

With its 7-dimension information, GPS is somewhat similar to INS when taken as an integration system. The differences are that the orientations of a GPS receiver are mathematically established in the WGS-84 system, and it is just a single integration from velocities to positions. To date, with SA switched off and with the improvement of receiver technologies, velocities of a GPS user can be determined at an accuracy level of centimetres per second, which are several orders of magnitude higher than that of the positions. Moreover, the velocities can be measured at a relative high sampling rate, say from 10Hz to 100Hz. The velocity, if resolved from the Doppler shift method, is an instant velocity at the measurement epoch. This gives the opportunity that a GPS receiver can be used like an INS to obtain precise positions in real time, with accuracy higher than that from the SPS, by integration of the precise velocities with high sampling rate over time.

Traditionally in navigation the positional accuracy has been of primary interest, however, precise velocity determination using GPS potentially is equally important as positioning, since velocities with better accuracy can be utilised for a better positional result or a longer effective integration time.

There are also many applications where the accuracy of velocity and acceleration is equally important as positioning, and is desired in standalone mode and in real time. A good example is in the sport of rowing where precise velocity and acceleration in conjunction with the position are sought after at a high output rate so that the in stroke velocities and accelerations of an elite rower can be measured and analysed. This information would in turn facilitate rowing performance improvement [Zhang et al.,

2003b]. In fact, due to the competitive nature of rowing, high accuracy of velocity and acceleration determination is preferable to accurate positioning.

Nowadays DGPS techniques are readily available for such an application, however, they require additional radio devices and reference stations. This makes the system expensive, heavy, and larger than otherwise necessary. A small real-time GPS system that can be used in standalone mode to achieve high positional accuracy is desirable for many applications.

In summary, this research seeks to explore autonomous, real-time, high accuracy velocity/acceleration determination using GPS in standalone mode, to facilitate the improvement of positional accuracy from SPS.

1.2 The Problem of Velocity and Acceleration Determination using GPS

The simplest way to get the velocity and acceleration from a GPS receiver is to differentiate the GPS determined positions with respect to time. Velocity is the first time derivative of positions and acceleration is the second time derivative. The trouble is that errors in positions may be amplified through differentiation, and this becomes worse when a high output rate is used, since the positional error remains the same but the time interval is decreased.

For precise velocity determination, GPS satellites are moving signal sources from which a GPS receiver senses the Doppler shifts that are generated due to the relative motion between the observed satellites and the receiver. The ground velocity of the receiver can then be determined if Doppler shifts can be measured from four or more GPS satellites.

GPS is designed for and operated by the US military as a radionavigation PVT system. Although the system has capabilities to provide precise real-time velocities, due to the deployment of SA prior to May 2000 and limitations of the receiver hardware, most methods of velocity determination were based on differential

techniques to account for the errors induced in satellite orbits and onboard GPS clocks.

Research groups at the Department of Geomatics Engineering, University of Calgary, Canada worked actively in the 1990s on precise velocity determination. Fenton and Townsend [1994] derived phase velocity, and phase acceleration from the carrier phase measurements of NovAtel GPS Card^{TM} receivers, and ±0.28mm/s was achieved in a 7km static baseline. Hebert et al. [1997] and Cannon et al. [1998] conducted in-depth analyses on GPS velocity determination using GPS signal simulators. Velocities with up to ±2mm/s 3D RMS accuracy were achieved under low dynamics (acceleration<0.5m/s^2). Szarmes et al. [1997] presented results of a series of differential GPS tests conducted for the purpose of high accuracy aircraft velocity determination. Compared with the velocities obtained through receiver generated raw Doppler measurements and the carrier phase derived 'precise' Doppler, it was concluded that under constant velocity or low acceleration conditions, the accuracy of the raw Doppler derived velocity estimates is at least as good as the velocity accuracy from the first-order central difference approximation of the carrier phase.

The removal of SA allows a significant velocity accuracy improvement. With SA off, ±0.2m/sec per axis (95%) accuracy is guaranteed by the GPS system [DoD, 1996]. A static velocity accuracy of ±3~5cm/s 3D RMS was first reported in Misra and Enge [2001] in standalone mode. Zhang et al. [2003a] conducted a comparison of real-time velocities obtained from an inexpensive code-only GPS receiver with the velocities from a Trimble 5700 RTK system, and reported that ±3cm/s accuracy was achieved using the code-only GPS receiver in both static and dynamic mode. Zhang et al. [2004] demonstrated that the same accuracy level had been achieved using an inexpensive 1 Hz GPS receiver as with a 10Hz sampling rate in standalone mode.

Van Graas and Soloview [2003] showed that sub-centimetre per second velocity accuracy is achievable whether in static or dynamic, stand-alone or relative mode. It

was concluded that what really matters is the receiver quality. Serrano et al. [2004] reported that by employing the first-order central difference approximation of the carrier phase measurements, better than ±1cm/s (2-sigma) accuracy under high-multipath conditions can be achieved from a low-cost GPS receiver. Based upon post processing of the kinematic data, they reported that accuracy of the phase measurements was degraded in a moving environment and there were biases in both the static and kinematic results.

Perhaps the best velocity accuracy published so far in standalone mode is from Septentrio's PolaRx2® receiver; ±1.5mm/s horizontal and ±2.8mm/s vertical precision are stated in the product specifications [PolaRx2, 2004]. This dual frequency GPS receiver achieves such a high accuracy through the introduction of a special scheme which accounts for the change-rate of tropospheric delay [Simsky and Boon, 2003], which is previously taken as negligible.

The author has continued and extended the investigation on precise velocity determination, primarily aimed at high output rate, real-time and autonomous applications. A series of papers have been published covering topics on real-time GPS satellite velocity/acceleration determination algorithms using the broadcast ephemeris, error budgets of the precise velocity determination using Doppler shift measurements, and the most accurate Doppler shift observation equation with the relativistic effects corrected [Zhang et al., 2005a; 2005b; 2006a; 2006b]. A thorough and extensive investigation on precise velocity determination using GPS on which these publications were based is presented in this thesis.

Much of the investigation on precise acceleration determination using GPS has been carried out in the context of airborne gravimetry. Many issues concerning the position differentiation method were examined by Van Dierendonck et al. [1994] and Bruton [2000]. Acceleration determination using GPS Doppler rates dates back to the early 1990s initially by Kleusberg et al. [1990], followed by Jekeli [1994] and Jekeli and

Carcia [1997]. Since Doppler rate is not a direct observable in GPS, it must be derived from either the Doppler shifts or the carrier phase measurements. A recent report on precise acceleration determination in this context is from Kennedy [2003], who used differentiators of the 5th order central difference of a Taylor series to derive Doppler rates from the carrier phase measurements. Essentially Kennedy's method is a differential GPS approach mainly for post processing applications due to concern that the accuracy of GPS satellite orbital accelerations from the broadcast ephemeris might not be accurate enough for real-time applications.

Other investigations on acceleration determination are in the context of attitude determination using GPS. Kornfeld et al. [1998] proposed that accelerations derived from a single GPS receiver can be used to determine a "2-axis pseudo attitude" for aircraft. Psiaki et al. [1999; 2000] extended Kornfeld's method to a 3-axis absolute attitude reference system. They also examined the accuracy of the GPS-derived acceleration vectors by position differentiation. The acceleration derivation using GPS also appeared in the work by Ellum and Sheimy [2002], which is also based on position differentiation.

This thesis investigates the real-time, high output rate acceleration determination using GPS in standalone mode. Satellite acceleration determination algorithms, differentiator designs to derive Doppler rate "observables", and "abnormal" controls are investigated. An alternative method is also proposed in the measurement domain by extending the state equation to accommodate acceleration as an unknown to be determined, which is different from the position differentiation method and the Doppler rate approach.

1.3 Research Scope, Objectives and Methodologies

As a navigation problem, a position can be obtained from integration of velocities over time, as mentioned earlier. The higher the sampling rate of the velocity

measurements and the higher the precision of velocity determination can be, the higher the positional accuracy is achievable. Thus this research is restricted to real-time applications where precise velocity and acceleration are required in standalone mode and in high dynamics; it is further assumed that the observations can be made at a high sampling rate. The ultimate objective of this research is to establish the theoretical basis for such applications.

In GPS, the Doppler frequency shifts directly relate to the relative velocities between a GPS receiver and in-view satellites. Since the carrier phase measurement comes from the integration of measured Doppler frequency shifts over time, and the Doppler shift is the first derivative of the carrier phase measurements with respect to time, the inter-relationship is exploited to analyse the error characteristics in both the velocity and acceleration determination using GPS. This is to mitigate/eliminate the errors in the Doppler shift measurements.

The change-rate of the Doppler shifts, or the carrier phase acceleration, relates to the kinematic acceleration. However, there is no such observable in GPS, and thus appropriate differentiators to derive the change-rate of the Doppler shifts are investigated. Such differentiators are designed for real-time and dynamic applications.

Previous Doppler effect in GPS only accounts the correction from the special relativity. As the GPS observation system is in the Earth's gravitational field, the general relativity theory is to be used to extend the theory of GPS relativistic Doppler effects.

Therefore, the objectives of this research can be listed as follows:

- To study the problem of real-time precise velocity and acceleration determination in standalone mode
- To identify all the error sources in GPS Doppler shift measurements
- To develop models to mitigate/eliminate these errors

- To design appropriate differentiators to derive Doppler shifts or their change-rates for real-time and high sampling-rate applications
- To investigate the operational issues of velocity and acceleration determination in the case of measurements are made at high sampling rates

1.4 Contributions of This Research

The primary outcomes of this research are:

- A new algorithm to compute satellite Earth-Centred-Earth-Fixed (ECEF) velocity using the broadcast ephemeris is derived. Polynomial interpolations are also proposed to improve the speed of the ECEF satellite velocity calculations, which is important for those applications where the receiver velocities are estimated at a high sampling rate;
- A closed-form formula to calculate satellite ECEF acceleration is developed using the broadcast ephemeris, which has an accuracy level equivalent to that from the precise ephemeris. This assures accurate accelerations of a GPS receiver can be determined in real time;
- A comprehensive error analysis for the range measurement has been conducted for PPP. With this knowledge, error corrections for the Doppler shift measurements and the errors that affect the change-rate of the Doppler frequencies have been elucidated. Several new formulae are derived to account for these errors, which benefit the accuracy improvement of ground velocity and acceleration determination;
- An intensive investigation of the Doppler effect on GPS has been carried out in order to improve the velocity estimation accuracy. The relativistic Doppler effects are elaborated and an accurate Doppler shift observation equation is developed;

- Methods of differentiator design have been investigated for real-time applications of velocity and acceleration determination using GPS. A class of first-order Infinite Impulse Response (IIR) differentiators has been derived;
- It has been demonstrated that, at least in the velocity domain (or the Doppler frequency shift domain), the receiver to satellite line-of-sight direction has been changed due to the high-speed motion of the satellite in its orbit. This discovery potentially contributes a better observation equation for GPS range measurements.

1.5 Thesis Organisation

Chapter One is an introduction to the research and thesis. The research background and problems are stated, and the research objectives are specified. The contributions and outcomes of this research are listed. The outline of the thesis is presented.

Chapter Two gives an overview of the GPS system. An introduction is given for the system development history, segmentations, signal structures, and the signal transmission and reception. The observables of the GPS system are elucidated, and the error sources in GPS are discussed. The concept of Doppler frequency shift is introduced and its role in the reception and reconstruction of GPS signal is discussed in brief.

Chapter Three describes the theory of GPS absolute positioning. Fundamentals of precise point positioning are introduced in conjunction with the description of different coordinate and time systems. Error analyses and modelling are discussed in depth for the range measurements.

Chapter Four elaborates the precise velocity and acceleration determination using GPS. The Doppler effect and its relationship with the relative velocity between a receiver and transmitter are introduced. Since a GPS satellite moves at a high speed in its orbit and the GPS signals propagate in the Earth's gravity field, the relativistic

Doppler effect is discussed with both the special relativity and general relativity theories considered. Based on Einstein's relativity theory, a theoretical GPS relativistic Doppler frequency shift formula is derived. Accounting for the error sources during signal transmission, propagation and reception, a state-of-the-art Doppler shift observation equation is developed. The principles of precise acceleration determination using Doppler rates from GPS are also discussed.

Similar to GPS point positioning, the ECEF velocities and accelerations of the in-view GPS satellites should be calculated and made known prior to the ground velocity and acceleration being determined. The algorithm for ECEF satellite position determination using the broadcast ephemeris has been presented in the interface control document of ICD-GPS-200, however, the algorithms for ECEF satellite velocity and acceleration determination are not included. **Chapter Five** discusses real-time GPS satellite velocity and acceleration determination in the ECEF coordinate system using the broadcast ephemeris.

The errors of GPS range measurement have been extensively investigated, especially in the context of PPP. These error sources have effects on the GPS Doppler shift measurement as well as the change-rate of Doppler shifts, and consequently degrade the accuracy of ground velocity and acceleration estimates. **Chapter Six** deals with the errors associated with the Doppler shift measurement and its change-rate. The properties of errors are analysed and the methods to eliminate or mitigate these errors are discussed.

Precise GPS velocity and acceleration determination relies on the Doppler shift and its change-rate observables. However, there are no direct Doppler shift rate measurements in GPS. Although every GPS receiver measures Doppler shifts, some only output 'raw' measurements, and some don't output Doppler shifts at all. In the absence of raw Doppler measurements, a differentiator is necessary to derive it from the GPS carrier phase observables. For real-time and dynamic application, an ideal

differentiator should have a wideband frequency response to cover the dynamics and have a group delay as short as possible, and a low order differentiator is preferred for easy implementation. **Chapter Seven** discusses differentiator designs specific to real-time Doppler or Doppler rate derivation for dynamic applications.

Miscellanea of velocity and acceleration determination are included in **Chapter Eight**. In the case of receiver clock resets, signal loss-of-locks and cycle slips, the derived Doppler shifts and Doppler rates will have "jumps" that in turn deteriorate the velocity or acceleration estimates. A treatment of these problems is presented. Alternative methods are proposed for dealing with the situation when the sampling rate is very high.

Chapter Nine summarises the major conclusions drawn as outcome of this research. It also provides recommendations for future work.

Chapter Two

GPS SYSTEM OVERVIEW

This chapter gives an overview of the GPS system. An introduction is provided of the system development history and the principle of positioning. GPS segments, service standards, signal structures, and the signal modulation and processing are elucidated. GPS error sources are described, as are the GPS observables. A short description of reference systems is also given. The primary purpose of this chapter is to consolidate the knowledge of a reader in preparation for the later chapters.

2.1 Introduction to GPS

The Navigation System with Timing and Ranging (NAVSTAR) Global Positioning System is a satellite-based radionavigation and time-transfer system managed and operated by the US Department of Defense [DoD]. Designed as a dual-use system with the primary purpose for enhancing the effectiveness of US and allied military forces, GPS satisfies the requirements from the military to accurately determine position, velocity and time (PVT) in the World Geodetic System 1984 (WGS-84), anywhere on or near the Earth's surface in all weather and all time. It also contains features that limit the full accuracy of the services only to authorised users and protect it from malicious interference through an implementation of Anti-Spoofing (AS).

The first GPS satellite was launched on the 22nd February 1978 and became operational on the 29th March. Initial Operational Capability (IOC) was declared on the 8th December 1993 when 18 GPS satellites were operating in their designated

orbits, available for navigation use and providing the Standard Positioning Service (SPS). The standards of SPS performance can be found in a document signed by Stenbit [2001]. The US government provides the SPS for peaceful civilian, commercial, and scientific uses on a continuous and worldwide basis, free of direct user fees [Goodman and Jacques, 2000].

On April 27, 1995, the US Air Force Space Command (AFSC) formally declared the GPS satellite constellation as having met the requirements for Full Operational Capability (FOC). These requirements include 24 operational satellites (Block II/IIA) functioning in their assigned orbits and successful testing completed for operational military functionality, known as the Precise Positioning Service (PPS).

Nominally, GPS consists of 24 operational satellites that are arranged so that four satellites are placed in each of the six orbital planes. With such constellation geometry, there are at least four satellites visible anywhere on the Earth surface above an elevation angle of 10 degrees, at all times.

Each GPS satellite carries a set of precise atomic clocks to keep time and provide the signal standard of the fundamental L-band frequency at f_0=10.23MHz. The highly accurate time and frequency standards are the heart of the GPS system. Multiplying the fundamental frequency by 154 and 120 respectively yields the primary L1 carrier frequency at f_1=1575.42MHz and the secondary L2 carrier frequency at f_2=1227.60MHz. These dual frequencies are essential for reducing the ionospheric delay, which is one of the major error sources in GPS positioning.

The GPS system exploits the pseudo-random noise (PRN) coding technique for ranging. A Coarse/Acquisition (C/A) code that has a wavelength of 293m is modulated upon the L1 carrier for civilian users. Precise code (P1), which has a wavelength of 29.3m, is modulated on the L1 carrier as well, but for military users only (PPS). Navigation data known as the broadcast ephemeris is also modulated on the L1 carrier, providing data from which satellite positions can be calculated as well

as satellite clock corrections. Precise code (P2) is modulated on the L2 carrier for PPS only.

A GPS receiver receives the signals transmitted from GPS satellites. For each satellite, the signal travel time can be measured by comparing the replica PRN code from the receiver with the PRN code from the received GPS signal. However, due to the inaccuracies of the receiver clock (crystal oscillator), the measured time has a bias known as the receiver clock delay, which is common to all tracked satellites. Multiplying the travel time by c, i.e., the speed of light in vacuum, the receiver to satellite distance can be determined. This distance is referred to as a "pseudorange", since it is biased by the receiver clock error.

At one measurement epoch, the positions of the in-view GPS satellites in WGS-84 are known from the broadcast navigation messages. With the measured pseudoranges, the three components of the receiver coordinate vector and the biased receiver clock term can be determined if at least four GPS satellites are observed simultaneously.

2.2 GPS Segments

The GPS system comprises three segments, known as the Space Segment, the Control Segment and the User Segment.

2.2.1 Space Segment

The space segment consists of nominally twenty-four GPS satellites (there may be more). GPS satellites operate in six nearly circular orbits with a height near 20,200km, an inclination angle of 55º , and a period of about 11 hours and 58 minutes. Under FOC, the space segment provides for a global coverage with four to eight satellites simultaneously observable above 15º elevation angle at any time. If the elevation mask is reduced to 10º, occasionally up to ten satellites will be in view [Hofmann-Wellenhof et al., 2001].

Each GPS satellite is a platform of radio transmitters, atomic clocks, computers and various ancillary equipment for the purpose of transmitting the dual-frequency signals to ground users. The signals include two carriers, C/A, P1 and P2 PRN codes, and the navigation messages (generally known as the broadcast ephemeris). The broadcast ephemeris contains ionospheric correction parameters, Keplerian orbital parameters and the corresponding corrections, the system time, satellite clock corrections, and the satellite status messages. In addition, an almanac is provided which gives the approximate navigation data for each active satellite. This allows a receiver to more easily find all satellites in view once the first has been acquired, using the approximate positions of the other satellites calculated from the almanac.

2.2.2 Control Segment

The control segment at the early times consisted of five globally distributed Monitor Stations (Hawaii, Kwajalein, Ascension Island, Diego Garcia, Colorado Springs), three Ground Antennas (Ascension Island, Diego Garcia, Kwajalein), and a Master Control Station (MCS) located at Schriever AFB in Colorado [SMC/GP, 2004].

The primary task of the operational control segment is to track GPS satellites in order to determine and predict satellite positions, the system integrity and behaviour of the satellite atomic clocks, atmospheric data, and the satellite almanac. The control segment is also responsible for satellite control and operation.

The monitor stations track all satellites in view passively, accumulating ranging data. The measurements are then transmitted to the MCS where the satellite ephemeris and clock parameters are estimated, and predicted forward in time. The MCS utilises the ground antennas to periodically upload the ephemeris and clock information to each GPS satellite for subsequent broadcasting. The MCS also functions to control satellite manoeuvres, reconfigure redundant satellite equipment, monitor satellite health, etc.

2.2.3 User Segment

The user segment includes all military and civilian users equipped with GPS receivers. These receivers vary significantly in design and functionality, depending on different applications in navigation, surveying, time transfer, or attitude determination.

It is in this segment that the GPS service is categorised into two different levels, i.e. the standard positioning service (SPS) and the precise positioning service (PPS).

The PPS is an accurate positioning, velocity determination and timing service that is available only to authorised users. Based on US defence requirements domestically and internationally, the DoD determines the authorisation. The authorised PPS users include US military forces, NATO military users, and other selected military and civilian users such as the Australian Defence Forces and the US National Geospatial Intelligence Agency.

Access to the PPS is controlled by two cryptographic technologies, namely, Selective Availability and Anti-Spoofing (AS). The AS has been exclusively used for this purpose since May 1, 2000 when the SA was terminated. The AS is activated on all satellites to negate potential spoofing of the ranging signals, by encrypting the P-code into the Y-code. PPS receivers can use either the P(Y) or C/A code, or both, to obtain the maximum GPS accuracy.

The SPS is intended to meet most civilian application requirements. SPS users access the C/A code only. However, with the development of GPS receivers, some users may be able to obtain P-code-like accurate measurements through advanced technology such as the Cross-correlation by Trimble [Rizos, 1999] or the Z-tracking presented by Ashtech [Ashjaee and Lorenz, 1992]. The US government reserves the right to degrade the SPS if it is necessary, for example, to deny accuracy to a potential enemy in time of crisis or war, through reactivating the SA.

2.3 GPS Signal Structures

The GPS satellite transmits two Right Hand Circularly Polarised (RHCP) L-band signals known as the L1 at 1575.42MHz and L2 at 1227.60MHz. Three pseudo-random noise (PRN) ranging codes are in use, namely C/A-code, P-code, and Y-code. Appropriate code-division-multiple–access (CDMA) techniques have been used to differentiate the satellites even though they all transmit at the same L-band frequencies. The adoption of CDMA techniques allows GPS signals to be received and processed with the same set of front-end components. This makes it possible for end user equipment to be relatively light, small, and low-cost.

2.3.1 Coarse Acquisition Code

The C/A code has a 1.023 MHz chip rate, a period of 1 millisecond (ms) and is used for the SPS, or as means to acquire the P-code for PPS. Each satellite transmits a unique C/A code that is from a Gold code family (PRN codes that are distinguished by a very low cross-correlation between any two codes, that is, they are nearly orthogonal).

2.3.2 Precise Code

The P-code has a 10.23 MHz rate, which is ten times faster than the C/A code. The whole length of the P-code sequence is about 266.4 days and each satellite has its own weeklong segment of the P-code sequence. This unique one-week segment of the P-code can be used to identify GPS satellites; for example, a GPS satellite with an ID of PRN 10 refers to the GPS satellite that is assigned the tenth-week segment of the PRN P-code. P-code is the principal navigation ranging code for PPS, reserved for authorised users only.

2.3.3 Y-Code

The Y-code is used in replace of the P-code whenever the AS mode operation is activated. In this case, a W-code is used to encrypt the P-code to generate the Y-code.

2.3.4 Navigation Data

The navigation data includes satellite ephemeris, GPS system time, conversion parameters to the Coordinated Universal Time (UTC), on-board atomic clock behaviour data, the satellite status message, the C/A to P(Y)-code handover word, etc. The 50 bits per second navigation data are modulo-2 added to the C/A and P(Y) codes. The resultant bit-trains are further used to modulate the L1 and L2 carriers. For each satellite, the data trains are common for C/A and P(Y) codes on both L1 and L2 frequencies.

2.3.5 Signal Modulation

Current GPS satellites broadcast signals on two L-band frequencies. These signals have three components: a carrier signal at the centre frequency, a bi-phase shift key (BPSK: phase modulation with $\varphi=\pm\pi$) modulated PRN code(s), and binary navigation data. The C/A code is modulated by a PRN Gold code of 1023 chips at a chipping rate of 1.023 MHz, resulting in a null-to-null bandwidth of 2.046 MHz and a repetition rate of 1ms. The C/A code is designed for civilian access, but can also be used to hand over to the longer P(Y) code which is generated by a modulo-2 addition of two code sequences of 15,345,000 chips and 15,345,037 chips respectively. At a chipping rate of 10.23 MHz, the P(Y) code has a null-to-null bandwidth of 20.46 MHz. The 50 bits per second (bps) navigation data are modulo-2 added to both the C/A and P(Y) codes. On L1, the C/A code has a phase lag of 90º to the P(Y) code, known as phase quadrature. The L2 frequency is modulated with the P(Y) code only, though, at the time of writing, there is an operational Block IIR-M satellite that has a L2C code

modulated on L2. This satellite was launched on 25th September 2005 as part of the GPS "modernization program".

With a 20ms chip (bit) width, the navigation message requires 12.5 minutes to be transmitted in its entirety, although the ephemeris and clock information required for navigation are repeated every 30 seconds.

An analysis of the signal structures and modulations are beyond the scope of this research. However, there is an official definition in the GPS Interface Control Document ICD-GPS-200c [ARINC, 2000], and a thorough description has been given by Spilker [1996b] .

Following Hofmann-Wellenhof et al. [2001], a representation equation is given in Eq. 2-1, with a slight modification to take the phase noise and oscillator drift component φ into consideration

$$\begin{aligned} L_1(t) &= a_1 P(t) W(t) D(t) \cos(2\pi f_1 t + \varphi) + a_1 \cdot C/A(t) \cdot D(t) \cdot \sin(2\pi f_1 t + \varphi) \\ L_2(t) &= a_2 P(t) W(t) D(t) \cos(2\pi f_2 t + \varphi) \end{aligned} \qquad 2\text{-}1$$

where:

- $a_i.cos(2\pi f_i t)$ denotes the unmodulated carriers of L1 and L2;
- f_i is the L1 at 1575.42MHz and L2 frequency at 1227.60MHz respectively;
- *P(t)* represents the P-code sequence, with 1 or 0 as state;
- *C/A(t)* is the C/A-code sequence, with 1 or 0 as state;
- *W(t)* is the W-code sequence, with 1 or 0 as state;
- *D(t)* stands for the navigation message;
- φ represents the phase noise and the oscillator drift component.

Note that the first term on the right hand side of Eq.2-1 is referred to as the inphase signal, and the second term is termed as the quadrature signal, which has a 90º-phase lag. Therefore, GPS L1 signal contains both inphase and quadrature components.

It should be noted that with the GPS signal "modernization program", a C/A code named L2C, is added on the L2 carrier, which is designed to be available for general use in non-safety critical applications. The first Block IIR-M satellite with the L2C had been launched and a third civil signal known as L5 at 1176.45MHz will be transmitted by the satellites in 2007. More details about the GPS "modernization" can be found from the website of the Navigation Center of the US Coast Guard at http://www.navcen.uscg.gov/gps/modernization/default.htm.

2.4 GPS Signal Processing

The context of GPS signal processing is widely published. A brief description is provided here, with an aim of providing an understanding of the roles played by the Doppler frequency shift in GPS signal processing and the way to measure the Doppler shift. For an in-depth knowledge on GPS signal processing, readers are referred to Spilker [1996a], Van Dierendonck [1996], Rizos [1999], and Misra and Enge [2001].

The primary task of signal processing for a GPS receiver is to retrieve the range codes and navigation message from the tracked GPS signals. For a geodetic type receiver, reconstruction of the carrier waves to measure the precise phase is required, and multipath mitigation is essential.

The description of the incoming GPS signal processing is given following Misra and Enge [2001]. The general signal processing begins with the GPS signal reception. A GPS receiver antenna captures the signal and converts it into electrical voltages and currents, which are passed to the Radio-Frequency (RF) front end. The captured, very weak signal is amplified while the carrier frequency is down converted to a lower Intermediate Frequency (IF). Meanwhile, interfering signals in adjacent frequencies

are filtered out. This process is called conditioning. After conditioning, the processed signals are digitised by an analog-to-digital (A/D) converter. At this stage, the digitised IF signals are ready to pass to the 'signal processing' channels where signals from each satellite in-view will be treated separately. Figure 2-1 is a simplified receiver diagram illustrating the flow of signal reception and processing, where LO stands for the local oscillator, I and Q denote inphase and quadrature respectively.

The signal processor section includes carrier and code tracking loops in Fig.2-2. Both the loops work together to deliver output. The output of the signal processing section is the raw GPS measurements, which include pseudoranges, Doppler shifts, and the carrier phase measurements (integrated Doppler). These measurements are then processed by the navigation algorithms to determine the receiver's position, velocity, and time.

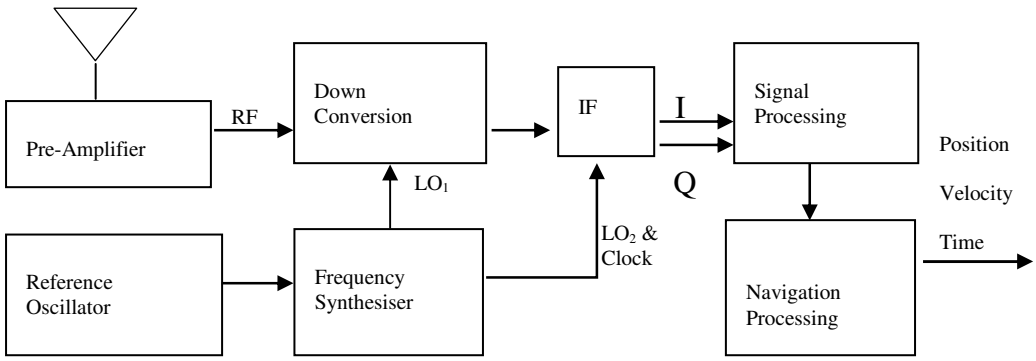

Figure 2- 1: A simplified GPS receiver diagram, from Misra and Enge [2001]

The reference oscillator has the key role of providing the time and frequency reference. The output of the reference oscillator is used in the frequency synthesiser to derive local oscillators and clocks (Numerically Controlled Oscillator, NCO) in each signal-processing channel, as can be seen in Figure 2-1.

In GPS signal acquisition, a receiver needs rudimentary knowledge of the Doppler frequency and code arrival time for each satellite. This can be demonstrated in the

received signal property. The received L1 signal at time t relates to the L1 signal transmitted from a satellite as

$$\begin{aligned} L_1(t) &= a_1 \cdot P(t-\tau)W(t-\tau)D(t-\tau)\cos[2\pi(f_1+f_d)t+\varphi] \\ &\quad + a_1 \cdot C/A(t-\tau)\cdot D(t-\tau)\sin[2\pi(f_1+f_d)t+\varphi] \end{aligned} \qquad 2\text{-}2$$

where:

- τ is the signal travel time from the satellite to the receiver;
- f_d is the Doppler shift due to the relative motion between the satellite and the receiver, as well as the frequency error, and drifts of the satellite and the receiver clocks.

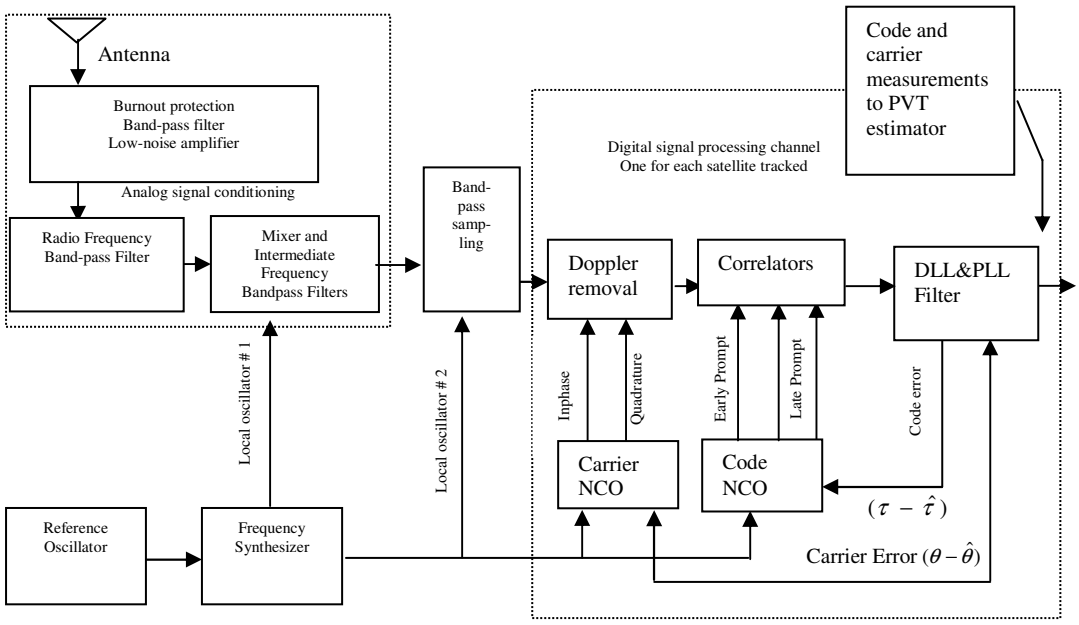

Figure 2- 2: An illustration of GPS signal processing

With rough estimations of f_d and τ, the signal can be put into the appropriate tuneable working window in a signal processing channel. In the channel, the Doppler shift can be estimated in the Doppler removal module through a low pass filter for inphase and quadrature processing. The sign of the Doppler shift may also be resolved here, and it is possible to estimate the Doppler shift without a good estimate of the carrier phase (ibid).

After the Doppler removal, the inphase and quadrature signals are fed into a Delay Lock Loop (DLL). The DLL is used to align the PRN code sequence (C/A or P-code) that is contained in the inphase and quadrature signals with the receiver generated identical PRN code replica sequence. A correlator in DLL continuously cross-correlates the local code stream, shifting stepwise against the received code sequence until the maximum correlation (i.e. alignment) is achieved. The time used in the alignment is equal to the signal transmission time τ if the receiver clock bias is neglected. This is achieved through the unique property that a PRN code sequence has in cross-correlation

$$C(\Delta t) = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} PRN(t)\cdot PRN(t+\Delta t)\cdot dt = \begin{cases} 1 & \Delta t = 0 \\ \approx 0 & \Delta t \neq 0 \end{cases} \qquad 2\text{-}3$$

The measured time is then converted into a pseudorange by multiplying the speed of light in vacuum. The code pseudorange measurement is sometimes referred to as the code phase measurement because the code alignment is actually a measure of the phase of the received code.

Once the code-tracking loop is aligned, the PRN code can be removed from the satellite signal. The stripped signal is then passed to the phase-tracking loop where the satellite message is extracted. Once the local oscillator is locked onto the satellite signals it will continue to follow the variations in the phase of the carrier as the satellite-receiver distance changes continuously. The integrated carrier beat phase

observable is obtained by simply counting the whole elapsed cycles (by noting the "zero crossings" of the beat wave) and by measuring the fractional phase of the locked local oscillator signal [Rizos, 1999].

To close this section, we list the outcome of the GPS signal processing as follows

- C/A code and P(Y) code pseudoranges;
- Doppler shifts;
- Carrier phase measurements;
- Navigation data (broadcast ephemeris);
- Signal-to-noise ratio (SNR) information.

2.5 GPS Error Sources

There are generally six classes of errors in GPS measurements according to Parkinson [1996]. They are ephemeris and satellite clock errors, ionosphere errors, troposphere errors, multipath errors, and receiver errors. A short introduction to the errors is provided here. In-depth properties of the errors will be analysed, and the error modelling methods will be investigated in Chapter Three. The time-varying behaviour of error affects the velocity and acceleration determination using GPS, which is the main topic of this thesis.

2.5.1 Satellite Dependent Errors

The source of ephemeris and satellite clock errors is the GPS control and space segments. Although GPS orbits have been carefully observed and calculated, and the GPS time has been kept precisely, the predicted satellite orbit and the broadcast satellite clock parameters differ from their true values. These errors are in the navigation data broadcast by GPS satellites, influence the range measurements, and consequently affect GPS positioning.

Errors induced by the relativistic effects are also in this category. The relativistic effect causes an apparent frequency shift in the satellite oscillator due to the presence of the Earth's gravity field. This frequency shift may be almost compensated by reducing by 0.00457Hz the nominal satellite oscillator's frequency at 10.23MHz. A GPS satellite orbits the Earth in a non-circular orbit with a relatively high-speed of about 3.8km/s, this produces another periodic error that must be subtracted from the satellite clock value [Hofmann-Wellenhof et al., 2001]. More details will be discussed in the following chapters.

2.5.2 Propagation Errors

GPS signal propagation errors include the ionosphere error, troposphere error and the multipath error.

Due to the presence of free electrons in the ionosphere, GPS signals do not travel at the speed of light in vacuum, nor as straight lines, but with bent signal paths. The modulation on the signal is delayed in proportion to the total electron number and is inversely proportional to the squared carrier frequency. The phase of the radio frequency carrier is advanced which is termed as the "phase advance", the PRN code is delayed by the same amount, and is referred to as the "group delay". The ionosphere is thus called a dispersive medium due to this characteristic. The GPS system takes advantage of the dispersive property to use two L-band frequencies to eliminate the first-order ionosphere range and range-rate errors [Klobuchar, 1996].

The troposphere is an electrically neutral atmospheric region that extends up to about 50km from the surface of the Earth [EL-Rabbany, 2002]. It is a non-dispersive medium for radio frequencies below 15GHz [Hofmann-Wellenhof et al., 2001]. As a result, the troposphere delays the GPS carrier and code measurements identically, and the measured satellite to receiver distance is longer than the true geometric distance. The tropospheric delay depends on the temperature, pressure and humidity along the

signal path. This path delay is of the order of 2~25m [Spilker, 1996c]. The tropospheric delay varies with elevation angles because a lower elevation angle produces a longer signal path delay through the troposphere.

Multipath error occurs when a GPS signal arrives at a receiver antenna through more than one path. These paths may be the line-of-sight signal and reflected signals from objects surrounding the receiver antenna. The interference by the reflected signals at the GPS antenna affects both the carrier phase and code measurements. The magnitude of the carrier phase multipath can reach a maximum of a quarter of cycle [Hofmann-Wellenhof et al., 2001] while the pseudorange multipath may amount to 10-20m [Wells et al., 1987]. The multipath error is a function of the wavelength, and has a periodic characteristic. More discussions on multipath effect can be found in Braasch [1996]. Note that some authors may group the multipath error into the receiver dependent error category.

2.5.3 Receiver Errors

Errors in a GPS receiver consist of the thermal noise, antenna phase centre variation, inter-channel and inter-frequency biases, and the receiver clock error that is the offset between the receiver clock and the GPS system time. Receiver errors do depend to some extent on the quality of design and manufacturing. In this research, only the receiver clock error is of interest.

2.6 GPS Observables

There are three types of GPS measurements, namely code pseudorange, carrier phase and the Doppler frequency shift. The code pseudorange and carrier phase measurements are generally considered as the two basic range observables between the observed satellites and the receiver. The Doppler shift observable, as has been introduced in §2.4, is a by-product of receiver signal processing.

2.6.1 Code Pseudorange Measurement

In GPS, code pseudorange measurements are often called "pseudoranges", however this is inaccurate since the carrier phase measurements may also be regarded as pseudoranges when the integer ambiguities are considered. In this thesis, when the term pseudorange is used, it refers to the *code* pseudorange. Code or code range is used as an alias of the code pseudorange.

Code pseudorange measurements are made by comparing the incoming signal from a GPS satellite with the receiver replicated PRN code. The time shift required in correlating the two signals is equivalent to the signal propagation time between the GPS satellite and the receiver. The distance between the GPS satellite and the receiver can therefore be measured; however it is biased due to the lack of synchronisation between the receiver and satellite clocks to the GPS time. The code phase observable may be expressed as [Teunissen and Kleusberg, 1998]

$$\begin{aligned} P_{r,i}^{s}(t) = & \left\| (\mathbf{r}^{s}(t-\tau_{r}^{s}) + d\mathbf{r}^{s}(t-\tau_{r}^{s})) - (\mathbf{r}_{r}(t) + d\mathbf{r}_{r}(t)) \right\| + dI_{r,i}^{s} + dT_{r}^{s} + \\ & c \cdot dt_{r}(t) - c \cdot dt^{s}(t-\tau_{r}^{s}) + dm_{r}^{s} - dR_{r}^{s} + \varepsilon_{r}^{s} \end{aligned} \qquad 2\text{-}4$$

where:

- Subscript *r* denotes the receiver where the measurement is made;
- Subscript *i* denotes the frequency band of the observation, *i=1* for L1, and *i=2* for L2;
- Superscript *s* represents the satellite being observed;
- *t* is the true epoch time in the GPS time system;
- $P_{r}^{s}(t)$ is the code pseudorange between receiver *r* and satellite *s*, at GPS time *t*

- $\boldsymbol{r}$ is a position vector of the three receiver coordinate components (x ,y, z), defaults in the WGS-84 system;
- d is used to represent an error term;
- $d\mathbf{r}^s$ represents the positional error of satellite s;
- $d\mathbf{r}_r$ represents the positional error of receiver r;
- τ_r^s denotes the signal transmission time from satellite s to receiver r;
- $\| \ \|$ is the norm of a vector, or simply the length of the vector;
- dI denotes the ionospheric delay;
- dT denotes the tropospheric delay;
- c is the speed of light in vacuum;
- dt_r is the receiver clock delay;
- dt^s is the satellite clock delay;
- dm is the multipath effect on the code range;
- dR is the relativistic effect due to the satellite motion and the presence of the Earth gravitational field;
- ε is the measurement error associated with the observation.

The code pseudorange measurement $P_{r,i}^s(t)$ is biased with respect to the satellite-to-receiver geometric distance by the above-listed errors. Note that in Eq.2-4 all error terms are in their absolute values, being scaled into unit of metres.

GPS code signal in the ionosphere is delayed due to the group delay effect, resulting in an increased distance measurement, and therefore a positive sign is assigned to the ionospheric correction. A time delay also occurs when the signal travels in the

troposphere, so the tropospheric correction is also positive. Similarly, a positive sign is assigned to the delay due to the receiver clock as it contributes to an extra signal travel distance, but a negative sign is with the delay of satellite clock, as this would decrease the measured distance. The relativistic effect slows down a satellite clock, and therefore has a negative sign (which is the same as the satellite clock correction). The multipath effect is caused by indirect signals coming into the receiver antenna by reflection. When multipath occurs, the measured distance will always be longer than the direct distance, and thus the error has a positive sign.

In GPS, an observable is referred to as the measurement of signal from its transmission to its reception. The signal travel time τ relates a transmitted signal to a received signal in the standard GPS time frame. In Eq.2-4, the geometric distance is between the position of receiver at reception time t and the satellite position at transmission time t-τ.

2.6.2 Carrier Phase Measurement

The measure of the carrier phase of the GPS signal does not require knowledge of the actual information being transmitted. Once a receiver locks onto a GPS satellite signal, the carrier phase is measured by phase comparison between the received carrier signal and the receiver generated reference carrier waves. In the Phase Lock Loop (PLL), the fractional part of the received phase is precisely measured and an integrated Doppler counter accumulates the changes of integer wavelength. The sum of the integer cycle count and the fractional phase part is the carrier phase measurement. Scaling with the wavelength of the carrier signal and introducing an unknown initial integer number of carrier cycles referred to as the *integer ambiguity*, the carrier phase measurement relates to the receiver and satellite position as follows [Teunissen and Kleusberg, 1998]

$$\lambda_i \varphi_{r,i}^s(t) = \left\| (\mathbf{r}^s(t-\tau_r^s) + d\mathbf{r}^s(t-\tau_r^s)) - (\mathbf{r}_r(t) + d\mathbf{r}_r(t)) \right\| - dI_{r,i}^s + dT_r^s + c \cdot dt_r(t) - c \cdot dt^s(t-\tau_r^s) - \lambda_i N_{r,i}^s + \lambda_i \varphi_{r,i}^s(0) + dM_r^s - dR_r^s + \varepsilon_r^s \quad \text{2- 5}$$

where

- λ_i is the wave length of carrier phase, 19.03cm for L1 and 24.42cm for L2;
- $\varphi_{r,i}^s(t)$ is the carrier phase observable on the *i*-th frequency for receiver *r* and satellite *s* at epoch *t*;
- $N_{r,i}^s$ is the integer ambiguity of the carrier phase measurement of the *i-th* frequency;
- $\varphi_{r,i}^s(0)$ is the initial fraction of carrier phase on the *i*-th frequency for receiver *r* and satellite *s* when the receiver achieves initial lock on;
- ε is the noise of the carrier phase measurement.

Comparing with the code phase measurement, one may find that a negative sign is given for the ionospheric delay; this is due to the "*phase advance*" (as has been already discussed). The multipath effect is capitalised to highlight the significant difference in magnitude. The carrier phase pseudorange makes sense if the $N \cdot \lambda$ term is moved from the right hand side of Eq.2-5 to the left hand side. This is depicted in Fig.2-3.

The integer ambiguity, *N*, is typically not known and is different for each receiver-satellite pair. As long as the tracking of the satellite is not lost, *N* remains constant, while the fractional phase and the integrated integer Doppler counter change over time. The integer ambiguity either can be solved using the code phase measurement in the measurement domain or be estimated in the coordinate domain, or be "searched" in the ambiguity domain. The change of the integer number of a receiver-satellite pair

is referred to as "cycle slip." When a cycle slip occurs, it is necessary to introduce a new integer ambiguity [Rizos, 1999].

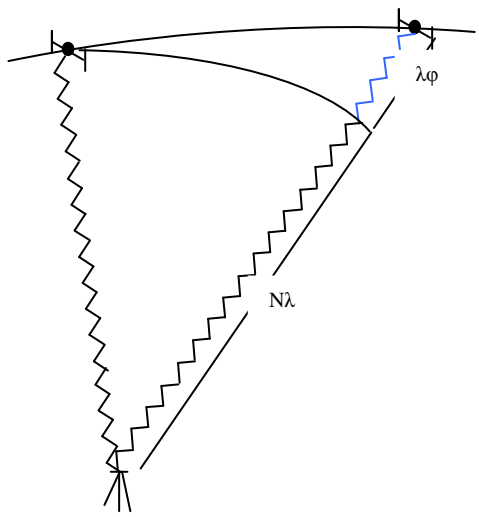

Figure 2- 3: The carrier phase measurement vs. the integer ambiguity

2.6.3 Doppler Shift Measurement

The Doppler shift is the frequency difference between the received signal and the source signal due to the relative motion between a receiver and a transmitter. As the carrier phase measurement is the difference between the phase of the receiver-generated carrier signal and the received carrier from a satellite at the instant of measurement, a more precise name for the carrier phase measurement is the “carrier beat phase measurement”. A GPS signal reaching the antenna of a receiver is Doppler shifted in frequency and it is necessary for the GPS receiver to have an estimate of the Doppler shift in order to bring the received signal (RF/IF) into the signal tracking loops. Otherwise the signal would be out of the working “window”, or rather, the receiver’s working bandwidth. As a result, there is no phase to beat! The more accurate the Doppler shift is estimated, the smoother a carrier-tracking loop will work, and consequently the more precise a carrier phase will be measured.

The carrier phase measurement of a GPS receiver is actually the integrated Doppler shift measurements [Seeber, 1993; Rizos, 1999; Hofmann-Wellenhof et al., 2001]. Thus, the Doppler shift, a "by-product" of the GPS receiver signal processing, relates to the carrier phase measurement as its first derivative with respect to time. With this, the Doppler shift observation can be written

$$\begin{aligned}\lambda_i D_{r,i}^s(t) = \lambda_i \dot{\varphi}_{r,i}^s(t) = \left\| (\dot{\mathbf{r}}^s(t-\tau_r^s) + d\dot{\mathbf{r}}^s(t-\tau_r^s)) - (\dot{\mathbf{r}}_r(t) + d\dot{\mathbf{r}}(t)) \right\| - d\dot{I}_{r,i}^s + d\dot{T}_r^s + \\ c \cdot d\dot{t}_r(t) - c \cdot d\dot{t}^s(t-\tau_r^s) + d\dot{M}_r^s - d\dot{r}_r^s + \varepsilon_r^s\end{aligned} \qquad 2\text{-}6$$

where a dot over a variable represents its first derivative with respect to time. Compared to Eq.2-5, the carrier phase observation, it can be seen that the integer ambiguity term and the initial fraction part have vanished, or in other words, the Doppler shift is free of integer ambiguity problems.

GPS Doppler shift observables in general are much "cleaner" than the carrier phase observables, since the errors and biases are the time derivatives of the error sources in the carrier phase measurements.

2.7 Coordinate Systems

World Geodetic System 1984 (WGS-84) is officially used by the GPS system as the datum. The broadcast ephemeris is calculated in WGS-84 and therefore, by default, user positions from GPS are generated in the WGS-84 coordinate system.

The WGS-84 coordinate system is a realisation of the Earth-Centred-Earth-Fixed (ECEF) coordinate system, which is fixed to the Earth and rotates with the Earth. The ECEF WGS-84 is defined as [ARINC, 2000]:

- Origin at the Earth's centre of mass;
- Z-axis extends through the IERS (International Earth Rotation Service) reference pole;

- X-axis directs to the intersection of the IERS reference meridian (IRM) and the plane passing through the origin and normal to the Z-axis;
- Y-axis completes the right-hand coordinate system, passing through the equator and 90° longitude east of the X-axis.

The major parameters of the WGS-84 system are given in Table 2-1 [cf. Seeber, 2003]

Parameter	Notation	Value	Unit
Semi-major axis	a	6378137	m
Flattening	f	1/298.257223563	
Angular velocity	ω	$7.292115 \cdot 10^{-5}$	$rad \cdot s^{-1}$
Geocentric gravitational constant	GM	398600.4418	$km^3 s^{-2}$
Second zonal harmonic	$\overline{C}_{2,0}$	$-484.16685 \cdot 10^{-6}$	

Table 2- 1: Major Parameters of WGS-84

Newtonian mechanics work in inertial systems that are neither rotating, nor accelerating. For the description and prediction of the motion of a GPS satellite, an Earth-centred inertial reference system is needed. Due to the presence of the Earth's gravitational field (and therefore gravitational accelerations), such an Earth centred inertial system can only be obtained by approximation and careful designation of the origin and 3-axis directions such that the resultant system is non-rotational with each axis directed to a fixed celestial point.

An implementation of such an inertial system is the Earth-Centred-Inertial (ECI) system, which is defined as

- Origin at the centre of mass of the Earth;
- Z-axis along the Earth's rotation axis, pointing to the north pole;

- X-axis in the equatorial plane directing to the vernal equinox;
- Y-axis completes the right-hand coordinate system, passing through the equator and 90˚ longitude.

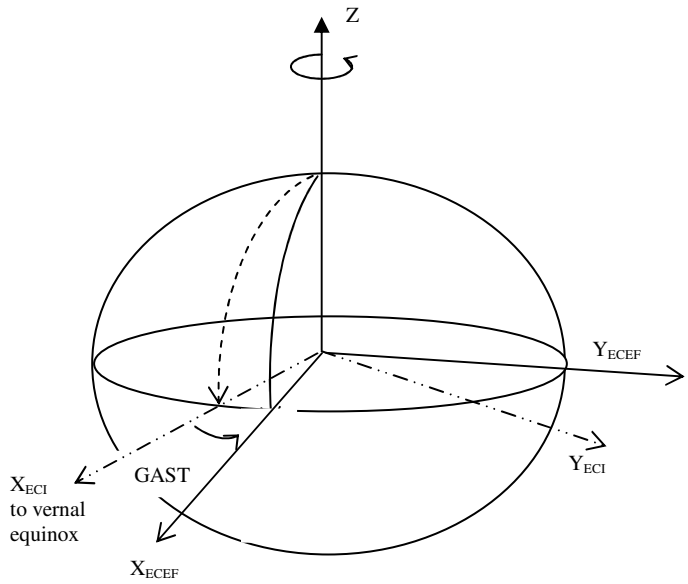

Figure 2- 4: ECEF and ECI coordinate systems

Figure 2-4 depicts the relationship between the ECEF and ECI systems where angle GAST stands for the Greenwich Apparent Sidereal Time (GAST), which changes with respect to time and is related to the Earth rotation rate by

$$GAST(t) \approx GAST(t_0) + \omega(t - t_0) \qquad 2\text{-}7$$

where t_0 is a reference time.

Note that approximations have been made in both Fig.2-4 and the GAST. Due to the polar motion, nutation and precession of the vernal equinox, the precise definition of ECI requires the specification of the positions of the pole and vernal equinox at a specific convention epoch such as the J2000. This allows correction of the

instantaneous pole and vernal equinox to their conventional positions fixed in the celestial system. Readers are referred to the work by Bock [1998], and for space geodesy applications, to McCarthy [2000]. It is sufficient to use the GAST to distinguish ECEF from ECI for most real-time applications using GPS.

Chapter Three

PRECISE ABSOLUTE GPS POSITIONING

This chapter elucidates the precise point positioning for real-time applications. From the simplest navigation solution, the least squares adjustment and accuracy assessment are then introduced. Error modelling of one-way un-differenced observation is elaborated using the methods developed in PPP. The error analyses are important as all the error sources in PPP have effects on precise velocity and acceleration determination. The relativistic effects and troposphere modelling are emphasised since they are the two major error sources limiting the accuracy of precise velocity determination using GPS.

3.1 GPS Point Positioning

To begin with, all the inherent errors in GPS measurements are neglected except the receiver clock error. This simplification assumes that those neglected errors have been properly modelled and accounted for. As such, the code pseudorange measurement from receiver *r* to satellite *s* at epoch *t* is expressed as

$$P_r^s(t) = \left\| \mathbf{r}^s(t - \tau_r^s) - \mathbf{r}_r(t) \right\| + c \cdot dt_r + \varepsilon \qquad 3\text{-}1$$

where:

- $P_r^s(t)$ is the code pseudorange between receiver *r* and satellite *s* at GPS time *t*;

- τ_r^s denotes the signal propagation time from satellite *s* to receiver *r*;
- $\mathbf{r}^s$ is the satellite position vector $(x^s, y^s, z^s)^T$ in the orbit, which is assumed known;
- $\mathbf{r}_r$ is the receiver position vector $(x_r, y_r, z_r)^T$ on the Earth, which is the unknown to be estimated;
- $\| \ \|$ is the norm of a vector, or simply its length;
- c is the speed of light in vacuum;
- dt_r is the receiver clock delay;
- ε is the measurement error associated with the observation.

In the above simplified observation equation there are four unknowns: three components in the receiver position vector and one receiver clock delay. At least four observations are required in order to solve the four unknowns at one epoch. Redundant measurements exist if more than four measurements are available. In the case of redundant measurements, the least squares adjustment procedure can be applied to obtain an improved positioning estimation.

3.1.1 Linearisation

However, the least squares adjustment requires the observation equation system to be linear. For the geometric distance between receiver *r* and the satellite *s*

$$P_{rg}^s = \left\| \mathbf{r}^s(t - \tau_r^s) - \mathbf{r}_r(t) \right\| = \sqrt{(x^s - x_r)^2 + (y^s - y_r)^2 + (z^s - z_r)^2} \qquad 3\text{-}2$$

the linearisation of P_{rg}^s requires approximate values of the receiver coordinate components, i.e. $\boldsymbol{r}_r^0 = (x_r^0, y_r^0, z_r^0)^T$, through which, P_{rg}^{s0}, the approximation of P_{rg}^s,

can be calculated. The linearisation can be carried out using the Taylor series approximation, with higher order terms neglected

$$
\begin{aligned}
P_{r\ g}^{s} &= P_{r\ g}^{s\,0} + \frac{\partial}{\partial x} P_{r\ g}^{s} \cdot \Delta x + \frac{\partial}{\partial y} P_{r\ g}^{s} \cdot \Delta y + \frac{\partial}{\partial z} P_{r\ g}^{s} \cdot \Delta z \\
&= P_{r\ g}^{s\,0} - \begin{bmatrix} \dfrac{x^{s} - x_{r}^{0}}{P_{r\ g}^{s\,0}} & \dfrac{y^{s} - y_{r}^{0}}{P_{r\ g}^{s\,0}} & \dfrac{z^{s} - z_{r}^{0}}{P_{r\ g}^{s\,0}} \end{bmatrix} \cdot \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} \\
&= P_{r\ g}^{s\,0} - \mathbf{n}_{r}^{s} \cdot \Delta \mathbf{r}_{r}
\end{aligned}
\qquad 3\text{-}3
$$

where

- $\boldsymbol{n}_r^s$ is the line-of-sight unit vector, a normal vector sometimes referred to as the direction cosine vector;
- $\Delta \boldsymbol{r}_r$ is the receiver position discrepancy vector, i.e. the position correction.

With this, the GPS code pseudorange observation equation can be rewritten into the following linear form

$$
P_{r}^{s}(t) = P_{r\ g}^{s\,0} - \mathbf{n}_{r}^{s} \cdot \Delta \mathbf{r}_{r} + c \cdot dt_{r} + \boldsymbol{\varepsilon} \qquad 3\text{-}4
$$

Usually it is assumed that the measurement error ε is of normal distribution in statistics, with a zero mean and a variance of σ^2, i.e. $\varepsilon \sim N[0, \sigma^2]$. The equivalent error equation for the code pseudorange measurement is then given as

$$
\begin{aligned}
v_{r}^{s} &= -\mathbf{n}_{r}^{s} \cdot d\mathbf{r} + c \cdot dt_{r} - (P_{r}^{s}(t) - P_{r\ g}^{s\,0}(t)) \\
&= \begin{bmatrix} -\mathbf{n}_{r}^{s} & 1 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{r}_{r} \\ c \cdot dt_{r} \end{bmatrix} - f_{r}^{s}
\end{aligned}
\qquad 3\text{-}5
$$

where:

- v is the residual of an observation, i.e. the difference between the observation and its estimated value;
- f is a constant term which is the difference between the measurement and its nominal value calculated using the approximations of the unknown parameters.

3.1.2 Least Squares Solution

When a GPS receiver observes more than four satellites simultaneously, a set of linearised observation error equations can be formed based on Eq.3-5

$$\begin{bmatrix} v_r^1 \\ v_r^2 \\ \vdots \\ v_r^n \end{bmatrix} = \begin{bmatrix} -\mathbf{n}_r^1 & 1 \\ -\mathbf{n}_r^2 & 1 \\ \vdots & \vdots \\ -\mathbf{n}_r^n & 1 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{r}_r \\ c \cdot dt_r \end{bmatrix} - \begin{bmatrix} f_r^1 \\ f_r^2 \\ \vdots \\ f_r^n \end{bmatrix}$$

or 3- 6

$$\mathbf{V} = \mathbf{GX} - \mathbf{F}$$

where n denotes the number of the observed satellites (n>4), $\boldsymbol{G}$ is a $n \times 4$ matrix characterising the receiver-satellite geometry through the line-of-sight unit vectors.

The observations are assumed to be independent on each other, with equal measurement weight $\boldsymbol{P=I_n}$, (unit matrix), i.e. equal measurement precision σ^2. The least squares estimation of the unknown corrections of the receiver position and clock delay can be obtained by assuming that the estimated parameters minimise the sum of weighted square residuals, i.e. $\boldsymbol{V^TPV}=min$

$$\hat{\mathbf{X}} = (\mathbf{G}^T\mathbf{PG})^{-1} \cdot \mathbf{G}^T\mathbf{PF} = (\mathbf{G}^T\mathbf{G})^{-1} \cdot \mathbf{G}^T\mathbf{F} \qquad 3\text{- }7$$

The least squares estimation is the best linear unbiased estimates (BLUE) of the above equation system. Iterations may be needed for least squares estimation; however, typically just one or two iterations may be sufficient for convergence.

3.1.3 Positioning Precision Assessment

The precision of the least squares estimates can be evaluated through the variance-covariance matrix of the estimates by

$$\mathrm{cov}\begin{bmatrix} \Delta\hat{\mathbf{r}}_r \\ c\hat{dt}_r \end{bmatrix} = \mathrm{cov}\begin{bmatrix} \hat{\mathbf{r}}_r \\ c\hat{dt}_r \end{bmatrix} = \hat{\sigma}^2 \cdot (\mathbf{G}^T\mathbf{G})^{-1} \qquad 3\text{-}8$$

where $\hat{\sigma}^2 = \dfrac{\boldsymbol{V}^T\boldsymbol{V}}{n-4}$ is the estimated measurement variance of σ^2.

Equation 3-8 shows that the accuracy of estimates of a user's position and clock is in terms of the code pseudorange measurement precision and the ***G*** matrix, which is the design matrix containing the components of line-of-sight unit vectors. In other words, a user's positional estimates are dependent of the measurement accuracy and the satellite geometry at the instant of observation.

In GPS applications, Dilution of Precision (DOP) is often used to quantify the satellite geometry, and the quality of the estimates that may be obtained under such geometry. To introduce DOP, ***H*** matrix from Eq.3-8 is defined as

$$\mathbf{H} \equiv (\mathbf{G}^T\mathbf{G})^{-1} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{x,cdt} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{y,cdt} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{z,cdt} \\ \sigma_{cdt,x} & \sigma_{cdt,y} & \sigma_{cdt,z} & \sigma_{cdt}^2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{11} & h_{11} & h_{11} \\ h_{21} & h_{22} & h_{23} & h_{23} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{42} & h_{43} & h_{44} \end{bmatrix} \qquad 3\text{-}9$$

and elements of h_{ii}, the i–th on the diagonal of ***H*** matrix are used to define different DOPs, for instance:

- Position: $PDOP \equiv \sqrt{h_{11} + h_{22} + h_{33}}$ 3- 10
- Geometric: $GDOP \equiv \sqrt{h_{11} + h_{22} + h_{33} + h_{44}}$ 3- 11
- Time: $TDOP \equiv \sqrt{h_{44}}$ 3- 12

Other DOPs such as the HDOP (H for horizontal) and VDOP (V for vertical) in GPS will need coordinate rotations to the $\boldsymbol{H}$ matrix to conduct transform from the ECEF system to a local navigation systems such as the NEU (North-East-Upward) system.

GDOP is the geometric effect of the spatial relationship of the in-view satellites in relation to a user receiver. It can be viewed as the "strength of figure" of the trilateration computation. A low GDOP value normally reflects good satellite geometry and thus a potential good position fix.

In summary, the navigation solution through the least squares approach requires the *a priori* position of a receiver, more than four satellites in view, and an iterated solution. A non-iterative analytical navigation solution is provided by Bancroft [1985], which is capable of solving the position unknowns and clock bias when there are only four satellites. More details of this approach and its development can be found in Goad [1998], Yang and Chen [2001].

3.1.4 Earth Rotation Correction

A GPS satellite has an orbit height of almost 20,200 kilometres above the Earth surface. It takes about 0.075 seconds for GPS signals to reach a ground receiver. The reception signal has travelled an extra distance during this short signal propagation time due to the Earth rotation. Ashby and Spilker [1996] considered this as a relativistic error known as the Sagnac effect, and provided a formula to correct for it. The Sagnac correction will be discussed later. In this section, a simple iterative treatment for the navigation solution is introduced as an alternative.

The vertical bars in Eq.3-2 represent the geometric distance between the satellite and receiver, which is valid in an inertial system [ARINC, 2000]. Because the satellite positions are calculated from the broadcast ephemeris, which is provided in ECEF system, and the user position is in ECEF, the Earth rotation must be accounted for to make sense of this inertial geometric distance. In-depth descriptions of the calculation of the geometric distance term when using ECEF coordinates are provided by Goad [1998], the following gives the procedure to compensate for the Earth rotation

$$P_{r\ g}^{s} = \left\| \mathbf{R}_3(\omega \cdot \tau_r^s)\mathrm{r}^s(t-\tau_r^s)_{ECEF} - \mathrm{r}_r(t)_{ECEF} \right\| \qquad 3\text{-}13$$

where $\boldsymbol{R}_3(\theta)$ is the rotation matrix for rotating angle θ counter-clockwise around Z-axis (the third axis that subscript 3 stands for), which is given as

$$\mathbf{R}_3(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad 3\text{-}14$$

Such a treatment fixes the instantaneous ECEF orientation at time t as the orientation of the inertial system.

Therefore, in a practical navigation solution, one needs to apply a rotation first to the ECEF satellite position at the transmitting time using Eq.3-13 to obtain the correct geometric distance. More specifically, this process is an iterative process since the propagation time is unknown. A commonly used method begins with setting τ_0=0.075s, followed by calculating the satellite ECEF position using the ICD-200-GPSc algorithm [ARINC, 2000], and applying the rotation to the satellite position to account for the Earth rotation, and finally calculates the travel time τ_1. The procedure is iterated until τ_i converges to a certain level. Leick [1995] has given a very intuitive description in this regard.

3.2 GPS Errors and Modelling

The navigation solution for GPS point positioning described in § 3.1 does not account for the observation errors. As has been mentioned in Chapter Two, GPS measurements manifest many errors and biases. In order to achieve a higher accuracy from the absolute point positioning, these errors and biases should be appropriately modelled and corrected for. The modelling methods are discussed, following the sequence of satellite dependent errors, propagation errors, and receiver dependent errors, as was done in Chapter Two.

3.2.1 Satellite Dependent Errors

3.2.1.1 Ephemeris Errors

Since this research mainly concentrates on real-time GPS applications, the discussion is constrained only to the broadcast ephemeris errors. These errors are the same for both code pseudoranges and carrier phase pseudorange measurements, as can be seen from the term $d\mathbf{r}^s$ in both Eq.3-1 and Eq.3-2. With SA off, the quality of the clock and ephemeris parameters in the broadcast navigation message has been significantly improved. Currently it is widely recognised that the broadcast orbit error is of the order of ±2~5m RMS.

The ephemeris error contains three components, i.e. radial, tangential and cross-track errors. The tangential and cross-track errors are much larger than the radial error [Roulston, 2001]. The satellite orbit error affects a user position fix in two ways: first, it affects the range measurement accuracy through a projection of the orbital position error components along the line-of-sight direction; secondly, it influences the user position accuracy through intersection of the biased positions of GPS satellites.

3.2.1.2 Satellite Clock Errors

GPS onboard satellite clocks, although being highly accurate, have errors that must be corrected. The GPS control segment utilises a set of precise atomic clocks to define the GPS time standard and monitors the onboard clocks. The satellite clock errors are calculated, predicted into the future, and uploaded to GPS satellites for broadcasting. Polynomials are used to determine the effective SV PRN code phase offset referenced to the phase centre of the antenna (Δt_{sv}) with respect to GPS system time, at the time of signal transmission. The broadcast parameters of the satellite clock account for the deterministic SV clock error characteristics of bias, drift and aging, as well as for the SV implementation characteristics of the group delay bias and the mean differential group delay. The algorithm to correct satellite clock errors is given in ICD-GPS-200c [ARINC, 2000]. A user receiver should correct the time received from a SV with the following equation in seconds

$$t = t^s - dt^s \qquad 3\text{-}15$$

where:

- t is the true GPS time in seconds;
- t^s represents the effective PRN code phase time at the transmission time in seconds;
- dt^s is the SV PRN code phase time offset in seconds.

The satellite PRN code phase offset is given by

$$dt^s = a_{f0} + a_{f1}(t - t_{oc}) + a_{f2}(t - t_{oc})^2 \qquad 3\text{-}16$$

where:

- a_{f0}, a_{f1}, a_{f2} are the polynomial coefficients in subframe one of the navigation message;
- t_{oc} represents the clock data reference time in seconds.

The satellite group delay T_{gd} is due to the satellite hardware bias. It is similar to the initial phase bias in GPS receiver errors (see §2.6.2). T_{gd} is initially calibrated by the satellite manufacturer to account for the effect of satellite group delay differential between L1 P(Y) and L2 P(Y), and then updated by the GPS control segment to reflect the actual orbital group delay difference. Single-frequency users (using only L1 or L2 measurements) need to apply this correction to the satellite clock error dt^s by

$$\begin{aligned} (dt^s)_{L1P(Y)} &= dt^s - T_{gd} \\ (dt^s)_{L2P(Y)} &= dt^s - (\frac{77}{66})^2 T_{gd} \end{aligned} \qquad 3\text{-}17$$

3.2.1.3 Relativistic Effect Corrections

Since a GPS satellite orbits the Earth at high speed, transmitting signals that travel through the Earth's gravity field at the speed of light, Einstein's theories of general and special relativities are applicable. Details of the relativistic effects in GPS is referred to Ashby and Spilker [1996], Jaldehag et al. [1998], and Ashby [2003]. In general, the relativistic effects are on GPS signal propagation and behaviour of the board satellite clocks.

The special and general relativistic effects on GPS onboard clocks due to the motion in circular orbits have been already accounted for by the GPS system. This is accomplished by shifting the system fundamental frequency of 10.23 MHz to 10.229 999 999 543MHz, i.e. (-0.004647 Hz correction) [Ashby and Spilker, 1996]. Thus, a user needs not be concerned about this error. In this section, only the significant relativistic effects that must be corrected for PPP are discussed. The research focus,

however, is to investigate their effects on Doppler measurements, which will be detailed in the following chapters.

Orbit Eccentricity Correction for Satellite Clock

The slight eccentricity of a satellite orbit causes a periodic relativistic effect on satellite clocks. When the satellite is at the perigee, it travels faster and has a lower gravitational potential. As a result, the satellite clocks run slower. Since the clock parameters in the navigation message do not include corrections for this relativistic effect, a user must determine the requisite relativistic corrections. In the ICD-GPS-200c, this relativistic correction is calculated and applied as an additional term in the satellite clock correction in seconds [ARINC, 2000]. As a separate term in the GPS pseudorange observation equation, however, it is more convenient to scale it into an equivalent distance in metres as

$$\delta R_{eccentricity} = -\frac{2\sqrt{GM}}{c} e \cdot \sqrt{a} \cdot \sin E_k \qquad 3\text{-}18$$

where : c

- $GM = 3.986005 \times 10^{14}\ m^3/s^2$, is the product of the universal gravitation constant G and the Earth's mass M;
- e represents the eccentricity of the satellite orbit;
- a is the semi-major axis of the satellite orbit; and
- E_k is the orbital eccentric anomaly.

The above relativistic correction is normally obtained after solving the Kepler's equation, however, in this thesis and its associated programming, an alternative but equivalent form is used, which is expressed in terms of the satellite position and velocity by

$$\delta R_{eccentricity} = -\frac{2}{c} \cdot \mathbf{r}^s \bullet \dot{\mathbf{r}}^s \qquad 3\text{- }19$$

where $\dot{\boldsymbol{r}}^s$ is the instantaneous velocity vector of the GPS satellite.

The GPS control segment utilises Eq.3-19 to estimate the navigation parameters. Since GPS satellite velocity and acceleration will always be estimated in this research, it is straightforward to calculate this correction. Moreover, one may find it advantageous to analyse the change-rate, which will be discussed in Chapter Six.

This correction may reach a maximum of 21m in range measurements according to Leva et al. [1996].

Sagnac Effect Correction

The Sagnac effect due to the Earth rotation has been mentioned in §3.1.4, where an iterative treatment is given. Here an alternative treatment is introduced than can apply corrections to the raw pseudorange measurements directly.

Aberration is the alternative term used for the Sagnac effect in space geodesy. A theoretical formula of the aberration correction for the integrated Doppler measurement can be found in Seeber [2003]

$$\delta \mathrm{R}_{\mathrm{aberration}} = \frac{1}{\mathrm{c}}(\boldsymbol{\Omega}_{\mathrm{e}} \times \mathbf{r}^{\mathrm{s}}) \bullet (\mathbf{r}^{\mathrm{s}} - \mathbf{r}_{\mathrm{r}}) \qquad 3\text{- }20$$

where $\boldsymbol{\Omega_e}=(0,0,\ \omega_e)^T$ is the vector of the angular rotation rate of the Earth. By applying vector operations, the range correction can be simplified as

$$\begin{aligned} \delta R_{aberration} &= \frac{\omega_e}{c}[x^s(y^s - y_r) - y^s(x^s - x_r)] \\ &= \frac{\omega_e}{c}(x_r y^s - y_r x^s) \end{aligned} \qquad 3\text{- }21$$

The Sagnac effect correction given by Ashby and Spilker [1996] is in the form of

$$\delta R_{\text{Sagnac}} = \frac{2\mathbf{\Omega}_{\text{e}} \bullet \mathbf{A}}{\text{c}} \qquad \text{3- 22}$$

where $\boldsymbol{A}$ is the shading area of the triangles swept out by an arrow with its "tail" at the Earth centre and its "head" following the electromagnetic signal wave. Nevertheless, the two corrections are identical. A simple proof is given as follows

$$\begin{aligned}\delta R_{aberration} &= \frac{1}{c}(\mathbf{\Omega}_e \times \mathbf{r}^s) \bullet (\mathbf{r}^s - \mathbf{r}_r) = \frac{1}{c}\left(\mathbf{\Omega}_e \bullet (\mathbf{r}_r \times \mathbf{r}^s)\right) = \frac{2\mathbf{\Omega}_e}{c} \bullet \frac{\mathbf{r}_r \times \mathbf{r}^s}{2} = \frac{2\mathbf{\Omega}_e \bullet \mathbf{A}}{c} \\ &= \delta R_{Sagnac}\end{aligned} \qquad \text{3- 23}$$

The Sagnac effect is an error that must be accounted for due to the extra signal propagation arising from the Earth rotation in PPP. However, whether or not to apply this correction depends on how the Earth rotation is treated. If the satellite ECEF position at the signal transmission time has been rotated and the iteration scheme as described in § 3.1.4 has been applied, the Sagnac effect will be automatically eliminated.

Secondary Relativistic Effects

Several additional significant relativistic effects should be taken into consideration when analysing the Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR) and GPS measurements. These include the signal propagation delay effect on the geometric distance, phase wind-up, and the effects from other solar system bodies. Ashby [2003] classified the following relativistic errors as secondary.

(1) Signal Propagation Delay

The signal propagation delay for a user in the ECEF is given in the form of (ibid)

$$\delta R_{spd} = \frac{\Phi_0 \cdot \left\| \mathbf{r}_r^s \right\|}{c^2} - \frac{2GM}{c^2} \ln \frac{\left\| \mathbf{r}^s \right\| + \left\| \mathbf{r}_r \right\| + \left\| \mathbf{r}_r^s \right\|}{\left\| \mathbf{r}^s \right\| + \left\| \mathbf{r}_r \right\| - \left\| \mathbf{r}_r^s \right\|} \qquad 3\text{-}24$$

where Φ_0 is the effective gravitational potential on the geoid, which consists of both the Newtonian mass gravitational potential and the centrifugal potential due to Earth rotation. $\frac{\Phi_0}{c^2}$ =6.969290134×10^{-10} is the value adopted by the International Astronomical Union [IAU, 2000].

This correction is also referred to as the Shapiro time delay (when scaled into seconds). The second logarithm term on the right hand side has a maximum value of 18.7 millimetres according to Hofmann-Wellenhof et al. [2001], who refer to it as the "space-time curvature correction", which is significant only in an inertial system. For ECEF systems, the signal propagation delay tends to be small because the presence of the first term nearly cancels out the logarithm term. Thus, it is a secondary effect.

(2) Effect on Geometric Distance

A spatial curvature effect should be considered if requiring accuracy at the few millimetre level. The correction formula is given as (ibid)

$$\delta R_{sc} = \frac{GM}{c^2} \ln \frac{\left\| \mathbf{r}^s \right\|}{\left\| \mathbf{r}_r \right\|} \qquad 3\text{-}25$$

The value of this correction is approximately 4.43•*ln*(4.2)=6.3 millimetres. This is caused by the fact that coordinates and distances are different in different relativistic frames. It becomes important if one compares distances computed in one frame with those in another frame [Mueller and Seeber, 2004].

3.2.1.4 Satellite Attitude Effects

According to Kouba and Héroux [2001], there are two satellite related error sources that have not been covered yet, one is the satellite antenna offset, the other is the phase wind-up error.

Satellite Antenna Offset

Satellite antenna offset error is due to the offset between the satellite's mass centre and the phase centre of its antenna(s). The observations made by a user are relative to the antenna phase centre of a satellite, while the precise ephemeris provided by IGS is actually relative to the satellite mass centre because the force model used for satellite orbit determination refers to that mass centre. The IGS provides the offsets table, and a correction method can be found in the thesis of Witchayangkoon [2000].

Since a satellite position from the broadcast ephemeris is with respect to the phase centre of the satellite antenna, GPS users for real-time applications using the broadcast ephemeris are free of this error. However, attention should be paid when comparing the satellite positions from the precise ephemeris and the broadcast ephemeris.

Phase Wind-up Correction

The relative orientation of the satellite and receiver antenna has an effect on GPS measurements because the GPS signal is a right hand circularly polarised radio wave. A rotation of either receiver or satellite antenna around its bore axis would affect the carrier phase measurement, and the maximum effect may theoretically reach one cycle [Ashby, 2003]. As the apparent satellite movement is relatively slow, this effect remains even for a stationary GPS receiver. The phase wind-up correction formula is provided by Wu et al. [1993]. Ashby [2003] grouped the phase wind-up error as a

secondary relativistic effect and reported that this error had been experimentally detected with a GPS receiver spinning at a rotation rate of 8 circles per second.

3.2.2 Signal Propagation Errors

3.2.2.1 Ionosphere Errors

The ionosphere is a dispersive medium, that is, the effect of ionosphere on signal is a function of the signal frequency. More precisely, by neglecting the effects of higher order terms of the ionospheric refractivity, the ionospheric delay is proportional to the inverse of signal frequency squared. As has been stated earlier, the free electrons in the ionosphere affect the GPS signal propagation. The total free electrons along the signal path may be expressed by an integral of the local electron density N_e that is in units of electrons per cubic metres (el/m^3), as

$$STEC = \int_{path} N_e \cdot ds \qquad 3\text{-}26$$

where

- *STEC* stands for the Slant Total Electron Content;
- *ds* is an infinitesimal length along the signal path.

The ionospheric delay is given as [cf. Leick, 1995]

$$dI_{r,i}^{s} = \frac{40.30}{f_i^2} \cdot STEC \qquad 3\text{-}27$$

where i specifies the GPS signal frequency, 1 for $f_1 = 1575.42MHz$ and 2 for $f_2 = 1227.60MHz$ respectively.

Linear Combination of GPS Observables to Extract/Eliminate Ionospheric Error

With measurements from a dual-frequency GPS receiver, it is possible to retrieve the ionospheric delay, i.e. *STEC* which could be further used to establish the TEC model based on the observed real ionosphere data. From Eq.2-4 , P2-P1 relates to *STEC* as

$$\begin{aligned} P_{r,2}^{s}(t) - P_{r,1}^{s}(t) &= dI_{r,2}^{s} - dI_{r,1}^{s} + \varepsilon = \frac{40.30}{f_2^2} \cdot STEC - \frac{40.30}{f_1^2} \cdot STEC + \varepsilon \\ &= 40.30 \cdot \frac{f_1^2 - f_2^2}{f_1^2 f_2^2} \cdot STEC + \varepsilon \end{aligned} \qquad 3\text{-}28$$

STEC information may also be extracted from the more precise carrier phase observables through the following linear combination of carrier phase measurements

$$\begin{aligned} \lambda_1 \varphi_{r,1}^{s}(t) - \lambda_2 \varphi_{r,2}^{s}(t) &= dI_{r,2}^{s} - dI_{r,1}^{s} + N_{r,2}^{s}\lambda_2 - N_{r,1}^{s}\lambda_1 + \varepsilon \\ &= 40.30 \cdot \frac{f_1^2 - f_2^2}{f_1^2 f_2^2} \cdot STEC + N_{r,2}^{s}\lambda_2 - N_{r,1}^{s}\lambda_1 + \varepsilon \end{aligned} \qquad 3\text{-}29$$

Although the *STEC* from the carrier phase measurements may be more accurate than that from P2-P1, a constant ambiguity bias is present due to the integer ambiguity term in both the carrier phase observables.

The above linear combination is popularly referred to as the geometry-free combination, which contributes to the precise ionosphere modelling. As the ionosphere changes smoothly, in general the geometry-free carrier phase observation can be used to detect cycle slips very efficiently, provided there are no concurrent cycle slips at the same integer number on both L1 and L2.

Neglecting the higher order terms of ionospheric refractivity, the ionospheric delay may be practically eliminated by forming so-called ionosphere-free combinations applicable for both the code and the carrier phase measurements. The ionosphere-free code measurement is formed as

$$P^{s}_{r,ion\_free} = \frac{f_1^2}{f_1^2 - f_2^2} P^{s}_{r,1} - \frac{f_2^2}{f_1^2 - f_2^2} P^{s}_{r,2} \qquad 3\text{-}30$$

and the ionosphere-free carrier phase measurement on L1 is

$$\lambda_1 \varphi^{s}_{r,1-ion\_free} = \frac{f_1^2}{f_1^2 - f_2^2} \lambda_1 \varphi^{s}_{r,1} - \frac{f_1 f_2}{f_1^2 - f_2^2} \lambda_2 \varphi^{s}_{r,2} \qquad 3\text{-}31$$

The ionosphere-free carrier phase measurement has been widely used in GPS baseline processing over long distances and in PPP applications, however the integer property of the ambiguity has been lost and the measurement has a larger noise level than the original L1 and L2 carrier phase measurements. In fact, all the above linear combinations amplify the noise magnitudes.

The different sign but equal magnitude of the group delay and the phase advance makes it possible to eliminate the ionosphere by a sum of the code phase and the carrier phase measurements. However, this method suffers from the poor accuracy of code measurements and the presence of the integer ambiguity in the carrier phase measurements.

Standard Ionosphere Correction Model

For a single-frequency GPS user, a standard ionospheric correction algorithm is used by the GPS system and the eight parameters for this algorithm are broadcast in the navigation message. The algorithm is based on Klobuchar's single layer ionosphere model, which assumes that the ionospheric delay can be represented by projecting the zenith delay in the single layer to the signal line-of-sight direction using a slant factor. The algorithm is capable of correcting for approximately 50% of the ionospheric range error [Klobuchar, 1996; ARINC, 2000].

The Klobuchar model uses a constant plus half-cosine waveform to represent the diurnal variation of TEC for the simplified ionospheric layer. The constant is presumed to be 5.0×10^{-9}, and the peak TEC to appear at 14:00h local time. Four of the eight model parameters are used to represent the amplitude of the cosine wave; the other four model the period of the cosine wave. The algorithm requires receiver knowledge of the geodetic latitude ϕ_u, longitude λ_u, and the azimuth A and elevation E to each satellite. They are used to calculate the sub-ionospheric point (ionospheric pierce point) and the slant projection factor. The algorithm is given in the Appendix.

3.2.2.2 Troposphere Errors

The troposphere is a neutral, non-dispersive medium. The delay of a GPS signal passing through is not dependent upon frequency, which means that there is no distinction between the carrier phase and code phase measurements, nor is it possible to eliminate it through combination of dual-frequency observations. The troposphere error is one of the main accuracy-limiting factors in GPS positioning, especially for a standalone user. Thus appropriate modelling is required. Simsky and Boon [2003] showed that the change of troposphere delay over time is the most important parameter to model in order to achieve sub-centimetre per second velocity accuracy. In this section, a thorough literature review of the troposphere modelling is provided. For more comprehensive discussions, the reader is referred to Spilker [1996c], Mendes [1999], and Hofmann-Wellenhof et al. [2001].

For modelling purposes, the troposphere delay can be simply considered as having a hydrostatic (dry) component and a non-hydrostatic (wet) component. The hydrostatic component depends on the locality, season and altitude; and is relative stable. On the other hand, the non-hydrostatic component varies with the local weather conditions and changes quickly. As a result it is very hard to model the wet troposphere component. Fortunately, the wet component only contributes approximately 10% of

the tropospheric delay; hence 90% of the delay is due to the more predictable hydrostatic air. Figure 3-1, which is modified from Hofmann-Wellenhof et al. [2001], is given to depict the thickness of the polytropic layers of the troposphere as well as the bent signal path.

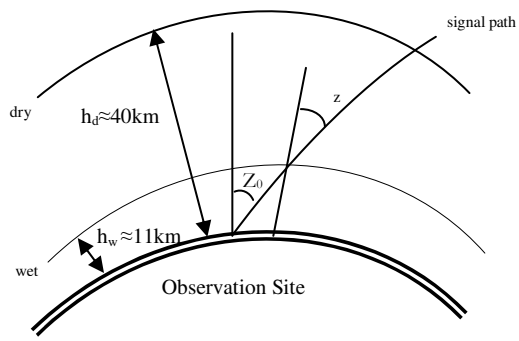

Figure 3- 1: Troposphere layers and the signal path, adapted from [Hofmann-Wellenhof et al., 2001]

The tropospheric path delay experienced by a GPS signal is an integral of the tropospheric refractivity N_T along the signal path

$$\begin{aligned} dT &= 10^{-6} \int_{path} N_T \cdot ds \\ &= 10^{-6} \int_{dry\_path} N_d \cdot ds + 10^{-6} \int_{wet\_path} N_w \cdot ds \end{aligned} \qquad 3\text{- }32$$

where N_T has been partitioned into a hydrostatic component N_d and a wet component N_w. The N_d changes with height, temperature and pressure; and the N_w changes with height, temperature and partial pressure of water vapour. The difficulty lies in that it is impractical to measure these parameters along a real signal path, and to precisely determine the troposphere delay through the above integration. However, the refractivity at the Earth's surface is well known and can be expressed in the following forms

$$N_d^0 = c_1 \cdot \frac{P_0}{T_0}$$
$$N_w^0 = c_2 \cdot \frac{e_0}{T_0} + c_3 \frac{e_0}{T_0^2} \qquad 3\text{- }33$$

where :

- P_0 is the ground atmospheric pressure in millibars (mb);
- T_0 represents the ground temperature in Kelvin (K);
- e_0 is the partial pressure of water vapour in mb;
- c_i are empirical coefficients which may change with location.

The knowledge accumulated through vertical radiosondes by atmospheric researchers enables us to relate vertical refractivity to the surface refractivity through empirical formulae. With the zenith refractivity model, analogous to the ionosphere modelling, slant factors can be introduced to represent the tropospheric delay as

$$dT \approx 10^{-6} \cdot \int_{h_0}^{h_d} N_d^z(h) \cdot dh \cdot m_d(Z) + 10^{-6} \cdot \int_{h_0}^{h_w} N_w^z(h) \cdot dh \cdot m_w(Z) \qquad 3\text{- }34$$

where:

- N_d^z is the dry zenith refractivity, a function of the altitude and surface meteorological parameters;
- N_d^z is the wet zenith refractivity, changing with altitude and surface meteorological data;
- $m_d(Z)$ and $m_w(Z)$ are mapping functions (slant factors) to scale the zenith delays to the signal path.

With such an approximation it is feasible to complete the integration along the vertical tropospheric path and therefore model the troposphere delay using surface meteorological data. Then the tropospheric delay may be written in the following succinct form

$$dT = dT_d^z \cdot m_d(Z) + dT_w^z \cdot m_w(Z) \qquad 3\text{-}35$$

Hopfield Model

Hopfield [1969] proposed a representation of the zenith dry refractivity empirically based on global meteorological observations. The vertical dry refractivity in height h is expressed in terms of surface dry refractivity N_d^0 by a quartic form

$$N_d^z(h) = N_d^0 \cdot \left(1 - \frac{h}{h_d}\right)^4 \qquad 3\text{-}36$$

where h_d is assumed to be 40.136km. The Hopfield model for wet refractivity assumes a similar relationship

$$N_w^z(h) = N_w^0 \cdot \left(1 - \frac{h}{h_w}\right)^4 \qquad 3\text{-}37$$

where the mean value of h_w is set 11.0km. An Eq.3-33-like Hopfield model is given in Misra and Enge [2001] with

$$\begin{aligned} \mathrm{dT_d^z} &= 77.6\times10^{-6}\,\frac{\mathrm{P_0}}{\mathrm{T_0}}\cdot\frac{\mathrm{h_d}}{5} \\ \mathrm{dT_w^z} &= 0.373\,\frac{\mathrm{e_0}}{\mathrm{T_0}}\cdot\frac{\mathrm{h_w}}{5} \end{aligned} \qquad 3\text{-}38$$

There are many variations of the Hopfield model that can be obtained by choosing different h_d and h_w values, varying from heights above sea level to lengths of position vectors, and using different mapping functions as well. The modified Hopfield model [Goad and Goodman, 1974] and the simplified Hopfield model [Wells, 1974] are typical examples and have been implemented as optional models in the Bernese software [Rothacher et al., 1996]

Saastamoinen Model

The Saastamoinen model has been widely used in geodesy. The Bernese software, for example, takes the Saastamoinen model as the *a priori* model to account for tropospheric delays (ibid). This model is based on the laws associated with ideal gas refraction. Saastamoinen [1972; 1973] presented two tropospheric delay models, one is the standard tropospheric model, and the other is a more precise model. The standard tropospheric delay model is for radio frequency ranging to satellites with elevation angles greater than 10 degrees, and has the following form

$$dT = \frac{0.002277}{\cos Z}[P_0 + (\frac{1255}{T_0} + 0.05)e - \tan^2 Z] \qquad 3\text{-}39$$

Similar to the Hopfield model, the Saastamoinen model has been refined to account for signal bending, height and even the location of the observation site by introducing more terms into the above equation [Spilker, 1996c], for example

$$dT = \frac{0.002277}{\cos Z}(1+D)[P_0 + (\frac{1255}{T_0} + 0.05)e - B\tan^2 Z] + \delta R \qquad 3\text{-}40$$

where:

- $D=0.0026\ cos2\varphi+0.00028h$, is a term that accounts for the variations of the receiver in latitude and height (h is in unit of kilometres);

- B is a function of the height in mb;
- δR term depends on the height and elevation of a radio signal source.

The B and δR variables can be obtained by interpolation of tables given in Spilker [1996c], Hofmann-Wellenhof et al. [2001].

Neglecting δR, the refined Saastamoinen model can be easily partitioned into dry and wet zenith delays like the form of Eq.3-33, see [Misra and Enge, 2001, p.147],

$$\begin{aligned} dT_d^z &= 0.002277(1+0.0026\cos 2\varphi + 2.8\times 10^{-4} h)\cdot(P_0 - B\tan^2 Z) \\ dT_w^z &= 0.002277 e_0 (\frac{1255}{T_0} + 0.05) \end{aligned} \qquad 3\text{-}41$$

In the Bernese software, only the B correction term in the refined equation has been implemented [Rothacher et al., 1996].

UNB Models

Based on the precise Saastamoinen model, researchers at the University of New Brunswick developed a series of tropospheric delay models designated as UNBx models. Rather than just using the three ground meteorological parameters of T_0, e_0 and P_0, another two parameters, the temperature lapse rate β and water vapour decreasing rate ,λ are introduced to model the changes with height along the vertical tropospheric profile.

The temperature lapse rate β is assumed linear with height h

$$T(h) = T_0 - \beta \cdot h \qquad 3\text{-}42$$

while the water vapour lapse rate λ is in the power form of

$$e(h) = e_0 (\frac{P}{P_0})^{\lambda+1} \qquad 3\text{-}43$$

Saastamoinen's hydrostatic zenith delay is in the following form [Davis et al., 1985]

$$dT_d^z = \frac{10^{-6} k_1 \cdot R_d}{g_m} \cdot P \qquad 3\text{-}44$$

and the Saastamoinen wet zenith delay is in the form given by Thayer [1974], Askne and Nordius [1987]

$$dT_w^z = \frac{10^{-6}(T_m \cdot k_2^{'} + k_3)R_d}{g_m(\lambda+1) - \beta \cdot R_d} \cdot \frac{e}{T} \qquad 3\text{-}45$$

In the above equations

- R_d is the gas constant for dry air;
- g_m is the gravitational acceleration at the centroid of the vertical tropospheric profile;
- T_m represents the mean temperature of the water vapour, $T_m = (1 - \frac{\beta \cdot R_d}{g_m(\lambda+1)})T$;
- k_1, k'_2, k_3 are refractivity constants [Thayer, 1974].

With these five parameters and the equations 3-41 to 3-44, the hydrostatic zenith delay can be given in terms of surface meteorological data as

$$dT_d^z = \frac{10^{-6} k_1 \cdot R_d}{g_m} \cdot \left(1 - \frac{\beta h}{T_0}\right)^{\frac{g_m}{R_d \cdot \beta}} \cdot P_0 \equiv \Gamma_d(T_0, \beta) \cdot P_0 \qquad 3\text{-}46$$

and the zenith wet delay is given by

$$dT_w^z = \frac{10^{-6}(T_m k_2^{'} + k_3)R_d}{g_m(\lambda+1) - \beta \cdot R_d} \cdot \left(1 - \frac{\beta \cdot h}{T_0}\right)^{\frac{(\lambda+1)g_m}{R_d \cdot \beta} - 1} \cdot \frac{e_0}{T_0} \equiv \Gamma_w(T_0, \beta, \lambda) \cdot \frac{e_0}{T_0} \qquad 3\text{-}47$$

Thorough reviews of the development of UNB models can be found in a series of publications [Collins and Langley, 1996; Collins et al., 1996; Collins and Langley, 1997], where tables associated with the models are also provided. The UNB1 adopts a set of 5 parameters with their global mean values; in the UNB2 model, the 5 parameters are provided by a table that reflects their changes in latitude with an interval of 10˚. The improvement of UNB3 model is that the annual variation of the 5 parameters have been accounted for using a sinusoidal function of day-of-the-year, which is taken from the concept of Niell's mapping function [Niell, 1996]. In UNB3 model the mean values of the 5 parameters are given at 15˚ intervals in latitude, accompanied by their variation amplitudes. This allows a user to calculate each of the 5 parameters from linear interpolations by specifying the latitude and day-of-the-year

$$\begin{aligned}\xi(\varphi, DoY) &= \xi_m(\varphi_i) + [\xi_m(\varphi_{i+1}) - \xi_m(\varphi_i)] \cdot \frac{\varphi - \varphi_i}{\varphi_{i+1} - \varphi_i} \\ &\quad - \{\xi_m(\varphi_i) + [\xi_m(\varphi_{i+1}) - \xi_m(\varphi_i)] \cdot \frac{\varphi - \varphi_i}{\varphi_{i+1} - \varphi_i}\} \cdot \cos\left(\frac{2\pi(DoY - 28)}{365.25}\right)\end{aligned}$$

3- 48

where:

- ξ represents a required parameter among the 5 parameters;
- φ, and DoY are the user latitude and the time in day-of-the-year;
- $\xi_m(\varphi_i)$ is the mean value of the parameter in the corresponding nearest latitude specified in the UNB3 table, $\xi_m(\varphi_{i+1})$ is the mean value of the parameter in the next nearest latitude specified in the UNB3 table;

- *28* indicates that the peak variation appears on January 28.

As temperature lapse rate has little variation above a certain troposphere boundary layer across all latitudes, a constant temperature lapse rate of 6.5K/km is adopted, and zero annual amplitude is assumed. This is the distinction between the UNB4 and UNB3 models. Last notes for the UNB models are that the unit of height h is in kilometres rather than metres, and it might be proper to set the peak variation time in the southern hemisphere to July 28, since the climate is contrary to that in the northern hemisphere.

3.2.2.2.4 Mapping Functions

There are numerous mapping functions to scale the modelled zenith delay to a signal path delay. The elevation angle E which relates to the zenith angle Z with $E=90^\circ-Z$ is popularly used in the literature of mapping functions. The simplest mapping function model is to apply $m(E)=1/\sin E$ for both the zenith hydrostatic and water vapour components, which assumes that the Earth is planar, and thus is not suitable for low elevation satellites. This mapping function may be improved using [Misra and Enge, 2001]

$$m(E) = \frac{1}{\sqrt{1-(\cos E/1.001)^2}} \qquad 3\text{-}49$$

Most of geodetic-quality mapping functions use a continued fraction form. This functional form for the mapping functions was first proposed by Marini [1972] and further developed by Chao [1974] , Davis et al. [1985], Ifadis [2000] , Herring [1992] and Niell [1996; 2000]. Marini's original form is given as

$$m(E) = \cfrac{1}{\sin E + \cfrac{a}{\sin E + \cfrac{b}{\sin E + \cfrac{c}{\sin E + \cdots}}}} \qquad 3\text{-}50$$

where *a, b, c*, etc. are the mapping function parameters to be determined. They may be constants or functions of other variables such as the latitude, height, surface temperature and pressure, and the day-of-the-year. Herring [1992] specified Marini's formula with three constants and normalised this function to be unity at the zenith

$$m(E) = \frac{1 + \cfrac{a}{1 + \cfrac{b}{1 + c}}}{\sin E + \cfrac{a}{\sin E + \cfrac{b}{\sin E + c}}} \qquad 3\text{-}51$$

However, accurate mapping functions distinguish the difference in dry and wet delays. Chao [1974] adopted Marini's form but separated mapping functions for the hydrostatic and dry components. Only the first two terms are used in Chao's functions, with a replacement of the second *sinE* by *tanE*. The specific dry and wet mapping functions by Chao are as follows

$$\begin{aligned} m_d(E) &= \frac{1}{\sin E + \cfrac{0.00143}{\tan E + 0.0445}} \\ m_w(E) &= \frac{1}{\sin E + \cfrac{0.00035}{\tan E + 0.017}} \end{aligned} \qquad 3\text{-}52$$

Chao's mapping functions can be used down to 10° of elevation. The merits of Chao's formulae are simple in forms, and independent on location and user height. As no

meteorological data are required, these mapping functions are suitable for kinematic applications.

Niell [1996] kept the form of the Herring's formula, but introduced an extra height correction term. He assumed that the elevation dependence was a function of the site position (latitude and height above sea level) and day-of-the-year. The proposed mapping functions are

$$m_d(E) = \frac{1 + \dfrac{a_d}{1 + \dfrac{b_d}{1 + c_d}}}{\sin E + \dfrac{a_d}{\sin E + \dfrac{b_d}{\sin E + c_d}}} + M_r \cdot h_r \qquad 3\text{-}53$$

$$M_r = \frac{1}{\sin E} - \frac{1 + \dfrac{a_h}{1 + \dfrac{b_h}{1 + c_h}}}{\sin E + \dfrac{a_h}{\sin E + \dfrac{b_h}{\sin E + c_h}}} \qquad 3\text{-}54$$

where:

- h_s is the receiver height above sea level in unit of kilometres;
- M_r is a mapping function to account for station height;
- a_h, b_h, c_h are constants of 2.53×10^{-5}, 5.49×10^{-3}, 1.14×10^{-3} respectively;
- a_d, b_d, c_d are Niell's dry parameters, which change with latitude and observation time in day-of-the-year.

As has been introduced in the UNB3 model, tabulated mean values of a_d, b_d, c_d are given at an interval of 15°, and the temporal variation amplitudes are provided. Linear

interpolation is required for arbitrary latitude of a user using the same scheme as in the UNB3 model

$$\xi(\varphi, DoY) = \xi_m(\varphi_i) + [\xi_m(\varphi_{i+1}) - \xi_m(\varphi_i)] \cdot \frac{\varphi - \varphi_i}{\varphi_{i+1} - \varphi_i} - \{\xi_m(\varphi_i) + [\xi_m(\varphi_{i+1}) - \xi_m(\varphi_i)] \cdot \frac{\varphi - \varphi_i}{\varphi_{i+1} - \varphi_i}\} \cdot \cos\left(\frac{2\pi(DoY - 28)}{365.25}\right) \quad 3\text{-}55$$

Half-year days should be added to *DoY* (day-of-the-year) if a location is in the southern hemisphere to account for the weather difference. Neill's wet zenith mapping function takes the form of Herring (see Eq.3-40). The wet mapping parameters of *a, b,* and *c* only depend on site latitude, and no annual variations are assumed. A user's mapping parameters can be obtained by a simple linear interpolation from the given table.

These precise mapping functions are developed for geodetic applications such as VLBI, SLR, and precise permanent GPS tracking. The accuracy is at the millimetre level even down to elevation of 3~5˚, which is suitable for extracting water vapour information contained in low elevation GPS signals. "GPS meteorology" for water vapour monitoring has therefore attracted much research attention.

3.2.2.3 Multipath Errors

Multipath error occurs when a GPS signal reaches an antenna via more than one path. The reflected signal travels a longer distance than the direct signal and is thus delayed. The errors of the multipath in GPS signals, both code and carrier phase, depend upon the reflected signal strength and the way that the signal is reflected. Although it affects both the carrier and code measurements, the magnitudes of the error differ significantly. Signals from low satellite elevation manifest greater multipath errors than signals from high elevation, and therefore a simple mitigation method is to raise the allowable elevation cut-off angle.

The multipath error in code measurements varies from ±1m in a benign environment to ±10~20m, and it may even grow to about 100m or more in the vicinity of buildings [Hofmann-Wellenhof et al., 2001]. However, the typical range of multipath error in code pseudorange measurements is ±1~5m [Misra and Enge, 2001] which is still two orders of magnitude higher than the multipath error in the carrier phase measurements (±1~5cm).

A mathematical scheme to illustrate the multipath effect on the carrier phase is given in Hofmann-Wellenhof et al. [2001]. The interfered receiving signal contains a direct line-of-sight part, and a reflected signal part with a phase shift $\Delta\varphi$ and an amplitude attenuation β

$$signal_{received} = A\cos\varphi + \beta \cdot A\cos(\varphi + \Delta\varphi) \qquad 3\text{-}56$$

Through trigonometric relationships, the errors in the carrier phase due to the multipath $\delta\varphi$ can be determined

$$\delta\varphi = \arctan\left(\frac{\beta\sin\Delta\varphi}{1+\beta\cdot\cos\Delta\varphi}\right) \qquad 3\text{-}57$$

The strongest possible reflection is defined by $\beta=1$. From Eq.3-57 it is clear that the error in the carrier phase measurement due to the multipath does not exceed a quarter cycle. For the GPS L1 carrier phase signal that has a wavelength of about 20cm, this is equivalent to ±5cm. It also indicates that the multipath effect has a periodic property, thus, a longer observation time would benefit the multipath reduction in static GPS applications.

More accurate modelling of multipath is based on Eq.3-57, but taking all incoming signals into consideration [Leick, 1995]

$$\delta\varphi = \arctan\left(\frac{\sum_{i=1}^{n}\beta_i \sin\Delta\varphi_i}{1+\sum_{i=1}^{n}\beta_i \cos\Delta\varphi_i}\right) \qquad 3\text{-}58$$

And a multipath estimation approach is also described by Blewitt [1998].

Multipath mitigation can be achieved by locating the antenna away from reflectors, raising the height of the antenna, applying a ground plate to the antenna to mask the reflecting signal from lower angles, adopting an expensive choke ring antenna if it is feasible, or all of the above. The antenna and receiver designs from manufacturers play key roles in multipath error mitigation. Techniques such as the narrow correlator spacing, the strobe correlator multipath rejection, and the multipath estimating delay lock loop may reduce most of the error [Hofmann-Wellenhof et al., 2001]. The phase signal to noise ratio (SNR) has also been used as an indicator of the magnitude of multipath effects, and contributes to a better multipath estimation. References for further study are in Georgiadou and Kleusberg [1988], Braash [1996], and Comp and Axelrad [1996].

Multipath effects in kinematic applications are more severe than in static applications because the motion of a receiver antenna causes large fluctuations in the signals reflected from the surroundings [Grewal et al., 2001]. Therefore, special treatment of multipath is always required in order to improve the positioning accuracy.

3.2.3 Receiver Dependent Errors

3.2.3.1 Receiver Clock Error

Subject to the cost of manufacturing, a GPS receiver clock is normally made from a quartz crystal oscillator, which is different from the GPS time by a clock offset, and this offset tends to drift. The receiver clock offset at the instant of measurement has

the same effect on all tracked satellite observables, thus it becomes the fourth unknown in addition to the three coordinate components for GPS point positioning.

The receiver clock error may be estimated as the fourth unknown when more than four satellites are observed; it may also be eliminated by double-differencing or between-satellite single-differencing. However, the residual of the clock error may remain. Unlike satellite clocks, the receiver clock information is not readily known and consequently a good estimation benefits almost all GPS applications. In this research, however, the receiver clock drifts are of particular interest.

There are two schemes for manufacturers to deal with the receiver clock drift. One is to steer the receiver clock continuously, an algorithm that can be implemented in the receiver's embedded software. The clock drift is adjusted epoch by epoch. The other approach is to allow the drift to reach a certain threshold (typically 1ms), and then reset the clock with a jump to a new epoch. The former is referred to as receiver clock "steering", and the latter is referred to as receiver clock "reset". The receiver clock reset has a detrimental effect on some applications in terms of signal sampling since many discrete signal-processing (DSP) techniques are based on equal spaced time intervals. Whenever there is a clock reset, there is a change of the time intervals of the sampled signals.

There are many approaches to model the receiver clock error. The simplest way is to model the receiver clock in a similar way to the correction of the satellite clock errors, that is, to use a polynomial to represent the error of the receiver clock by

$$dt_r(t) = a_0 + a_1(t - t_{ref}) + a_2(t - t_{ref})^2 \qquad 3\text{-}59$$

where:

- a_0, a_1, a_2 are the polynomial coefficients, some times referred to as the clock bias, clock rate, and clock drift respectively;

- t_{ref} is the reference time; the observation beginning time may be chosen as the reference time.

Another way to model the clock error is a variation of Eq.3-59, given as

$$dt_r(t) = a_0 + a_1(t - t_{ref}) + 0.5a_2(t - t_{ref})^2 \qquad 3\text{-}60$$

where a dynamic modelling concept is used by introducing the change-rate and acceleration of the receiver clock.

An alternative to model the clock error utilises the physical character of a crystal clock. The clock error comes from the errors in the frequency and phase of a crystal oscillator, which are assumed to be random walk processes over a reasonable time span [Brown and Hwang, 1992; Farrell and Barth, 1999]. A two-state Kalman filter is used and a discrete dynamic state equation is given as

$$\begin{bmatrix} dt_r \\ d\dot{t}_r \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} t_r(t) \\ \dot{t}_r(t) \end{bmatrix} + \begin{bmatrix} \omega_f \\ \omega_g \end{bmatrix} \qquad 3\text{-}61$$

where ω_f is the frequency noise and ω_g is the phase noise. The noise behaviours can be modelled by the following driving noise matrix

$$Q = \begin{bmatrix} S_f \Delta t + \frac{1}{3} \cdot S_g \Delta t^3 & \frac{1}{2} S_g \Delta t^2 \\ \frac{1}{2} S_g \Delta t^2 & S_g \Delta t \end{bmatrix} \qquad 3\text{-}62$$

where S_f and S_g are the power spectral densities of the frequency and phase error respectively, which can be determined via the Allan variance parameters. Typical Allan variance parameters can be found in Brown and Hwang [1992].

3.2.3.2 Receiver Site Displacement

Precisely speaking, a receiver position changes with time, even though it is static on the ground. This is because in a global sense, the station undergoes real or apparent periodic movement due to tectonic motions, tidal effects, and loadings from the atmosphere and oceans. The movements may reach a few decimetres level, and thus must be modelled by adding the site displacement correction terms.

Kouba [2001], Kouba and Héroux [2003] list the displacement terms for effects with magnitudes more than ±1cm:

- Solid Earth tides: are the deformations of the solid Earth as it rotates within the gravitational fields of the Sun and the Moon. Similar to ocean tides, the Earth deforms because it has a certain degree of elasticity. The solid Earth tidal correction can be mathematically expressed in terms of: 1) the gravitational parameters of the Earth, the Moon and the Sun; 2) the position vectors of the station, the Moon and the Sun. 3) Spherical harmonics of degree and order (n,m) characterised by the Love number h_{nm} and the Shida number l_{mn} . The magnitudes may reach about 30cm in the radical and 5cm in the horizontal directions. Consequently neglecting this correction may result in systematic position errors up to 12.5cm in the radial and 5 cm in the north directions.
- Ocean loading: the Earth's surface dips under the load of the ocean tide and thus the magnitude of this deformation is dominated by diurnal and semidiurnal periods of the tide. Ocean loading may exist everywhere on the Earth but has significant effects on costal areas, with the same temporal frequencies as the tide. However, for single epoch positioning at ±5cm level or static positioning longer than 24 hours, or inland surveying, the effects of ocean loading can be neglected.

- Rotational deformation due to polar motion: Similar to the solid Earth tide effect due to attractions from the Sun and the Moon, the Earth polar motion (changes of the Earth rotation axis) results in station positional deformation with periodical characteristics due to minute changes in the Earth centrifugal potential. Analogous spherical harmonic expansions are used with second degree Love and Shida numbers. The polar motion may reach up to 0.8 arcsecond, and the corresponding site displacements may reach about 0.7 cm in horizontal directions, and over 2.5cm in vertical.
- Earth rotation parameters (ERP): the pole position (*Xp*, *Yp)* and *UT1-UTC*, along with the conventions for sidereal time, precession and nutation facilitate accurate transformations between terrestrial and inertial reference frames. Users who utilise software working in an inertial system should apply this correction.

The correction models and equations for receiver site displacement may also be found in Kouba and Héroux [2001; 2003], Kouba [2003]. Another good reference is in Bock [1998].

3.3 Summary

The principles of absolute GPS positioning or PPP are introduced in this chapter, by further formation of the observation equations, performing the least squares adjustment and evaluating the position estimates. GPS errors and modelling methods are elaborated from the point of views of PPP.

Since the Doppler shift is the first-order derivative of the carrier phase, all errors in the carrier phase measurements would have some effects on the measured Doppler shifts. Amongst the various error sources in the range measurements, the relativistic effects and the troposphere errors are comprehensively discussed . This is because they are the two major error sources limiting the accuracy of precise velocity determination using the GPS.

Chapter Four

PRECISE VELOCITY AND ACCELERATION DETERMINATION USING GPS

Velocity is the first derivative of position with respect to time, and acceleration is the second-order derivative of the position with respect to time. Therefore, one may obtain the ground velocity and acceleration of a GPS receiver directly through position differentiation. However, the accuracy of the resultant velocity and acceleration is degraded due to the poor GPS positional accuracy in standalone mode under SPS. The objective of this chapter is to introduce an alternative method of determining precise velocity and acceleration using the GPS Doppler shift measurements.

A GPS receiver-satellite pair is in the Earth's gravity field and GPS signals from the satellite travel at the speed of light, hence both Einstein's special and general relativity theories are to be considered in the Doppler effects of GPS.

This chapter establishes the relationship between the measured Doppler shift and the user's ground velocity by considering both the special and general relativistic effects. A unified Doppler shift model is developed, which accommodates both the classical Doppler effect and the relativistic Doppler effect. A highly accurate GPS Doppler shift observation equation is presented with all major GPS error sources considered. The principle of acceleration determination using GPS is also discussed. This method uses a virtual GPS observable, i.e. the change-rate of the Doppler shift or simply the

Doppler shift rate. The methods to obtain the Doppler shift rate "observable" will be discussed in Chapter Seven.

4.1 Doppler Effect and Velocity Determination

Designed as a PVT system, GPS is capable of providing instantaneous velocity for a moving user [Hofmann-Wellenhof et al., 2001]. With SA turned off, ±0.2m/sec per axis accuracy (95%) is guaranteed by the GPS system for standalone GPS users [DoD, 1996]. To date, such an accuracy level of velocity can be easily achieved by code-only low-cost GPS receivers [Zhang et al., 2003a]. Ground velocity accuracy at the sub-centimetres per second level, has also been shown to be achievable in standalone mode [Van Graas and Soloview, 2003; Serrano et al., 2004].

Compared to the positional accuracy specifications under SPS, it is obvious that the above velocity accuracy is not obtained through direct differentiation of the GPS determined positions with respect to time.

For high accuracy velocity determination, GPS satellites are assumed to be known signal sources from which a GPS receiver senses the Doppler frequency shifts due to the relative motion between the observed satellites and the receiver. The velocity of the GPS receiver is resolved if Doppler shifts from at least four GPS satellites have been measured.

4.1.1 Classical Doppler Effect

The Doppler effect is the apparent change in frequency of a wave that is received by an observer moving with respect to the source of the emitted wave [Wikipedia, 2004]. The Doppler effect was named in honour of Christian Doppler, an Austrian mathematician who first proposed the idea in 1842 when he observed the coloured lights of a double star named Albireo (β Cygni). He proposed that the differences in

colour were due to the motion of the stars, since the frequency of the light could be increasing from an approaching star and decreasing from a receding star.

Doppler's conjecture was proven by an experiment designed by C. H. D. Buijs-Ballot in 1845 [Calvert, 2004]. He tested for sound waves by standing next to a rail line and listening to a rail-car full of musicians as they approached him and after they passed him. He confirmed that the frequency of the sound was higher as the sound source approached him, and lower as the sound source receded from him.

The changes of sound frequency in the air fall into the category of the classical Doppler effect wherein the velocity of the sound wave is much less than the speed of light. In the classical Doppler Effect, the following formula holds [Wikipedia, 2004]

$$f_r = f_s \cdot (1 + \frac{\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s}{c_0}) \qquad 4\text{-}1$$

where:

- A dot over a vector represents the first derivative of the vector with respect to time;
- s denotes the transmitter;
- r represents the receiver;
- f_r is the received frequency from the receiver;
- f_s is the original frequency of the sound wave (transmitter or sender);
- c_0 represents the speed of the wave (sound) in the air;
- $\mathbf{n}_r^s$ is the receiver-transmitter line-of-sight unit vector;
- $\mathbf{r}_r^s$ is the receiver-transmitter vector;
- $\dot{\mathbf{r}}_r^s$ is the relative velocity vector between the signal transmitter and the receiver.

The Doppler shift D can be related to the relative velocity by rewriting Eq.4-1 into the following form

$$D \equiv (f_r - f_s) = f_s \frac{\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s}{c_0} = \frac{1}{\lambda_s} \mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s \qquad 4\text{-}2$$

where $\lambda_s = \frac{c_0}{f}$ is the wavelength of the wave. Note that the Doppler shift D is defined as having a positive sign when the receiver and the transmitter approach each other and a negative sign when they depart from each other. This equation can be rewritten into the following form

$$D\lambda_s = \mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s \qquad 4\text{-}3$$

4.1.2 Relativistic Doppler Effect

4.1.2.1 Special Relativity Only

The GPS signal propagation is quite different from sound propagation in the air since GPS satellites orbit the Earth at high speed (approximately 3.8km/s), transmitting electromagnetic wave signals that travel at the speed of light. In this case Einstein's relativity theory applies. This is the relativistic Doppler effect, wherein the received signal frequency is [Wells et al., 1987; Seeber, 1993]

$$f_r = f_s \cdot \frac{(1 - \frac{v}{c} \cdot \cos\theta)}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad 4\text{-}4$$

where:

- v is the magnitude of the satellite velocity;

- c is the speed of light in vacuum;
- θ is the angle between the satellite velocity vector and the receiver to satellite line-of-sight vector.

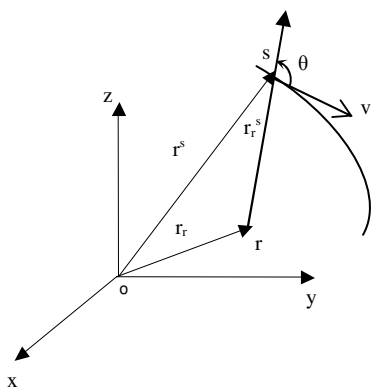

Figure 4- 1: A sketch of the receiver to satellite pair in space

A further explanation of the relativistic Doppler effect expressed by Eq.4-4 is given in Fig.4-1. It can be seen that the satellite velocity $\boldsymbol{v}$ may be mapped into the receiver to satellite range-rate by multiplying $cos\theta$

$$\dot{\mathbf{r}}_r^s = \mathbf{v} \cdot \cos\theta \qquad 4\text{-}5$$

Therefore, Eq.4-4 can be rewritten in the vector form by replacing $v\,cos\,\theta$ with $\boldsymbol{n}_r^s \bullet \dot{\boldsymbol{r}}_r^s$

$$f_r = f_s \left(\frac{1+\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s}{\sqrt{1-\frac{v^2}{c^2}}}\right) \qquad 4\text{-}6$$

The denominator in Eq.4-6 can be expanded using the binomial series

$$f_r \approx f_s \cdot \left(1+\frac{\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s}{c}\right) \cdot \left(1+\frac{v^2}{2c^2}+\frac{3v^4}{8c^4}+\cdots\right) \qquad 4\text{-}7$$

When the higher order terms in Eq.4-7 are neglected, the Doppler shift becomes

$$D \equiv f_r - f_s = \frac{\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s}{c} f_s = \frac{1}{\lambda} \cdot \mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s \qquad 4\text{-}8$$

where once again D is positive when satellite s is approaching and negative when the satellite is departing. Eq.4-8 gives the relationship between the Doppler shift and the receiver-satellite line-of-sight range-rate. This form of expression for Doppler shift D has been widely used in velocity determination from GPS measurements.

Equation 4-6 is the relativistic Doppler effect equation under special relativity. Since the satellite speed $v<<c$, the denominator in Eq.4-6 is approximately unity and a rearrangement of term $\boldsymbol{n}_{\mathrm{r}}^{\mathrm{s}} \bullet \dot{\boldsymbol{r}}_{\mathrm{r}}^{\mathrm{s}}$ with the substitution of $-\mathbf{v}\cos\theta$ gives the classical Doppler Effect in the form of Eq.4-2, as shown in Eq.4-8. Hence, the relativistic Doppler effect is in a more generic form.

Among the neglected terms in Eq.4-6, only the second-order v/c term may have some numerical effect on the frequency reception. Seeber [1993] terms it the "transversal" Doppler effect, whereas Ashby and Spilker [1996] refer to it as the "second-order Doppler effect". The magnitude of this effect may be evaluated from

$$\Delta f_r \approx f_s (1 + \frac{\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s}{c}) \cdot \frac{v^2}{2c^2} \approx \frac{1}{2}\frac{v^2}{c^2} \cdot f_s \qquad 4\text{-}9$$

Since GPS signals have very high frequencies (L1= 1575.42 MHz, L2=1227.60MHz) and as a result Δf_r is greater than 0.1Hz, which is equivalent to over 2cm/s error in the range-rate. Therefore, the second-order Doppler effect should be taken into consideration for precise applications.

4.1.2.2 GPS Relativistic Doppler Effect

The range vector from satellite s to receiver r is

$$\rho = \left\| \mathbf{r}^s - \mathbf{r}_r \right\| = \mathbf{n}_r^s \bullet \mathbf{r}_r^s \qquad 4\text{-}10$$

From the above equation, the difference between the coordinate time of reception and the coordinate time of transmission is

$$t_r - t^s = \frac{\rho}{c} \qquad \text{4- 11}$$

Differentiating Eq.4-11 with respect to coordinate time

$$\begin{aligned} 1 - \frac{dt^s}{dt_r} &= \frac{1}{c}\left(\frac{\partial\rho}{\partial t_r} + \frac{\partial\rho}{\partial t^s}\frac{dt^s}{dt_r}\right) \\ &= \frac{1}{c}\left[\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r - \mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s \frac{dt^s}{dt_r}\right] \end{aligned} \qquad \text{4- 12}$$

This can be formulated as

$$\frac{dt^s}{dt_r} = \frac{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r}{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s} \qquad \text{4- 13}$$

The ratio of the received frequency f_r and the transmitted frequency f_s can be expressed as

$$\frac{f_r}{f_s} = \frac{d\tau^s}{d\tau_r} = \frac{\frac{dt_r}{d\tau_r}}{\frac{dt^s}{d\tau^s}}\frac{dt^s}{dt_r} \qquad \text{4- 14}$$

At every space-time point, if the field is weak, then the proper time τ is related to the coordinate time t by [Ashby and Spilker, 1996]

$$d\tau^2 = -g_{00}(\mathbf{r}) \cdot dt^2 \qquad \Rightarrow \qquad d\tau = \sqrt{-g_{00}(\mathbf{r})} \cdot dt \qquad \text{4- 15}$$

where g_{00} is the classical approximation of the zero-zero component of the metric tensor in a rotating frame such as the ECEF system

$$g_{00}(\mathbf{r}) = -\left[1 + \frac{2\Phi(\mathbf{r})}{c^2}\right] \qquad \text{4- 16}$$

where Φ is the total energy that a receiver or satellite has.

For a ground receiver, the total energy is the sum of the effective gravitational potential, kinetic energy and the centrifugal potential. However, the kinetic energy term can be neglected, as the ground velocity is normally very small when compared with the speed of light (especially when squared). The metric tensor for the receiver is $g_{00} = -\left[1 + \frac{2V(\mathbf{r}_r)}{c^2} - \frac{\omega^2 \|\mathbf{r}_r\|^2 \cos^2 B_r}{c^2}\right]$, where $V(\mathbf{r}_r)$ is the receiver gravitational potential, which can be approximated by only taking into account the Earth equatorial oblateness as

$$V_r = -\frac{GM}{\|\mathbf{r}_r\|}\left[1 - J_2 (\frac{a_e}{\|\mathbf{r}_r\|})^2 P_2(\sin B_r)\right] \qquad 4\text{-}17$$

where

- GM is the product of the universal gravitation constant G and the Earth's mass M;
- J_2=1.08262998905H10^{-3} is the Earth's second zonal harmonic coefficient, also known as the dynamical form factor;
- B_r is the latitude of the receiver;
- P_2 is the Legendre polynomial of degree two, $P_2(x)=0.5(3x^2-1)$;
- a_e is the semi-major axis of the WGS-84 ellipsoid.

Substituting g_{00} into Eq.4-15 and expanding it with the binomial series, the following expression for receiver r is obtained

$$\frac{dt_r}{d\tau_r} = \left[1 + \frac{2\Phi(\mathbf{r}_r)}{c^2}\right]^{-\frac{1}{2}} = 1 - \frac{\Phi(\mathbf{r}_r)}{c^2}$$
$$= 1 - \frac{V_r - \frac{1}{2}\omega^2 \|\mathbf{r}_r\|^2 \cos^2 B_r}{c^2} \qquad 4\text{-}18$$

For satellite s in the orbit, the total energy is the sum of gravitational potential and the kinetic energy, i.e. $V^s - \frac{1}{2}\|\dot{\mathbf{r}}^s\|^2$, then

$$\frac{dt^s}{d\tau_s} = 1 - \frac{V^s - \frac{1}{2}\|\dot{\mathbf{r}}^s\|^2}{c^2} \qquad 4\text{-}19$$

In its high orbit, the Earth gravitational potential can be derived from the field of point mass. Substituting Eq.4-14, 4-18 and 4-19 into Eq.4-13, the relationship between the received frequency f_r and the transmitted frequency f_s can be formulated as

$$f_r = f_s \left[\frac{1 - \frac{1}{c^2}\left[V_r - \frac{1}{2}\omega^2 \|\mathbf{r}_r\|^2 (\cos B)^2\right]}{1 - \frac{1}{c^2}\left[V^s - \frac{1}{2}\|\dot{\mathbf{r}}^s\|^2\right]}\right] \cdot \frac{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r}{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s}$$
$$= f_s \left[(1 - \frac{\Phi(\mathbf{r}_r)}{c^2})(1 + \frac{V^s - \frac{1}{2}\|\dot{\vec{r}}^s\|^2}{c^2})\right] \cdot \frac{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r}{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s} \qquad 4\text{-}20$$

The above equation can be rewritten into

$$f_r = f_s \left[1 + \frac{-\Phi(\mathbf{r}_r)}{c^2} + \frac{V^s - \frac{1}{2}\|\dot{\mathbf{r}}^s\|^2}{c^2}\right] \cdot \frac{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r}{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s} \qquad 4\text{-}21$$

Since the sum of the potential energy and the kinetic energy for a GPS satellite is constant during the motion in its orbit, this gives [Montenbruck and Gill, 2000]

$$V^s + \frac{1}{2}\left\|\dot{\mathbf{r}}^s\right\|^2 = -\frac{GM}{2a_{orb}} \qquad \Rightarrow \qquad \frac{1}{2}\left\|\dot{\mathbf{r}}^s\right\|^2 = -\frac{GM}{2a_{orb}} - V^s \tag{4-22}$$

where a_{orb} is the semi-major axis of the satellite orbit.

Substituting Eq.4-22 into Eq.4-21 yields

$$f_r = f_s\left[1 + \frac{-\Phi(\mathbf{r}_r)}{c^2} + \frac{2V^s + \frac{GM}{2a_{orb}}}{c^2}\right]\frac{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r}{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s} \tag{4-23}$$

Since the GPS system has already adopted a rate adjustment [Ashby and Spilker, 1996] by

$$f_s = f_0\left[1 + \frac{3GM}{2a_{orb} \cdot c^2} + \frac{\Phi_0}{c^2}\right] \tag{4-24}$$

where f_0 is the nominal GPS frequency, after incorporating the rate adjustment, Eq.4-23 changes to

$$\begin{aligned} f_r &= f_0\left[1 + \frac{\Phi_0 - \Phi(\mathbf{r}_r)}{c^2} + \frac{\frac{2GM}{a_{orb}} + 2V^s}{c^2}\right]\frac{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r}{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s} \\ &= f_0\left[1 + \frac{\Phi_0 - \Phi(\mathbf{r}_r)}{c^2} + \frac{2GM}{c^2}\left(\frac{1}{a_{orb}} - \frac{1}{\left\|\mathbf{r}^s\right\|}\right)\right]\frac{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r}{1 - \frac{1}{c}\mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s} \end{aligned} \tag{4-25}$$

This is the theoretical relativistic Doppler effect on GPS. As has been stated, this equation is tailored to applications in ECEF frames. The equivalent relativistic Doppler frequency shift equation for applications in an inertial system is given by Ashby [2003, p.27], which is in the form of

$$f_r = f_0\left[1+\frac{-V_r+\frac{1}{2}\dot{\mathbf{r}}_r^{\,2}+\Phi_0+\frac{2GM}{a_{orb}}+2V^s}{c^2}\right]\cdot\frac{1-\frac{1}{c}\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r}{1-\frac{1}{c}\mathbf{n}_r^s\bullet\dot{\mathbf{r}}^s} \qquad 4\text{-}26$$

As most high accuracy velocity determination applications are in the ECEF frame, Eq.4-25 can be conveniently used along with the GPS positioning. Moreover, it has several intuitive meanings:

- the last term $\frac{1-\frac{1}{c}\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r}{1-\frac{1}{c}\mathbf{n}_r^s\bullet\dot{\mathbf{r}}^s}$ is due to the special relativity induced by the relative motion of satellite *s* and receiver *r*; and
- the terms in the square brackets are the contributions to the Doppler frequency shift from the general relativity, where the second term $\frac{\Phi_0-\Phi(\boldsymbol{r}_r)}{c^2}$ is the correction of the receiver potential difference from the geoid, and the third term $\frac{2GM}{c^2}(\frac{1}{a_{orb}}-\frac{1}{\|\boldsymbol{r}^s\|})$ is the effect from the orbit eccentricity.

Hence, one can see that in GPS signal reception, extra frequency shifts are introduced due to the presence of the satellite orbital eccentricity and the receiver gravity potential difference.

The special relativity term in Eq.4-24 can be further expanded into the following form

$$\begin{aligned}\frac{1-\frac{1}{c}\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r}{1-\frac{1}{c}\mathbf{n}_r^s\bullet\dot{\mathbf{r}}^s} &\approx (1-\frac{1}{c}\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r)\left[1+\frac{1}{c}\mathbf{n}_r^s\bullet\dot{\mathbf{r}}^s+\frac{1}{c^2}(\dot{\mathbf{r}}^s)^2\right] \\ &= 1+\frac{\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s}{c}+\frac{\dot{\mathbf{r}}^s\bullet\dot{\mathbf{r}}_r^s}{c^2}\end{aligned} \qquad 4\text{-}27$$

where $\dot{\mathbf{r}}_r^s = \dot{\mathbf{r}}^s - \dot{\mathbf{r}}_r$. Neglecting the last term on the right hand side of Eq.4-27 and the terms in the square brackets of Eq.4-26, then it agrees with the Doppler frequency shift under the condition of special relativity only.

Substituting Eq.4-27 into Eq.4-26

$$
\begin{aligned}
f_r &= f_0\left[1+\frac{\Phi_0-\Phi(\mathbf{r}_r)}{c^2}+\frac{2GM}{c^2}\left(\frac{1}{a_{orb}}-\frac{1}{\|\mathbf{r}^s\|}\right)\right]\left(1+\frac{\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s}{c}+\frac{\dot{\mathbf{r}}^s\bullet\dot{\mathbf{r}}_r^s}{c^2}\right)\\
&\approx f_0+f_0\left[\frac{\Phi_0-\Phi(\mathbf{r}_r)}{c^2}+\frac{2GM}{c^2}\left(\frac{1}{a_{orb}}-\frac{1}{\|\mathbf{r}^s\|}\right)+\frac{\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s}{c}\right.\\
&\quad\left.+\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s\frac{[\Phi_0-\Phi(\mathbf{r}_r)]}{c^3}+\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s\frac{2GM}{c^3}\left(\frac{1}{a_{orb}}-\frac{1}{\|\mathbf{r}^s\|}\right)+\frac{\dot{\mathbf{r}}^s\bullet\dot{\mathbf{r}}_r^s}{c^2}\right]
\end{aligned}
\qquad 4\text{-}28
$$

This can be further simplified as

$$
\begin{aligned}
f_r - f_0 &= f_0\left[\frac{\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s}{c}+\frac{\Phi_0-\Phi(\mathbf{r}_r)}{c^2}+\frac{2GM}{c^2}\left(\frac{1}{a_{orb}}-\frac{1}{\|\mathbf{r}^s\|}\right)+\right.\\
&\quad\left.\frac{\dot{\mathbf{r}}^s\bullet\dot{\mathbf{r}}_r^s}{c^2}+\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s\frac{[\Phi_0-\Phi(\mathbf{r}_r)]}{c^3}+\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s\frac{2GM}{c^3}\left(\frac{1}{a_{orb}}-\frac{1}{\|\mathbf{r}^s\|}\right)\right]
\end{aligned}
\qquad 4\text{-}29
$$

Scaling the Doppler frequency shift into the velocity domain

$$
\begin{aligned}
\mathrm{D}\lambda_i &= \mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s+\frac{\Phi_0-\Phi(\mathbf{r}_r)}{c}+\frac{2GM}{c}\left(\frac{1}{a_{orb}}-\frac{1}{\|\mathbf{r}^s\|}\right)+\frac{\dot{\mathbf{r}}^s\bullet\dot{\mathbf{r}}_r^s}{c}+\\
&\quad\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s\frac{[\Phi_0-\Phi(\mathbf{r}_r)]}{c^2}+\mathbf{n}_r^s\bullet\dot{\mathbf{r}}_r^s\frac{2GM}{c^2}\left(\frac{1}{a_{orb}}-\frac{1}{\|\mathbf{r}^s\|}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \left[\mathbf{n}_r^s + \frac{\dot{\mathbf{r}}^s}{c} + \mathbf{n}_r^s \frac{[\Phi_0 - \Phi(\mathbf{r}_r)]}{c^2} + \mathbf{n}_r^s \frac{2GM}{c^2} (\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|}) \right] \bullet \dot{\mathbf{r}}_r^s + \\
&\quad \frac{\Phi_0 - \Phi(\mathbf{r}_r)}{c} + \frac{2GM}{c} (\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|}) \\
&\approx \left[\mathbf{n}_r^s + \frac{\dot{\mathbf{r}}^s}{c} \right] \bullet \dot{\mathbf{r}}_r^s + \frac{\Phi_0 - \Phi(\mathbf{r}_r)}{c} + \frac{2GM}{c} (\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|})
\end{aligned}
\tag{4-30}
$$

where $\frac{\dot{\boldsymbol{r}}^s}{c}$ acts as a line-of-sight direction correction applied to the range-rate due to the satellite velocity. The meaning of Eq.4-30 is clear. It shows that the Doppler shift relates to a "direction changed" receiver-to-satellite range-rate with the relativistic biases of the receiver potential difference and the satellite orbital eccentricity.

4.1.3 Doppler Shift Observation Equation

Thus far, we have established the inter-relationship of the theoretical Doppler shift and the relative motion between satellite *s* and receiver *r*, with the Earth's gravity field considered. Due to imperfections of the satellite clock and receiver clock, the presence of non-vacuum media such as the ionosphere and troposphere, and other inherent errors in a GPS observation, there are other terms in the measured GPS Doppler shift. Taking account for all the aforementioned effects, the proposed Doppler shift observation equation is as follows

$$
\begin{aligned}
\lambda_i D_{r,i}^s(t) = (\mathbf{n}_r^s + \frac{\dot{\mathbf{r}}^s}{c}) \bullet [\dot{\mathbf{r}}^s(t - \tau_r^s) - \dot{\mathbf{r}}_r(t)] - d\dot{I}_{r,i}^s + d\dot{T}_r^s + c \cdot d\dot{t}_r(t) - c \cdot d\dot{t}^s(t - \tau_r^s) + \\
\frac{\Phi_0 - \Phi(\mathbf{r}_r)}{c} + \frac{2GM}{c} (\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|}) + d\dot{R}_{Sagnac} + d\dot{R}_r^s + \varepsilon_r^s
\end{aligned}
\tag{4-31}
$$

where:

- subscript *i* designates the frequency L1 or L2;

- dR_{Sagnac} is the Sagnac correction for range measurements, and $d\dot{R}_{Sagnac}$ is the corresponding correction for the Doppler shift observables; and
- $d\dot{R}$ is the remaining relativistic correction for the Doppler shift measurements.

4.1.4 Receiver Velocity Determination

The velocity of a GPS receiver can be determined in an analogous way to GPS point positioning. To simplify the observation equation 4-31, it is assumed that all the errors and biases have been well modelled and corrected for by neglecting all except the receiver clock rate, as

$$\lambda_i D_{r,i}^s(t) \approx \mathbf{n}_r^s \bullet [\dot{\mathbf{r}}^s(t-\tau_r^s) - \dot{\mathbf{r}}_r(t)] + c \cdot d\dot{t}_r(t) + \varepsilon_r^s \qquad 4\text{-}32$$

where:

- $\boldsymbol{n}_r^s$ is the receiver to satellite line-of-sight unit vector;
- $\dot{\boldsymbol{r}}_r^s \equiv \dot{\boldsymbol{r}}^s - \dot{\mathbf{r}}_r$ is the receiver to satellite range-rate vector;
- $d\dot{t}_r(t)$ is the receiver clock rate.

Since the satellite velocity in the orbit can be calculated using the broadcast ephemeris, which will be discussed in Chapter Five, there are only four unknowns in Eq.4-32. These unknowns are the three vector components of the user velocity and one receiver clock rate. The linearisation process is also analogous to the case of point positioning, using the calculated satellite velocity, receiver velocity approximation and the computed Doppler value. Neglecting the tedious process, the final error equation for the Doppler measurement can be expressed in the form of

$$\begin{aligned} v_r^s &= -\mathbf{n}_r^s \bullet \Delta\dot{\mathbf{r}}_r + c \cdot d\dot{t}_r - f_r^s \\ &= \begin{bmatrix} -\mathbf{n}_r^s & 1 \end{bmatrix} \begin{bmatrix} \Delta\dot{\mathbf{r}}_r \\ c \cdot d\dot{t}_r \end{bmatrix} - f_r^s \end{aligned} \qquad 4\text{-}33$$

where f_r^s is the difference between the Doppler measurement and its nominal value calculated using the a priori values of the unknown parameters.

With at least four Doppler shift measurements at one epoch, the error equation system can be formed as

$$\begin{bmatrix} v_r^1 \\ v_r^2 \\ \vdots \\ v_r^n \end{bmatrix} = \begin{bmatrix} -\mathbf{n}_r^1 & 1 \\ -\mathbf{n}_r^2 & 1 \\ \vdots & \vdots \\ -\mathbf{n}_r^n & 1 \end{bmatrix} \begin{bmatrix} \Delta\dot{\mathbf{r}}_r \\ c \cdot dt_r \end{bmatrix} - \begin{bmatrix} f_r^1 \\ f_r^2 \\ \vdots \\ f_r^n \end{bmatrix} \qquad 4\text{-}34$$

Comparing the above error equation with Eq.3-6, the design matrix is identical. This property eases the computational load for velocity determination. The velocity can be easily established along with the position because one can "reuse" the normal equation to save computational time.

When the Doppler measurements are used in a static mode, one can obtain a position fix with only three GPS satellites since there are six observables while there are only five unknowns, i.e. three position vector components, one clock bias and one clock rate [Hofmann-Wellenhof et al., 2001].

The above mentioned velocity determination method can be used to obtain the *a priori* receiver velocity. Rigorous velocity determination can then be carried out based on Eq.4-31, with all the errors in Doppler shift measurements corrected for.

4.2. Acceleration Determination Using GPS

There are generally two methods of acceleration determination using GPS. Both methods have been widely used in airborne gravimetry where differential GPS techniques are employed. Data from the GPS receiver whose antenna is mounted on an aircraft are processed together with data from the ground reference GPS receivers.

The most popular method of determining aircraft acceleration from such systems is the position method, whereby positions in the aircraft trajectory are twice differentiated with respect to time. Typical applications of this method can be found in the literature [Brozena et al., 1989; Wei and Schwarz, 1995; Bruton et al., 1999; Psiaki et al., 1999; 2000; Bruton and Schwarz, 2002]. The position method requires the positional accuracy to be as high as possible since the double differentiation amplifies the noise rather quickly.

The alternative method is the Doppler rate method where the change-rate of Doppler shift is derived from either the carrier phase (second-derivative) or the precise Doppler measurement (first-derivative). Although in the literature this method is generally referred to as the carrier acceleration method, it is not restricted to carrier measurements and the author believes it is more appropriate to refer to as the Doppler rate method. The change-rate of Doppler shifts has been derived and used to determine the receiver's acceleration.

4.2.1 Range Acceleration

Following Jekeli and Garcia [1997], vector operation is used to derive the range acceleration. It can be seen from Fig.4-2 that

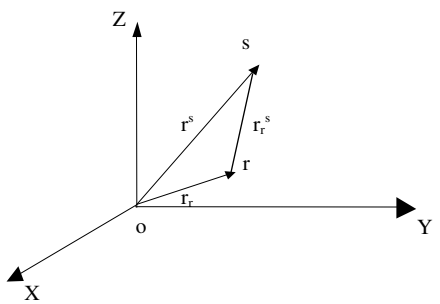

Figure 4- 2: Geometric relationships of receiver-satellite vectors

$$\mathbf{r}^s = \mathbf{r}_r + \mathbf{r}_r^s \qquad 4\text{-}35$$

where $\boldsymbol{r}_r^s$ can be expressed in terms of the unit vector $\boldsymbol{n}_r^s$ and the range ρ_r^s

$$\mathbf{r}_r^s = \mathbf{n}_r^s \cdot \rho_r^s \qquad 4\text{-}36$$

From Eq.4-36, the range ρ_r^s can be expressed by

$$\rho_r^s = \mathbf{n}_r^s \bullet \mathbf{r}_r^s \qquad 4\text{-}37$$

Differentiating Eq.4-37 with respecte to time gives

$$\dot{\rho}_r^s = \dot{\mathbf{n}}_r^s \bullet \mathbf{r}_r^s + \mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s \qquad 4\text{-}38$$

Note that $\dot{\boldsymbol{n}}_r^s$ is perpendicular to vector $\boldsymbol{r}_r^s$, therefore their dot product is zero. This leads to

$$\dot{\rho}_r^s = \mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s \qquad 4\text{-}39$$

In Eq.4-39, the range-rate is a dot product of the line-of-sight unit vector and the relative velocity between the satellite and receiver. This is the fundamental velocity equation used in the derivations in § 4.1.

From the fundamental velocity equation, the range change-rate between receiver r to satellite s, $\dot{\rho}_r^s$, can be rewritten as follows

$$\dot{\rho}_r^s = \mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r^s = \mathbf{n}_r^s \bullet \dot{\mathbf{r}}^s - \mathbf{n}_r^s \bullet \dot{\mathbf{r}}_r \qquad 4\text{-}40$$

Differentiating Eq.4-40 with respect to time yields

$$\begin{aligned} \ddot{\rho}_r^s &= \dot{\mathbf{n}}_r^s \bullet \dot{\mathbf{r}}^s - \dot{\mathbf{n}}_r^s \bullet \dot{\mathbf{r}}_r + (\mathbf{n}_r^s \bullet \ddot{\mathbf{r}}^s - \mathbf{n}_r^s \bullet \ddot{\mathbf{r}}_r) \\ &= \dot{\mathbf{n}}_r^s \bullet \dot{\mathbf{r}}_r^s + (\mathbf{n}_r^s \bullet \ddot{\mathbf{r}}^s - \mathbf{n}_r^s \bullet \ddot{\mathbf{r}}_r) \end{aligned} \qquad 4\text{-}41$$

There are only two unknowns in Eq.4-41: $\ddot{\boldsymbol{r}}_r$ and $\dot{\boldsymbol{n}}_r^s$. The latter is an intermediate unknown that must be solved first. To obtain it, differentiating the first term on the right hand side of equation 4-36, and then rearranging terms

$$\dot{\mathbf{n}}_r^s = \frac{1}{\rho_r^s} \cdot \left(\dot{\mathbf{r}}_r^s - \mathbf{n}_r^s \cdot \dot{\rho}_r^s\right) \qquad 4\text{-}42$$

Substituting $\dot{\boldsymbol{n}}_r^s$ into equation 4-36 and replacing $\dot{\boldsymbol{r}}_r^s$ with $\dot{\boldsymbol{n}}_r^s \bullet \rho_r^s + \boldsymbol{n}_r^s \cdot \dot{\rho}_r^s$, the range acceleration is

$$\begin{aligned}\ddot{\rho}_r^s &= \frac{1}{\rho_r^s} \cdot \left[(\dot{\mathbf{r}}_r^s)^2 - (\dot{\rho}_r^s)^2\right] + (\mathbf{n}_r^s \bullet \ddot{\mathbf{r}}^s - \mathbf{n}_r^s \bullet \ddot{\mathbf{r}}_r) \\ &= \mathbf{n}_r^s \bullet \ddot{\mathbf{r}}_r^s + \frac{1}{\rho_r^s} \cdot \left[(\dot{\mathbf{r}}_r^s)^2 - (\dot{\rho}_r^s)^2\right]\end{aligned} \qquad 4\text{-}43$$

where unit vector properties of $\boldsymbol{n}_\mathrm{r}^\mathrm{s} \bullet \boldsymbol{n}_\mathrm{r}^\mathrm{s} = 1$ and $\dot{\boldsymbol{n}}_\mathrm{r}^\mathrm{s} \bullet \boldsymbol{n}_\mathrm{r}^\mathrm{s} = 0$ have been employed in the derivation process. It can be seen that the receiver to satellite range acceleration consists of two parts: one is the acceleration caused by the relative motion of the receiver and satellite pair projected into the line-of-sight direction; and the other is a centrifugal-like acceleration (ibid).

The "centrifugal" acceleration term has very good tolerance for velocity errors owing to the long receiver to satellite distance in the denominator. Since the GPS satellite acceleration can be calculated with better than $\pm 0.1\mathrm{mm/s}^2$ level of accuracy (see Chapter Five), if the range acceleration can be precisely obtained, then a receiver's acceleration may be resolved with relatively high accuracy.

4.2.2 Doppler Rate Observation Equation

In GPS, unfortunately, there is no direct observation of the range acceleration $\ddot{\rho}_r^s$. However, the change-rate of the Doppler shift (Doppler rate hereafter) measurements with respect to time can be numerically obtained and used as a "virtual" measurement, which relates to the range acceleration by

$$\lambda_i \dot{D}_i(t) = \lambda_i \ddot{\Phi}_i(t) = \ddot{\rho}_r^s \qquad 4\text{-}44$$

This virtual observable has been introduced by Jekeli [1994], and followed by Jekeli and Garcia [1997] and Kennedy [2003]. The Doppler rate is the first-derivative of Doppler shift with respect to time, and the second-derivative of the carrier phase

measurement. In order to obtain precise Doppler rate, different differentiators have been adopted based on precise carrier phase observables. However, various methods of differentiation may have various effects on the resultant derivative, and their suitability varies from situation to situation. A good reference for differentiator design in GPS applications is Bruton et al. [1999]. However, a comprehensive discussion on differentiator designs is provided in Chapter Seven, with an emphasis on real-time and dynamic applications.

Similar to the GPS range and Doppler measurements, this virtual observable is biased by GPS errors, and therefore relates only to the "pseudo range acceleration" [Jekeli and Garcia., 1997]

$$\begin{aligned}\lambda_i \dot{D}_i(t) = \ddot{\rho}_r^s = \mathbf{n}_r^s \bullet \ddot{\mathbf{r}}_r^s + \frac{1}{\rho_r^s} \cdot \left[(\dot{\mathbf{r}}_r^s)^2 - (\dot{\rho}_r^s)^2 \right] - d\ddot{I}_{r,i}^s + d\ddot{T}_r^s + c \cdot d\ddot{t}_r(t) \\ - c \cdot d\ddot{t}^s(t - \tau_r^s) + d\ddot{M}_r^s - d\ddot{R}_r^s + \varepsilon\end{aligned} \qquad 4\text{-}45$$

where:

- $d\ddot{I}$ is the second derivative of the ionospheric errors with respect to time;
- $d\ddot{T}$ is the second derivative of the tropospheric errors with respect to time;
- dt_r, dt^s are the receiver clock and satellite clock errors respectively. The double dots over them represent the second derivatives with respect to time;
- $d\ddot{M}$ is the second derivative with respect to time of the multipath errors;
- $d\ddot{R}$ is the second derivative with respect to time of the relativistic effect errors.

Comparing Eq.4-45 with Eq.4-31, there are no terms corresponding to the line-of-sight correction and the receiver potential difference terms. One of the explanations is that they might be present in the Doppler rate "observable"; however, the magnitudes are so small that they are negligible.

4.2.3 Receiver Acceleration Determination

Receiver acceleration determination is analogous to point positioning and receiver velocity determination described in § 4.1.4. The acceleration of a GPS receiver *r* can be determined by assuming that the errors and biases in the Doppler rate "observations" at epoch *t* have been appropriately corrected for with the receiver clock acceleration as the only term remaining, and the Doppler rate observables have been derived. As such, the simplified Doppler rate "observation" equation for the receiver *r* and satellite *s* is

$$\lambda_i \dot{D}_i(t) = \mathbf{n}_r^s \bullet \ddot{\mathbf{r}}^s - \mathbf{n}_r^s \bullet \ddot{\mathbf{r}}_r + \frac{1}{\rho_r^s} \cdot \left[(\dot{\mathbf{r}}_r^s)^2 - (\dot{\rho}_r^s)^2 \right] + c \cdot \ddot{dt}_r(t) + \varepsilon \qquad 4\text{-}46$$

Since the satellite velocity and acceleration are known, and the receiver to satellite range-rate can be calculated, and this leaves four unknowns in Eq.4-46. They are the three receiver acceleration components and the second time derivative of receiver clock.

Therefore, at one epoch if there are at least four satellites being tracked, and for each satellite tracked the Doppler rate $\dot{D}$ can be derived from Doppler measurements or carrier phase measurements, then the instantaneous receiver acceleration can be determined.

4.3 Summary

The principles of velocity and acceleration determination using a GPS receiver in standalone mode have been presented in this chapter. Doppler and Doppler rate measurements are used in precise velocity and acceleration determination respectively. Taking the velocities and accelerations of the in-view satellites as known, the velocity and acceleration determination using GPS is very similar to GPS point positioning.

A highly accurate Doppler shift observation equation has been developed with both the special and general relativity theories taken into consideration. It is demonstrated that the line-of-sight direction has a slight change towards the satellite motion direction. An interesting research question is therefore raised to see if this also applies in the range measurement. It is also illustrated that there is a general relativistic effect term due to the receiver potential difference from that of the geoid, and a relativistic correction due to the satellite orbit eccentricity.

The principle of acceleration determination using GPS is discussed based on a virtual GPS observable, i.e. the change-rate of the Doppler shift or simply the Doppler shift rate. The methods to obtain the Doppler shift rate "observable" will be discussed in Chapter Seven.

The satellite ECEF velocity and acceleration determination using the broadcast ephemeris will be discussed in Chapter Five, and methods of error correction for the Doppler and Doppler rate measurements will be presented in Chapter Six.

Chapter Five

SATELLITE VELOCITY AND ACCELERATION DETERMINATION USING THE GPS BROADCAST EPHEMERIS

The principles of position, velocity and acceleration determination using GPS have been given in the previous chapters where the satellite ECEF position, velocity and acceleration (PVA) have been assumed to be known. The question is how to obtain the satellite PVA in the ECEF system. The satellite ECEF position algorithm using the broadcast ephemeris is described in the GPS interface control document ICD-GPS-200c. However, the algorithms for real-time satellite ECEF velocity and acceleration determination are not available. This chapter deals with the topic of satellite ECEF velocity and acceleration determination using the GPS broadcast ephemeris.

Section 5.1 introduces the natural satellite orbit system and addresses its distinction from the orbit system adopted by the GPS control segment. The method of GPS ECEF velocity determination using the ordinary rotation method is given in detail. The complexity of the rotation method has resulted in an easy alternative using a simple positional differentiation method.

A closed-form satellite ECEF acceleration algorithm is also presented, and the accuracy of this closed form formula is tested by a comparison with the accelerations from the IGS SP3 precise ephemeris. Many real-time applications require that the PVA information from GPS receivers can be output at a high rate, thus polynomial

representation schemes for the satellite PVA are discussed in order to speed up the satellite PVA determination.

5.1 Satellite Orbit Representation by Keplerian Parameters

A satellite orbiting the Earth can be described by Kelpler's laws. However, due to the Earth's oblate mass distribution, the gravitational attractions from the Sun and the Moon, and other disturbing forces, there are perturbations in the GPS satellite orbit.

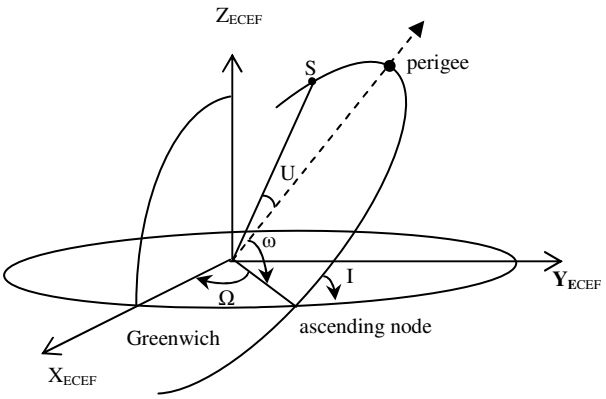

Figure 5- 1: Keplerian elements and satellite orbit

Figure 5-1 illustrates the representation of a satellite position using the Keplerian elements. Satellite positions in an ideal, non-perturbing orbit can be represented by:

- Size and shape of the ellipse: semi-major axis a and eccentricity e;
- Orientation of the orbital plane relative to the Earth: orbit inclination I and longitude of the ascending node Ω;
- Orientation of the ellipse in the orbital plane: argument of perigee ω;
- Satellite position in the ellipse: true anomaly U;
- Reference time: t (time of perigee passage), or t_{oe} (time of the reference ephemeris).

5.1.1 Satellite PVA in "Natural" Orbital Plane

The "natural" satellite orbital plane system is defined as follows: the origin is located at one focus of the elliptical orbit, which corresponds to the position of the mass centre of the Earth. The X-axis and Y-axis are defined as being along the major axis and parallel to the minor axis respectively. Figure 5-2 illustrates the natural orbital system where the perigee is in line with the X-axis.

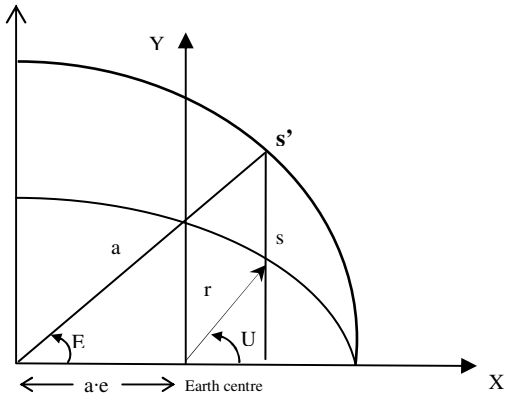

Figure 5- 2: Satellite position in the orbital plane coordinate system

A satellite position in the natural orbit plane system can be expressed [Beutler, 1998; Misra and Enge, 2001] by

$$\mathbf{r} = \begin{pmatrix} X \\ Y \\ 0 \end{pmatrix} = \begin{pmatrix} a\cos E - a \cdot e \\ a\sqrt{1-e^2}\,\sin E \\ 0 \end{pmatrix} = \begin{pmatrix} r\cos U \\ r\sin U \\ 0 \end{pmatrix} \qquad 5\text{-}1$$

where:

- a is the semi-major axis of the satellite orbit;
- e is the eccentricity of the orbit;
- E is the orbital eccentric anomaly;

- r is the instantaneous distance between the satellite and the centre of mass of the Earth; and
- U is the true anomaly.

The satellite velocity in the natural orbital plane system is given by (ibid)

$$\dot{\mathbf{r}} = \frac{a \cdot n}{1 - e \cdot \cos E} \cdot \begin{pmatrix} -\sin E \\ \sqrt{1-e^2} \cdot \cos E \\ 0 \end{pmatrix} \qquad 5\text{-}2$$

where n is the mean motion of the satellite. The satellite acceleration in the orbital plane system is (ibid)

$$\ddot{\mathbf{r}} = -\frac{GM}{r^3} \cdot \dot{\mathbf{r}} \qquad 5\text{-}3$$

where G is the universal gravitational constant and M is the Earth's mass.

5.1.2 Transform Satellite Position to ECEF System

It can be seen from Fig.5-1 that the transformation of a satellite position from the orbital coordinate system to the ECEF may be carried out by three rotations in the following way

- First rotate the argument of the perigee ω clockwise to the ascending node;
- Then rotate the angle of I clockwise to the equatorial plane;
- Finally rotate the angle of the longitude of ascending node Ω clockwise to the Greenwich prime meridian.

The corresponding transform equation is

$$\mathbf{r}_{ECEF} = \mathbf{R}_3(-\Omega) \cdot \mathbf{R}_1(-I) \cdot \mathbf{R}_3(-\omega) \cdot \mathbf{r} \qquad 5\text{-}4$$

where $\boldsymbol{R}_n(\theta)$ is the rotation matrix, the subscript n = 1,3 corresponding to the rotation axes of X and Z respectively. The rotation matrixes $\boldsymbol{R}_3(\theta)$ $\boldsymbol{R}_1(\theta)$ are given as [Farrell and Barth, 1999, p.34]

$$\mathbf{R}_1(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \qquad \mathbf{R}_3(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5-5}$$

5.2. ICD-200 Orbital Coordinate System

The algorithm for GPS satellite ECEF position determination from the broadcast ephemeris is presented in the ICD-GPS-200c. It is important to stress that the orbital coordinate system used in the ICD-GPS-200c (ICDorb) is slightly different from the above "natural" orbital system. The ICDorb system sets the X-axis aligned to the right ascending node rather than the perigee. The difference is reflected mainly in the parameterisation and calculation of the three pairs of amplitudes of harmonic correction terms for the sinusoid correction models of the argument of latitude, the orbit radius, and the inclination angle, see Eq.5-16 and Eq.5-17.

By such a coordinate system definition, there are only two rotations to transform the satellite position from the "natural" orbital system to the ECEF system, i.e.

$$\mathbf{r}_{\mathrm{ECEF}} = \mathbf{R}_3(-\Omega_c) \cdot \mathbf{R}_1(-\mathrm{I}) \cdot \mathbf{r}_{\mathrm{ICDorb}} \equiv \mathbf{R}_i^e \cdot \mathbf{r}_{\mathrm{ICDorb}} \tag{5-6}$$

where $\boldsymbol{R}_i^e$ is defined as the rotation matrix from the ICDorb to the ECEF. According to the ICD-GPS-200c, the rotation matrix has the following form

$$\mathbf{R}_i^e = \begin{pmatrix} \cos\Omega_c & -\sin\Omega_c\cos\mathrm{I} & \sin\Omega_c\sin\mathrm{I} \\ \sin\Omega_c & \cos\Omega_c\cos\mathrm{I} & -\cos\Omega_c\sin\mathrm{I} \\ 0 & \sin\mathrm{I} & \cos\mathrm{I} \end{pmatrix} \tag{5-7}$$

where Ω_c is the corrected right ascension of the ascending node, which is calculated by

$$\Omega_c = \Omega_0 + (\dot{\Omega} - \dot{\Omega}_e)t_k + \dot{\Omega}_e \cdot t_{oe} \qquad 5\text{-}8$$

where:

- Ω_0 is the right ascension of the ascending node at reference time;
- $\dot{\Omega}$ is the change-rate of the right ascension;
- $\dot{\Omega}_e$ is the angular change-rate of the Earth's rotation;
- t_k is the calculation time;
- t_{oe} is the reference time of the ephemeris parameters.

To better represent the satellite positions in the orbit, three change-rate parameters are used to describe the linear change characteristics of the satellite mean motion n, Ω_c, and I. These rate parameters are vitally important for the ECEF satellite velocity and acceleration determination.

The adoption of the ICD-GPS-200c orbit representation scheme by the GPS Control Segment has some advantages. Firstly, the satellite position calculation could become more effective given the fact that the GPS orbit is near circular and the effective orbit representation time is short (2~3 hours), there are only two rotations to transform an orbital position in the ECEF system. Secondly, there is no need to introduce another parameter, i.e. the change-rate of the argument of perigee thus reducing the payload of the navigation message. Since the perigee itself is hard to define in orbits with small eccentricities [Montenbruck and Gill, 2000: p.30], the adopted broadcast orbit representation scheme not only alleviates the navigation payload but also avoids the difficulty in defining the change-rate of the argument of perigee. However, as will be

discussed later, it is rather complicated when the satellite velocity needs to be determined.

5.3 Transform GPS Orbital Velocity to ECEF

The transformation of a GPS satellite velocity from the ICDorb system to the ECEF may be simply given through differentiation of Eq.5-6 by

$$\dot{\mathbf{r}}_{\mathrm{ECEF}} = \dot{\mathbf{R}}_{\mathrm{i}}^{\mathrm{e}} \cdot \mathbf{r}_{\mathrm{ICDorb}} + \mathbf{R}_{\mathbf{i}}^{\mathbf{e}} \cdot \dot{\mathbf{r}}_{\mathrm{ICDorb}} \qquad 5\text{-}9$$

where:

- $\dot{\boldsymbol{R}}_i^e$ is the first-derivative of the rotation matrix $\boldsymbol{R}_i^e$ with respect to time;
- $\dot{\boldsymbol{r}}_{ICDorb}$ is the velocity of the satellite in the ICDorb system.

These two unknowns must be resolved prior to obtaining the ECEF velocity.

5.3.1 Determination of $\dot{\boldsymbol{R}}_i^e$

To determine the first-derivative of the rotation matrix $\dot{\boldsymbol{R}}_i^e$, it is necessary to obtain the derivatives of the corrected longitude of ascending node Ω_c and the inclination I. These can be done simply through their own definitions by

$$\begin{aligned} \dot{\Omega}_c &= \dot{\Omega} - \dot{\Omega}_e \\ \dot{I} &= \dot{i} \end{aligned} \qquad 5\text{-}10$$

where $\dot{i}$ is IDOT, the inclination change-rate, which is one of the parameters in the broadcast ephemeris.

Differentiating Eq.5-7 with respect to time, and then substituting the above rates

$$\dot{\mathbf{R}}_i^e = \begin{pmatrix} -\sin\Omega_c \cdot \dot{\Omega}_c & -\cos\Omega_c \cos I \cdot \dot{\Omega}_c + \sin\Omega_c \sin I \cdot \dot{I} & \cos\Omega_c \sin I \cdot \dot{\Omega}_c + \sin\Omega_c \cos I \cdot \dot{I} \\ \cos\Omega_c \cdot \dot{\Omega}_c & -\sin\Omega_c \cos I \cdot \dot{\Omega}_c - \cos\Omega_c \sin I \cdot \dot{I} & \sin\Omega_c \sin I \cdot \dot{\Omega}_c - \cos\Omega_c \cos I \cdot \dot{I} \\ 0 & \cos I \cdot \dot{I} & -\sin I \cdot \dot{I} \end{pmatrix} \quad 5\text{- }11$$

The magnitude of $\dot{I}$ is of the order of 10^{-10} to 10^{-12} [Gurtner, 2001], while the magnitude of the *sinI* and *cosI* terms are less than 1.0, so Eq.5-11 can be simplified by neglecting those $\dot{I}$ terms without losing numerical precision to become

$$\dot{\mathbf{R}}_i^e = \begin{pmatrix} -\sin\Omega_c \cdot \dot{\Omega}_c & -\cos\Omega_c \cos I \cdot \dot{\Omega}_c & \cos\Omega_c \sin I \cdot \dot{\Omega}_c \\ \cos\Omega_c \cdot \dot{\Omega}_c & -\sin\Omega_c \cos I \cdot \dot{\Omega}_c & \sin\Omega_c \sin I \cdot \dot{\Omega}_c \\ 0.0 & 0.0 & 0.0 \end{pmatrix} \quad 5\text{- }12$$

5.3.2 Determination of $\dot{r}_{\mathrm{ICDorb}}$

The satellite coordinate in the ICDorb system is calculated by

$$\mathbf{r}_{ICDorb} = \begin{pmatrix} X_{ICDorb} \\ Y_{ICDorb} \end{pmatrix} = \begin{pmatrix} r_c \cdot \cos Vc \\ r_c \cdot \sin Vc \end{pmatrix} \quad 5\text{- }13$$

where r_c is the corrected radius and V_c is the corrected argument of latitude. Note that subscript c is used to indicate that they are "corrected" and calculated relative to the ascending node. These two variables can be calculated by

$$r_c = a(1 - e \cdot \cos E) + dr \quad 5\text{- }14$$

$$V_c = U + \omega + dU \quad 5\text{- }15$$

where dr, dU are the harmonic perturbation corrections given by

$$dr = crc \cdot \cos 2(U + \omega) + crs \cdot \sin 2(U + \omega) \quad 5\text{- }16$$

$$dU = cuc \cdot \cos 2(U + \omega) + cus \cdot \sin 2(U + \omega) \qquad 5\text{-}17$$

where *crc, crs, cuc, cus* are the harmonic perturbation parameters in the broadcast ephemeris for r_c and V_c respectively. It is evident from the above equations that the ICDorb system is actually defined with the X-axis pointing to the ascending node and the broadcast ephemeris is parameterised accordingly.

The orbital velocity can be obtained by differentiating Eq.5-13 with respect to time

$$\dot{\mathbf{r}}_{ICDorb} = \begin{pmatrix} \dot{X}_{ICDorb} \\ \dot{Y}_{ICDorb} \end{pmatrix} = \begin{pmatrix} \dot{r}_c \cos V_c - r_c \sin V_c \cdot \dot{V}_c \\ \dot{r}_c \sin V_c + r_c \cos V_c \cdot \dot{V}_c \end{pmatrix} \qquad 5\text{-}18$$

Equation 5-18 shows that the derivatives of r_c and V_c should be determined prior to the calculation of satellite velocity in the ICDorb coordinate system.

5.3.2.1 Derivation of $\dot{r}_c$

$\dot{r}_c$ may be derived by differentiating Eq.5-14 with respect to time. The differentiation could be carried out in two steps.

The first step is to differentiate the first term on the right hand side of Eq.5-14 as

$$[a(1 - e \cdot \cos E)]' = a \cdot e \cdot \sin E \cdot \dot{E} \qquad 5\text{-}19$$

where the superscript prime is a differentiation operator and $\dot{E}$ may be obtained from the Kepler's equation [Marshall, 2002], or simply from a comparison of Eq.5-1 and Eq.5-2

$$\dot{E} = \frac{n}{1 - e \cdot \cos E} \qquad 5\text{-}20$$

It can also be numerically computed along with the solution of Kepler's equation in an alternative form as

$$\dot{M} = \dot{E} \cdot (1 - e \cdot \cos E) \qquad \Rightarrow \qquad \dot{E} = \frac{\dot{M}}{1 - e \cdot \cos E} \tag{5-21}$$

The second step is to obtain the derivative of the second term, *dr*, on the right hand side of Eq.5-14 through differentiating Eq.5-16, assuming that ω is an invariant

$$[dr]' = -2[crc \cdot \sin 2(U + \omega) - crs \cdot \cos 2(U + \omega)]\dot{U} \tag{5-22}$$

where the change-rate of the true anomaly $\dot{U}$ is still unknown. To comply with Kepler's second law which states equal area in an infinitely small time interval *dt* by

$$r_c^2 \dot{U} \cdot dt = a\sqrt{1 - e^2} \cdot n \cdot dt \tag{5-23}$$

leads to

$$\dot{U} = \frac{a\sqrt{1 - e^2} \cdot n}{r_c^{\ 2}} \tag{5-24}$$

5.3.2.2 Derivation of $\dot{V}_c$

$\dot{V}_c$, the derivative of the argument of latitude with respect to time, can be derived in a similar manner by differentiating Eq.5-15 and holding ω as a constant

$$\dot{V}_c = \dot{U} + [dU]' \tag{5-25}$$

where the only unsolved quantity is the second term on the right hand side. This may be easily obtained by differentiating Eq.5-17, taking ω as an invariant

$$[dU]' = -2[cuc \cdot \sin 2(U + \omega) - cus \cdot \cos 2(U + \omega)]\dot{U} \qquad 5\text{-}26$$

A C++ implementation of the above velocity algorithm can be found in the source code from the National Geodetic Service (NGS) website[Marshall, 2002]. It is embedded in function *bccalc*(…) in the program file *skyplot.cpp*. Recently, an independent program using the same algorithm was presented by Remondi [2004] at the NGS website. Remondi's derivation and subsequent publication were to answer the questions originally raised by the author in this regard when consulting with Steve Hilla who is in charge of the NGS GPS Toolbox.

The algorithm is detailed in this chapter since firstly it is presented in the form of source code and pseudo-code respectively in both references, and secondly the derivation of the satellite ECEF acceleration requires some of the equations in the velocity determination.

The derivation of the GPS satellite velocity in the ICDorb system is "complicated", as commented in the source code *skyplot.cpp*. The complexity is caused mainly by the derivation of change-rates of the corrected radius and the corrected argument of latitude, which is due to the ICD-GPS-200's orbital system and its effects on the representation of the harmonic perturbation parameters. The process of calculating $\dot{E}$ also contributes to the complexity of the algorithm. Note that by taking ω as a constant several times in the derivation process, it may cause an extra error in the resultant ICDorb velocity.

Table 5-1 lists both the velocities and accelerations (for PRN 07 on August 20, 2002) derived using the IGS SP3 precise ephemeris. The accelerations were derived using the first-order central difference of a Taylor series of IGS SP3 velocities. The velocities in Table 5-2 are calculated using the rotation matrix method, and the accelerations are determined using the proposed acceleration formula. Table 5-1 and Table 5-2 have the same period for the same satellite so one can compare the accuracy

of velocity and acceleration from the broadcast ephemeris with the precise ephemeris directly. It is useful to point out that, although the precise ephemeris is defined with reference to the mass centre of a satellite, which differs from the antenna centre to which the broadcast ephemeris is referenced, the velocities can be directly compared with each other.

Time	A_x (m/s^2)	A_y (m/s^2)	A_z (m/s^2)	V_x (m/s)	V_y (m/s)	V_z (m/s)
22:19:01	0.2055	-0.3022	0.1413	329.9513	-888.5965	-2997.0555
22:19:02	0.2054	-0.3022	0.1414	330.1568	-888.8987	-2996.9141
22:19:03	0.2054	-0.3022	0.1414	330.3622	-889.2009	-2996.7727
22:19:04	0.2053	-0.3023	0.1415	330.5676	-889.5032	-2996.6312
22:19:05	0.2053	-0.3023	0.1415	330.7729	-889.8055	-2996.4897
22:19:06	0.2052	-0.3023	0.1416	330.9782	-890.1078	-2996.3481
22:19:07	0.2052	-0.3023	0.1417	331.1834	-890.4101	-2996.2064
22:19:08	0.2051	-0.3023	0.1418	331.3886	-890.7124	-2996.0646
22:19:09	0.2051	-0.3024	0.1418	331.5937	-891.0147	-2995.9228
22:19:10	0.2050	-0.3024	0.1419	331.7988	-891.3171	-2995.7809

Table 5- 1: Acceleration and velocity from the IGS SP3 precise (PRN07, 20 August 2002) positions and velocities

Time	A_x (m/s^2)	A_y (m/s^2)	A_z (m/s^2)	V_x (m/s)	V_y (m/s)	V_z (m/s)
22:19:01	0.2055	-0.3022	0.1412	329.9518	-888.5962	-2997.0552
22:19:02	0.2054	-0.3022	0.1413	330.1573	-888.8985	-2996.9138
22:19:03	0.2054	-0.3022	0.1413	330.3627	-889.2007	-2996.7724
22:19:04	0.2053	-0.3022	0.1414	330.5680	-889.5030	-2996.6309
22:19:05	0.2053	-0.3022	0.1415	330.7733	-889.8052	-2996.4894
22:19:06	0.2052	-0.3022	0.1415	330.9786	-890.1075	-2996.3478
22:19:07	0.2052	-0.3023	0.1416	331.1838	-890.4098	-2996.2061
22:19:08	0.2051	-0.3023	0.1417	331.3890	-890.7121	-2996.0643
22:19:09	0.2051	-0.3023	0.1417	331.5941	-891.0145	-2995.9225
22:19:10	0.2050	-0.3023	0.1418	331.7992	-891.3168	-2995.7807

Table 5- 2: Acceleration and velocity from the rotation method using the broadcast ephemeris

It can be seen from Table 5-1 and Table 5-2 that although there are small biases in each axis, the velocities derived from the broadcast ephemeris are close to those from the precise ephemeris, well within ±1mm/s for each axis component.

5.4 Transform GPS Orbital Acceleration to ECEF

Since the orbital acceleration formula is in terms of the position vector, which is independent on the orbital orientation, the transformation is much easier, see Eq.5-3. One only needs to account for the second-derivative of the rotation matrix $\boldsymbol{R}_{\mathrm{i}}^{\mathrm{e}}$ with respect to time, and then the acceleration transformation can be carried out through differentiating Eq.5-9 with respect to time. This yields

$$\ddot{\mathbf{r}}_{ECEF} = \ddot{\mathbf{R}}_i^e \cdot \mathbf{r}_{ICDorb} + 2\dot{\mathbf{R}}_i^e \cdot \dot{\mathbf{r}}_{ICDorb} + \mathbf{R}_i^e \cdot \ddot{\mathbf{r}}_{ICDorb} \qquad 5\text{-}27$$

where $\ddot{\boldsymbol{R}}_{\mathrm{i}}^{\mathrm{e}}$ is the second-derivative of the rotation matrix $\boldsymbol{R}_{\mathrm{i}}^{\mathrm{e}}$. Taking $\ddot{I}$, $\ddot{\Omega}$ and thus $\ddot{\Omega}_{\mathrm{c}}$ as zero, and those $\dot{I}$ terms also as zero, the second-derivative of the rotation matrix can be derived from Eq.5-11 as

$$\ddot{\mathbf{R}}_i^e = \begin{pmatrix} -\cos\Omega_c \cdot \dot{\Omega}_c^{\,2} & \sin\Omega_c \cdot \dot{\Omega}_c^{\,2} \cdot \cos I & -\sin\Omega_c \cdot \dot{\Omega}_c^{\,2} \cdot \sin I \\ -\sin\Omega_c \cdot \dot{\Omega}_c^{\,2} & -\cos\Omega_c \cdot \dot{\Omega}_c^{\,2} \cdot \cos I & \cos\Omega_c \cdot \dot{\Omega}_c^{\,2} \cdot \sin I \\ 0.0 & 0.0 & 0.0 \end{pmatrix} \qquad 5\text{-}28$$

In Eq.5-27, the satellite position vector $\boldsymbol{r}_{ICDorb}$ is readily derived from Eq.5-13, $\boldsymbol{R}_i^e$ and $\dot{\boldsymbol{R}}_i^e$ are obtained from Eq.5-7 and Eq.5-12 respectively, $\dot{\boldsymbol{r}}_{ICDorb}$ can be calculated using Eq.5-18, and the ICDorb acceleration can be obtained in terms of $\boldsymbol{r}_{ICDorb}$ using Eq.5-3. Since the perturbations of the satellite orbit have been accounted for by the harmonic correction terms provided by the broadcast ephemeris, Eq.5-27 is capable of delivering accurate GPS satellite accelerations.

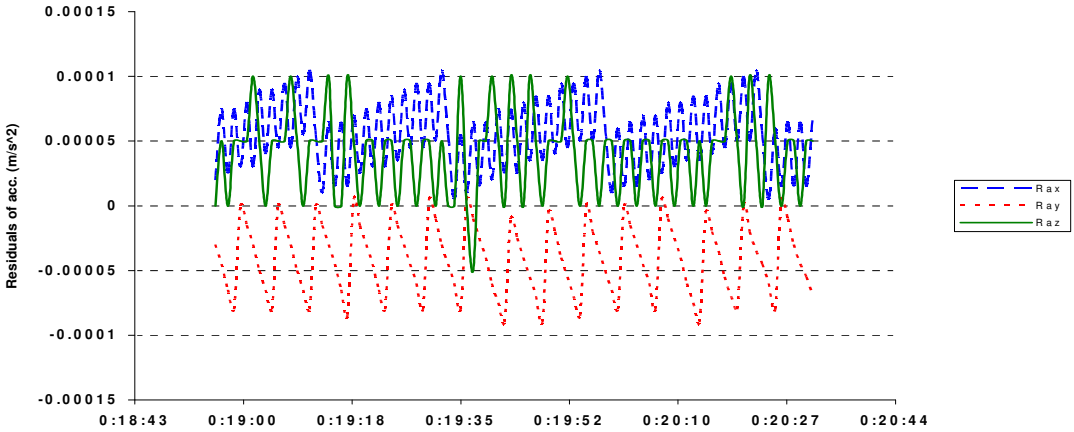

Figure 5- 3: Residuals of the accelerations from the closed-form formula and SP3 velocities using the first-order central difference of a Taylor series approximation (PRN=07, 22:18:56~22:20:32, 08/20/2002)

The acceleration from the IGS precise ephemeris is very accurate and thus can be considered as a "ground truth". The accelerations obtained using Eq.5-27 from the broadcast ephemeris have better than ±0.1mm/s^2 accuracy when compared with the accelerations from the SP3 precise ephemeris, which is evidenced by Fig.5-3 as well as Table 5-1 and Table 5-2. Thus, the derived formula can be confidently used in real-time applications. This overcomes the concerns raised by Kennedy [2003a] about the accuracy of acceleration determination using GPS due to the accuracy limitations of the broadcast ephemeris.

It is perhaps understandable that even the broadcast ephemeris is capable of achieving such high acceleration accuracy since the satellite motion in its high orbit is very stable; and mathematically the large radius of GPS orbit in the denominator of Eq.5-3 can greatly suppress the error propagation.

5.5 Alternatives to Get ECEF Satellite Velocity and Acceleration

It is well known that the satellite position from the broadcast ephemeris has an accuracy level of ±1m~5m. Due to the error propagation, the straightforward method of differentiating the ECEF positions to obtain the satellite ECEF velocity may lead to

large errors. The velocity could be too noisy because of the amplification due to the differentiation process. However, due to the complexity of the velocity algorithm, the position method was tested using the first-order central difference of a Taylor series approximation

$$\dot{\mathbf{r}}_{ECEF}(t) = \frac{\mathbf{r}_{ECEF}(t+\Delta t) - \mathbf{r}_{ECEF}(t-\Delta t)}{2\Delta t} \qquad 5\text{- }29$$

where t is the time of the calculation epoch, and Δt is the time interval, which is set as one second. Unexpectedly the results are close to the precise velocity of SP3 by better than ±1.0mm per second per axis, as reflected by Fig.5-4 which depicts the performance of this position differencing method. Once again the differentiation scenario works even in the satellite orbit.

The promising result of the ECEF satellite velocity obtained through differentiation is explained by the realisation that the errors associated with the orbit position of a satellite are not Gaussian white. The position may manifest systematic bias and correlated errors due to errors in the broadcast ephemeris, which have been significantly mitigated through differentiation.

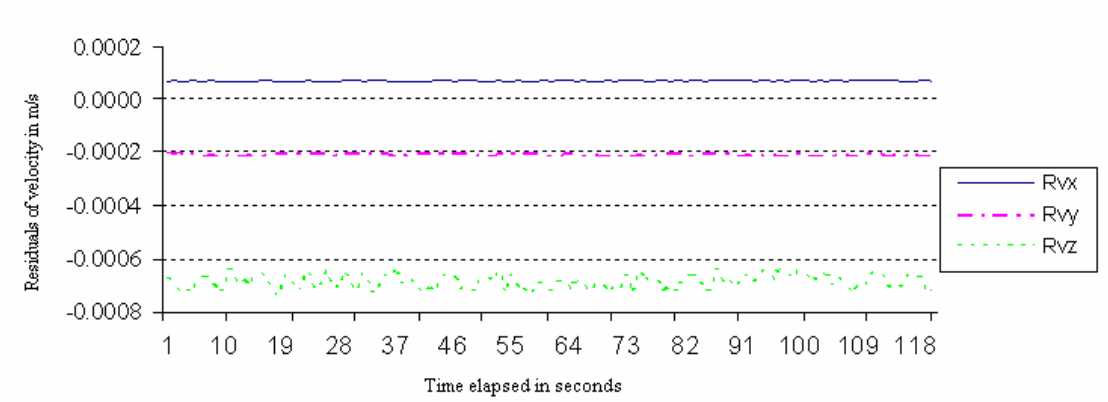

Figure 5- 4: Residuals of the position-differenced ECEF satellite velocities compared with the velocities from the SP3 precise ephemeris (PRN=07, 22:19:00~22:21:00, 08/20/2002)

Given the excellent performance of the position-differentiation method for ECEF velocity determination, it is sensible to test the performance of the differential method for acceleration determination, i.e.

$$\ddot{\mathbf{r}}_{ECEF}(t) = \frac{\dot{\mathbf{r}}_{ECEF}(t+\Delta t) - \dot{\mathbf{r}}_{ECEF}(t-\Delta t)}{2\Delta t} \qquad 5\text{- }30$$

where Δt is set to be again one second.

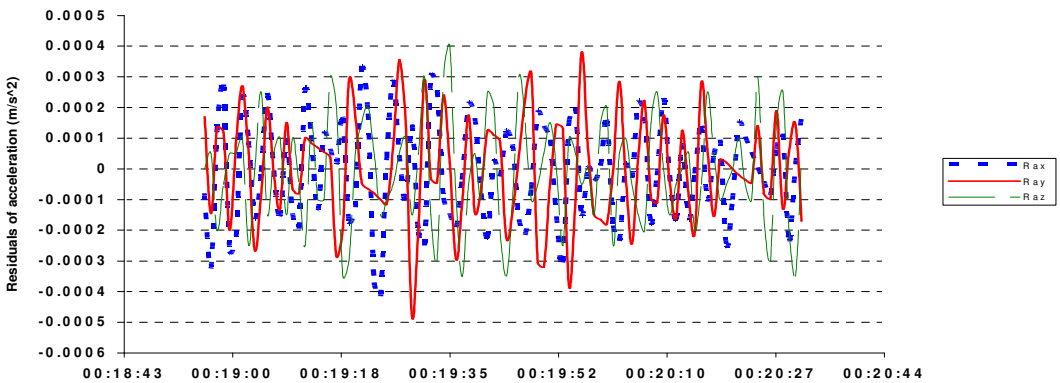

Figure 5- 5: Residuals of the position-differenced ECEF satellite accelerations and the accelerations obtained from the precise SP3 velocities using the first-order central difference of a Taylor series approximation (PRN=07, 22:18:56~22:20:32, 08/20/2002)

Figure 5-5 illustrates that the performance of the differentiation procedure is good. However, the quality of the results is inferior to the accelerations derived from the closed-form formula.

The differentiator used is a simple Finite Impulse Response (FIR) filter. The frequency response of the FIR differentiator is shown in Fig.5-6. It approximates the ideal differentiator for lower frequencies, which better suits the GPS satellite dynamics in a high orbit. Since even an ideal differentiator has a fixed phase delay, the central difference gives the velocity exactly at the desired epoch. Thanks to the high stability of the orbit of a GPS satellite, one can set Δt to be one second which gives ±0.5 as the filter coefficients, which can further suppress the errors associated

with the two positions. More details concerning differentiator design are given in Chapter Seven.

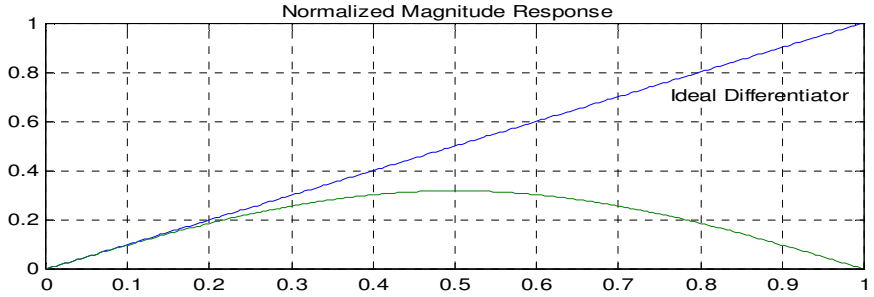

Figure 5- 6: Frequency response of the differentiator of the first-order central difference of a Taylor series approximation

Thus a practical alternative of satellite ECEF position, velocity and acceleration determination algorithm using the broadcast ephemeris is proposed as follows:

- Obtain positions of $\boldsymbol{r}_{ECEF}(t)$ and $\boldsymbol{r}_{ECEF}(t \pm \Delta t)$ by using the ICD-GPS-200c algorithm;
- Calculate velocity by using the position differentiation method, see Eq.5-29;
- Determine acceleration by using the closed-form rotation formula of Eq.5-27.

5.6 Polynomial Representation

There are many GPS applications where the position and velocity information is required to be output at a high sampling rate. In the sport project associated with this research, velocities and accelerations determined from GPS are to be output at an update rate of 10Hz. With advances in GPS receiver technologies, the sampling rate of GPS receivers has increased to as high as 100Hz. VBOX [2004] is a typical example of such receivers, which is designed for automobile breaking tests, where the positions and velocities of a car are output at a rate of 100Hz.

The determination of a GPS satellite position from the broadcast ephemeris requires cumbersome computations, where the Kepler's equation has to be resolved in a recursive way. Taking the VBOX receiver as an example, and assume that at every epoch, six GPS satellites have been tracked. Then in a one-second-observation period, the receiver needs to calculate 600 satellite positions and velocities, and to form and solve the 4×4 normal equation 2×100 times for position and velocity determination. With such a heavy computational load, it is desirable that the satellite position and velocity computation be significantly accelerated.

Since the movement of a GPS satellite is relatively stable and therefore predictable, polynomials can be used to represent GPS orbital positions. The satellite position at a given epoch within the representation period can be numerically interpolated. Remondi [1991] studied the polynomial representation of the orbit and concluded that it is sufficient for an accuracy of about 10^{-8} using a 30-minute interval and a 9th-order interpolator for the IGS precise ephemeris.

5.6.1 Lagrange Polynomials

Many polynomials can be used to represent satellite orbits. Among them, the Chebyshev and Lagrange polynomials are the two most frequently used. In this research, the Lagrange polynomial representation is adopted.

The general Lagrange polynomial is in the form of [cf. Xu, 2003]

$$y(t) = \sum_{j=0}^{m} L_j(t) \cdot y(t_j) \qquad 5\text{-}31$$

where:

$$L_j(t) = \prod_{k=0}^{m} \frac{(t - t_k)}{(t_j - t_k)}, \quad k \neq j \qquad 5\text{-}32$$

where

- $\prod$ is a multiplying operator from $k=0$ to $k=m$, m is the order of the polynomials;
- $y(t)$ is the given data at the time t;
- $L(t)$ is the base function of order m;
- t is the time at which data will be interpolated.

t should be placed in the middle of the time span (t_0, t_m) if possible, and m is usually selected as an odd number.

For the equal distance Lagrange interpolation

$$\begin{aligned} t_k &= t_0 + k\Delta t \\ t - t_k &= t - t_0 - k\Delta t \\ t_j - t_k &= (j-k)\Delta t \end{aligned} \qquad 5\text{-}33$$

where Δt is the data interval. Then a simplified form is given by

$$L_j(t) = \prod_{k=0}^{m} \frac{(t - t_0 - k\Delta t)}{(j-k)\Delta t}, \quad k \neq j \qquad 5\text{-}34$$

In the program developed in this research, the 9th-order Lagrange polynomials fitted with 5-minute intervals are used to represent GPS satellite positions derived from the broadcast ephemeris.

5.6.2 Two Schemes of Satellite PVA Representation

The simplest satellite PVA representation scheme uses three separated Lagrange polynomials to represent the satellite positions, velocities and accelerations. For the entire fitting period of t_0 to $t_0+m\cdot\Delta t$, the satellite position, velocity and acceleration are calculated at each fitting epoch. Since there are only a few fitting epochs, the rotation

matrix method of the satellite velocity calculation can be applied since the complexity of the algorithm doesn't matter in this case. This scheme has been adopted in this research.

The second scheme for satellite PVA representation is suggested by the success of ECEF velocity determination using the central difference of a Taylor series of the satellite positions. The success of the position differentiation method implies that, when the polynomial interpolation is employed to represent satellite positions, the precise satellite ECEF velocity can be directly obtained through numerical differentiation of the polynomial coefficients. In this scheme, it is sufficient to use only two sets of Lagrange polynomials, one representing the satellite positions and the other for acceleration representation.

5.7 Summary

This chapter supplements the ICD-GPS-200c algorithms for ECEF satellite position determination, with formulae for velocity and acceleration determination.

Both the conventional rotation method and the proposed positional differentiation method deliver better than ±1mm/s satellite velocity using the broadcast ephemeris. The proposed closed-form formula for ECEF satellite acceleration determination using the broadcast ephemeris produces ±0.1mm per second square accuracy when compared with the accelerations derived from the IGS precise ephemeris.

The polynomial representation of ECEF satellite PVA is discussed and the Lagrange polynomial interpolation is introduced for real-time and high output rate GPS applications. Two schemes for the PVA representation are proposed. It is concluded that when the polynomials are used for satellite position representation, the satellite velocity can be directly obtained through numerical differentiation of the polynomial coefficients.

Chapter Six

ERROR ANALYSIS AND MODELLING FOR DOPPLER SHIFT MEASUREMENTS

In the preceding chapters, the underlying principles of velocity and acceleration determination using GPS have been discussed by assuming that the errors in the GPS observables have been somehow accounted for. With the real-time calculation of satellite orbital velocity and acceleration in ECEF described, it is now appropriate to consider how to improve the real-time accuracy of velocity and acceleration determination using GPS in standalone mode.

There are over twenty kinds of errors and biases that have effects on a GPS range measurement. As the most important considerations in precise point positioning, these error effects have been identified, investigated and catalogued in §3.2. Since the carrier phase is the integrated Doppler shift over time, all errors in the carrier phase range measurement would have effects on the Doppler frequency shift measurement, through the first-order time derivatives of the source errors. This property will be exploited to analyse and model the inherent errors in the Doppler shift measurement, which forms the main content of this chapter. Many error correction formulae are derived and error elimination or mitigation methods are proposed. This chapter also discusses the error corrections for the Doppler rate "measurement".

6.1 Doppler Shift Observation Equation

In Chapter Four a highly accurate Doppler shift observation equation has been developed, with both the special and general relativity theories considered. The derived observation equation is in the form of (see Eq.4-41)

$$\lambda_i D_{r,i}^s(t) = (\mathbf{n}_r^s + \frac{\dot{\mathbf{r}}^s}{c}) \bullet [\dot{\mathbf{r}}^s(t-\tau_r^s) - \dot{\mathbf{r}}_r(t)] - d\dot{I}_{r,i}^s + d\dot{T}_r^s + c \cdot d\dot{t}_r(t) - c \cdot d\dot{t}^s(t-\tau_r^s) + \frac{\Phi_0 - \Phi(\mathbf{r}_r)}{c} + \frac{2GM}{c}(\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|}) + d\dot{R}_{Sagnac} + d\dot{R}_r^s + \varepsilon_r^s \qquad \text{6- 1}$$

For the purpose of error analysis, a modification of the above equation is more favourable. Analogous to the form of Eq.2-6, the errors in $\dot{\boldsymbol{r}}^s$ and $\dot{\boldsymbol{r}}_r$ as well as the multipath effects are considered. This leads to

$$\lambda_i D_{r,i}^s(t) = (\mathbf{n}_r^s + \frac{\dot{\mathbf{r}}^s}{c}) \bullet \{[\dot{\mathbf{r}}^s(t-\tau_r^s) + d\dot{\mathbf{r}}^s] - [\dot{\mathbf{r}}_r(t) + d\dot{\mathbf{r}}_r(t)]\} - d\dot{I}_{r,i}^s + d\dot{T}_r^s + c \cdot d\dot{t}_r(t) - c \cdot d\dot{t}^s(t-\tau_r^s) + \frac{\Phi_0 - \Phi(\mathbf{r}_r)}{c} + \frac{2GM}{c}(\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|}) + d\dot{R}_{Sagnac} + d\dot{R}_r^s + d\dot{m}_r^s + \varepsilon_r^s \qquad \text{6- 2}$$

where $d\dot{\boldsymbol{r}}^s$ and $d\dot{\boldsymbol{r}}_r$ can be interpreted as the first-order time derivative of the satellite positional error and the receiver positional error respectively.

6.2 Doppler Shift Error Analysis

The following sections discuss the error sources and error behaviours in the order of their appearance on the right hand side of Eq.6-2.

6.2.1 Line-of-Sight Correction

Line-of-sight correction is the correction applied to the receiver-satellite line-of-sight unit vector to account for the relativistic effect induced by the high speed of the satellite in its orbit. This is a special relativistic effect.

As can be seen from Eq.6-2, the "received" receiver-satellite range-rate differs from the line-of-sight direction by $\frac{\dot{\boldsymbol{r}}^s}{c}$. This correction changes with the satellite velocity over time, and the direction of the correction term coincides with the direction of satellite motion. Since it changes the line-of-sight direction toward the direction of movement, it is termed the line-of-sight correction.

Most of this correction is absorbed into the estimated receiver clock rate. The contribution of the line-of-sight correction to the receiver velocity estimation applies mainly in the vertical direction, since under good satellite geometry the horizontal errors might be averaged out. However, a bias of velocity estimation in the horizontal directions may be introduced when the satellite geometry is poor.

6.2.2 Satellite Velocity Error $d\dot{\boldsymbol{r}}^s$

$d\dot{\boldsymbol{r}}^s$, the first time-derivative of the satellite positional error, $d\boldsymbol{r}^s$, refers to the satellite velocity error due to satellite ephemeris. Since this research has been restricted to deal with real-time applications only, it is the error in the broadcast ephemeris. This has been discussed in Chapter Five.

Generally, the accuracy of ECEF satellite velocity from the broadcast ephemeris can be summarised as:

- In each axis, the velocity accuracy is better than ±0.5mm/s;
- The total accuracy is well within ± 1mm/s;

- There is a slight bias in each axis.

The above characteristics of the ECEF satellite velocity can be reflected by Fig.5-4, which is a typical representation of the accuracy of GPS orbit velocity. The biases will enter the determined velocity of a user, resulting in a biased solution.

6.2.3 Receiver Positional Error $d\dot{\mathbf{r}}_r$

The receiver error term $d\dot{\boldsymbol{r}}_r$, is the first-order time derivative of the receiver site displacement error which is due to the ocean load, atmospheric load, solid Earth tide and rotational deformation due to the polar motion. Section 3.2.3.2 gives the magnitude of each displacement. The magnitudes of their effects in the position domain are listed in Table 6-1.

Site Displacement	Maximum Magnitude	Length of the period
Solid Earth tides	12.5 cm in radial 5.0 cm in the north	combination of the long period, diurnal and semidiurnal periods
Ocean loading	M2: 5cm K1: 3cm	same as the solid Earth tide
Rotational deformation	0.7 cm in horizontal 2.5cm in vertical	same as the solid Earth tide

Table 6- 1: Maximum magnitudes and lengths of the periods of various site displacement errors

where *M2* and *K1* are the harmonic terms in the ocean loading. It can been seen from Table 6-1 that the site displacement components due to the solid Earth tide, ocean loading, and polar motion are associated with small magnitudes but very long periods. As a result their change-rates are negligible and consequentially have little effects on the instantaneous Doppler measurements. Therefore, it is concluded that the receiver site displacement error is negligible.

6.2.4 Ionospheric Error $d\dot{I}$

6.2.4.1 Dual Frequency Doppler Measurements

For a dual-frequency geodetic GPS receiver which is capable of outputting Doppler shift measurements on both frequency bands, the ionosphere-free linearised Doppler observable can be formed using the same coefficients as applied in the ionosphere-free carrier phase observable, i.e. [Rothacher et al., 1996:p.139]

$$D_{ion-free} \equiv \frac{f_1^2}{f_1^2 - f_2^2} D_1\lambda_1 - \frac{f_2^2}{f_1^2 - f_2^2} D_2\lambda_2 \approx 2.5457 D_1\lambda_1 - 1.5457 D_2\lambda_2 \qquad 6\text{-}3$$

Note that in the above ionosphere-free Doppler "measurement", the $d\dot{I}$ term has been deleted.

6.2.4.2 Dual-Frequency Receiver with L1 Doppler Measurement Only

For those dual-frequency receivers which make Doppler measurements on L_1 only, such as the Trimble 5700 GPS receivers, the ionospheric delay at the L_1 frequency can be precisely measured from the geometry-free carrier phase observable [Langley, 1998]. This delay, however, is biased by a constant due to the integer ambiguities, i.e.

$$\begin{aligned} dI_1 &= \frac{f_2^2}{f_1^2 - f_2^2}\left[(\lambda_1 N_1 - \lambda_2 N_2) + (\lambda_1\varphi_1 - \lambda_2\varphi_2)\right] + \varepsilon \approx 1.5456\left[(\lambda_1 N_1 - \lambda_2 N_2) + (\lambda_1\varphi_1 - \lambda_2\varphi_2)\right] \\ &= 1.5456[L_g + (\lambda_1 N_1 - \lambda_2 N_2)] \end{aligned} \qquad 6\text{-}4$$

where L_g is the geometry-free linear combination of L_1 and L_2 carrier phase observables. Thus, the ionospheric correction for Doppler measurements can be obtained by differentiating the above equation with respect to time. This leads to

$$d\dot{I}_1 = 1.5456\dot{L}_g \qquad 6\text{-}5$$

Figure 6-1 shows a 1.5456L_g observable time series of 200 seconds in an ionosphere-calm day, with the satellite elevation angle changing from 20.8º to 19.6º. It can be clearly seen that the ionosphere delay in the one-way observation is in a linear fashion, changing smoothly. There is about a ±5.0mm per second ionospheric rate correction for the Doppler shift measurements.

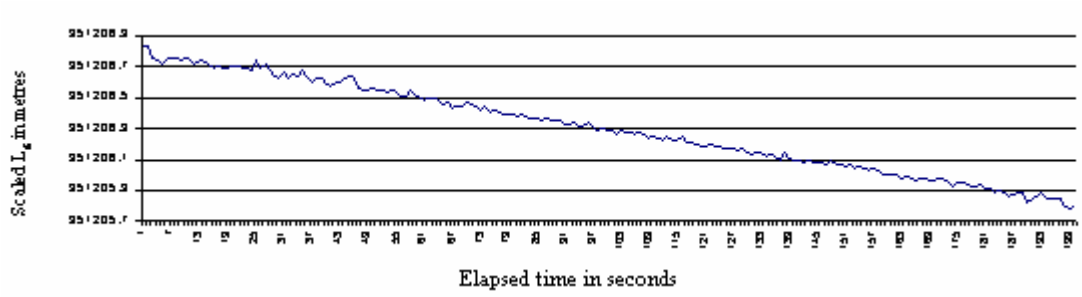

Figure 6- 1: A 1.5456 L_g 'observable' series of 200 seconds (elevation 20.8º~19.6 º)

With a magnitude as large as such, the ionospheric correction is critical for precise velocity determination at an accuracy level of sub-centimetre per second.

The ionospheric delay rate $d\dot{I}$ can be numerically obtained by using a differentiator with L_g time series as the filter input. Since the ionosphere changes smoothly and almost linearly, a low-pass differentiator is sufficient for this purpose. Details of differentiator design are given in Chapter Seven.

6.2.4.3 Single-Frequency Receiver

There are two methods to obtain the ionospheric correction $d\dot{I}$ for single-frequency GPS receiver. One is based on the actual code and carrier phase measurements; the other method is to use the standard Klobochar model.

Code-Carrier Method

This method takes advantage of the property of GPS signal propagation in the ionosphere. The different sign but equal magnitude of the group delay and the phase advance makes it possible to eliminate the ionosphere effect by a summation of the code and carrier phase measurements, as described in § 3.2.2.1.1. This characteristic also makes it possible to retrieve the ionospheric delay by a substraction of the code and carrier phase measurements.

Subtracting Eq.2-5 from Eq.2-4 and neglecting the difference in magnitude of the multipath effects between the code and carrier phase measurements leads to

$$P_1(t) - \lambda_1\varphi_1(t) = 2 \cdot dI + \lambda_1 N_1 + \lambda_1\varphi_1(0) \qquad \text{6- 6}$$

where N_1 is the integer ambiguity value and $\varphi_1(0)$ is the receiver initial phase bias. Note that both of them are constants.

Multiplying 0.5 to both side of Eq.6-6 and then differentiating with respect to time yields the desired $d\dot{I}$ as

$$d\dot{I} = 0.5 \cdot \frac{d}{dt}\left[P_1(t) - \lambda_1\varphi_1(t)\right] \qquad \text{6- 7}$$

Compared to Eq.6-7 and Eq.6-5, it is clear that the ionospheric delay rate derived from the geometry-free carrier phase combination has much higher accuracy. However, the code-carrier method is affected by the lower precision of the PRN code measurement and the presence of code multipath effect. Since the ionospheric delay changes smoothly and linearly during a short time period, $d\dot{I}$ can be obtained from a simple numerical differentiator with either the L_g or code-carrier time series as the filter input.

Standard Klobuchar Model

The Klobuchar model is an empirical ionospheric model developed for GPS single-frequency receivers to correct for approximately 50% of the ionospheric delay error [Klobuchar, 1996]. The GPS control segment has adopted it as the standard ionospheric model and broadcasts the 8 model parameters in the navigation message to GPS users.

The ionospheric delay rate can be analytically derived as

$$d\dot{I}_1 = \begin{cases} c \cdot F \cdot AMP \cdot \dot{x} \cdot \left(-x + \dfrac{x^3}{6} \right) & if \quad |x| < 1.57 \\ 0 & if \quad |x| \geq 1.57 \end{cases} \qquad 6\text{-}8$$

where according to Eq.A-3 in the Appendix, $\dot{x}$ can calculated as

$$\dot{x} = \frac{2\pi}{PER} \qquad 6\text{-}9$$

Details about the terms and algorithms can be found in the Appendix. In Eq.6-8, taking the ionospheric change-rate as zero when |x|≥1.75 is an approximation by neglecting the change-rate of the elevation angle. For brevity, no further explanation is provided here.

Alternatively and preferably, the ionospheric delay can also be numerically obtained by a differentiator with the time series of the calculated ionospheric corrections as the filter input.

6.2.5 Tropospheric Error $d\dot{T}$

Most reference books suggest that the contribution of the troposphere to the Doppler measurement is so small that it can be neglected. However, according to Simsky and

Boon [2003], not properly accounting for the tropospheric delay rate could result in nearly centimetre per second level noise and a significant bias.

A tropospheric delay model for range measurement corrections consists of a hydrostatic and a non-hydrostatic component. The delay can be modelled with a zenith delay model and a mapping function, in the form of

$$dT = dT_{dry}^{zenith} \cdot m_{dry} + dT_{wet}^{zenith} \cdot m_{wet} \qquad 6\text{-}10$$

The dry component contributes 90% of the total tropospheric delay while the wet component contributes approximately 10%. The tropospheric delays in the zenith direction, according to Kouba and Heroux [2001], vary in time by a relatively small amount, of the order of a few centimetres per hour and thus can be viewed as a constant for Doppler measurement. Hence it is sufficient to choose a standard model such as the Saastamoinen model [Saastamoinen, 1972] based on standard meteorological data to represent the zenith delays. In this case, what really matters is the adoption of mapping functions.

Ordinary mapping functions in simple cosecant E (elevation) forms are unable to reflect the tropospheric change-rate since the elevation angle changes very slowly, at a rate of about ±0.1millirad per second due to the slow change of satellite geometry [Simsky and Boon, 2003]. This requires that precise geodetic mapping functions should be used for the tropospheric delay rate modelling.

The global mapping function developed by Niell [1996] may be used to model the tropospheric delay rate. This model is precise even when the elevation angle is down to 3º to 5º; however its calculation is complicated. In kinematic applications, an elevation cut-off angle of 15º is normally set. Up to this elevation angle, other geodetic mapping functions such as Chao's mapping functions [Chao, 1974] may be considered

$$m_d(E) = \frac{1}{\sin E + \dfrac{0.00143}{\tan E + 0.0445}}$$
$$m_w(E) = \frac{1}{\sin E + \dfrac{0.00035}{\tan E + 0.017}} \qquad 6\text{- }11$$

Chao's formulae are simple in form and independent upon location and height. Since no meteorological data is required, these mapping functions are attractive for kinematic applications.

One may numerically obtain the tropospheric delay change-rate after applying the tropospheric correction to the range measurement, through a simple differentiator. That is, differentiate the calculated tropospheric delays with respect to time

$$d\dot{T} = \frac{d}{dt}(dT) \qquad 6\text{- }12$$

This avoids the computation of the bulky expressions of the derivatives of the mapping functions, cf. Simsky and Boon [2003]. As the tropospheric delay changes slowly, a low-pass differentiator is sufficient for this purpose.

Figure 6-2 illustrates the tropospheric delay rate from Chao's model. In this simulation, the vertical dry delay is set as 2.4m, and the vertical wet delay is set as 0.24m. This resembles the maximum tropospheric effect on GPS signals. Figure 6-2 shows the importance of troposphere modelling since even at higher elevation angles there are still contributions in millimetres per second level to the Doppler measurements. Thus, the tropospheric delay rate is a main error factor limiting precise velocity determination using GPS.

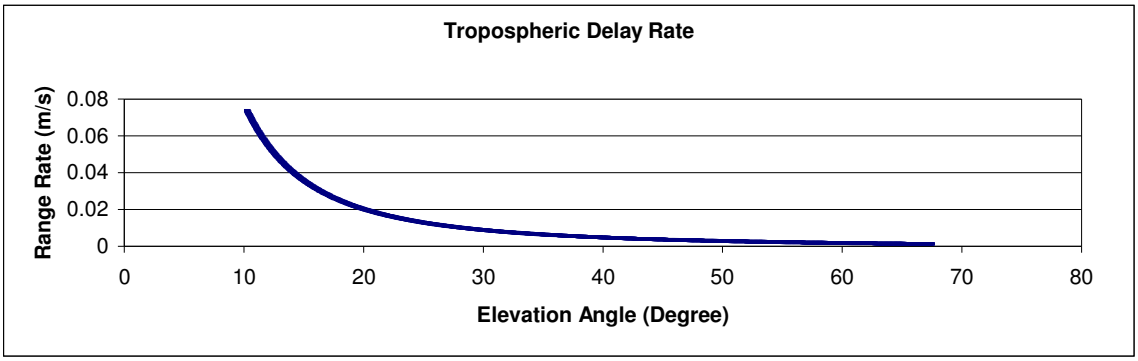

Figure 6- 2: Tropospheric delay rate from Chao's model

6.2.6 Clock Rate Corrections

6.2.6.1 Receiver Clock

Similar to navigation point positioning, the clock rate becomes the fourth unknown and is solved for along with the receiver velocity unknowns if at least four Doppler shift measurements are observed. A GPS receiver clock is typically a quartz crystal oscillator, thus the clock rate stability depends on the quality of the receiver's crystal oscillator. The errors introduced through the receiver oscillator affect the quality of the measured Doppler shifts, and therefore degrade the velocity estimations. Fortunately, with improvements in technology many quartz crystal oscillators used in GPS receivers are capable of providing relatively high short-term stability, which results in accurate Doppler shift measurements.

6.2.6.2 Satellite Clock

In the ICD-GPS-200c, the satellite PRN code phase time offset is given by

$$dt^s = a_{f0} + a_{f1}(t - t_{oc}) + a_{f2}(t - t_{oc})^2 \qquad 6\text{- }13$$

where:

- a_{f0}, a_{f1}, a_{f2} are the polynomial coefficients in sub-frame one of the navigation message;
- t_{oc} represents the clock data reference time in GPS seconds.

From Eq. 6-13, the satellite clock rate correction for the Doppler measurement can be derived as

$$\dot{dt}^s = a_{f1} + 2a_{f2}(t - t_{oc}) \qquad 6\text{-}14$$

Another error associated with the satellite clock is the satellite group delay T_{gd}, which is due to the satellite hardware bias. As a constant it cancels by differentiation with respect to time. Hence, the satellite group delay has no effect on Doppler measurements.

6.2.7 Receiver Potential Difference

The term $\frac{\Phi_0 - \Phi(\boldsymbol{r}_r)}{c}$ is a site-dependent general relativistic correction term attributed to the gravity potential of the receiver. It becomes zero when a receiver is located at the geoid.

The receiver potential difference is a common term in each Doppler shift measurement at a specific epoch and thus contributes to the user velocity in the same way as the receiver clock rate. Neglecting the receiver potential difference will cause a biased receiver clock rate estimate, but do little harm to the user velocity estimation. This may be the reason why this term has been historically neglected for velocity determination using GPS.

The potential difference can be simply estimated by the approximation $\Phi_0 - \Phi(\boldsymbol{r}_r) = gH_r$ where g is the gravity acceleration and H_r is the height of the receiver. The maximum effect of this term on the Earth's surface is at the peak of Mt.

Everest, which is less than 0.3mm/s. The magnitude of this term varies with the receiver position (mainly in height); it increases for aviation and space applications.

6.2.8 Orbital Eccentricity Correction

The orbital eccentricity correction is a general relativistic correction term. This relativistic correction for the range of receiver *r* and satellite *s* depends on the satellite orbit. The alternative expression of the orbital eccentricity correction for range measurement is in terms of the satellite position and velocity [ARINC, 2000, p.89]

$$dR_{eccentricity} = -\frac{2}{c} \cdot \mathbf{r}^s \bullet \dot{\mathbf{r}}^s \qquad 6\text{- }15$$

This correction is not sensitive to whether the vectors of position and velocity are expressed in an ECEF rotating coordinate system or in an ECI coordinate system (ibid). The eccentricity correction for range measurement is a periodic function. It is the biggest correction due to GPS relativity (see § 3.2.1.3.1).

Simply differentiating the above equation with respect to time gives

$$d\dot{R}_{eccentricity} = -\frac{2}{c} \cdot \left[(\mathbf{r}^s)^2 + \mathbf{r}^s \bullet \ddot{\mathbf{r}}^s \right] \qquad 6\text{- }16$$

where $\ddot{\boldsymbol{r}}^s$ is the satellite acceleration. Since a satellite conserves energy while in orbit, according to the Vis-Viva equation [Montenbruck and Gill, 2000, p.20]

$$(\dot{\mathbf{r}}^s)^2 = GM\left(\frac{2}{\left\|\mathbf{r}^s\right\|} - \frac{1}{a_{orb}}\right) \qquad 6\text{- }17$$

and the satellite acceleration is, for example [Misra and Enge, 2001, p.99]

$$\ddot{\mathbf{r}}^s = -\frac{GM}{\left\|\mathbf{r}^s\right\|^3}\mathbf{r}^s \qquad 6\text{-}18$$

Substituting Eq.6-17 and Eq.6-18 into the brackets of Eq.6-16 yields

$$d\dot{R}_{eccentricity} = -\frac{2}{c}\cdot[(\dot{\mathbf{r}}^s)^2 + \mathbf{r}^s\cdot\ddot{\mathbf{r}}^s] = \frac{2GM}{c}(\frac{1}{a_{orb}} - \frac{1}{\left\|\mathbf{r}^s\right\|}) \qquad 6\text{-}19$$

This proves that the term $\frac{2GM}{c}(\frac{1}{a_{orb}} - \frac{1}{\left\|\boldsymbol{r}^s\right\|})$ is the relativistic correction for the orbit eccentricity. Here the term "orbit eccentricity correction" is an intuitive term, since this term vanishes for a circular orbit where $a_{orb} = \left\|\boldsymbol{r}^s\right\|$.

6.2.9 Earth Rotation Correction (Sagnac Effect)

As described previously, the rotation of the Earth during the GPS signal propagation period causes another relativistic error that is known as the Sagnac effect. The incoming signal has an extra signal passage than would otherwise be the case.

The Sagnac Effect correction given by Ashby and Spilker [1996] is in the form of

$$dR_{Sagnac} = \frac{2\boldsymbol{\Omega}_e\cdot A}{c} \qquad 6\text{-}20$$

where $\boldsymbol{\Omega}_e=(0,0,\ \omega)^T$ is the vector of the angular rate of the Earth's rotation and $\boldsymbol{A}$ is the shading area of the triangles formed by the Earth's centre, the receiver and the GPS signal. As such, the Sagnac correction for range measurement can be calculated using

$$dR_{Sagnac} = \frac{2\boldsymbol{\Omega}_e^{\ T}}{c}\cdot\frac{\mathbf{r}_r\times\mathbf{r}^s}{2} = \frac{\omega}{c}(y^s x_r - x^s y_r) \qquad 6\text{-}21$$

which is identical to the aberration correction formula given by Seeber [2003, p.198].

With this, the Sagnac correction for Doppler can then be obtained by differentiating the above equation with respect to time. This leads to

$$d\dot{R}_{Sagnac} = \frac{\mathbf{\Omega}_e^{\ T}}{c} \cdot \frac{d}{dt}(\mathbf{r}_r \times \mathbf{r}^s) = \frac{\omega}{c}[\dot{x}_r y^s - \dot{y}_r x^s + x_r \cdot \dot{y}^s - y_r \dot{x}^s] \qquad 6\text{-}22$$

This correction is generally very small, but the maximum may reach several millimetres per second and therefore requires consideration.

It's worthwhile to point out that whether or not to apply the Sagnac correction in a real GPS application depends on different treatments of the signal propagation delay due to the Earth's rotation [Ashby and Spilker, 1996, p.675], see also the discussion in § 3.2.1.3.2. There is no need to apply the Sagnac effect correction if the real signal transmission time is determined recursively and the rotation has been applied to the satellite ECEF position at the signal transmission epoch.

6.2.10 Secondary Relativistic Effects

For integrity, the secondary relativistic Doppler correction term is included in Eq.6-2. This term corresponds to the first derivatives of the secondary relativistic correction terms for the GPS range measurement (see § 3.2.1.3.3).

Under this category, significant relativistic corrections for range measurements include the signal propagation delay, spatial curvature, satellite antenna offset, phase wind-up correction, and even the effects from other bodies in the solar system.

The secondary relativistic effects on the instantaneous Doppler shift measurement are too small to sense. This is due to the small magnitudes and the long periods associated with them. Their effects are therefore neglected.

6.2.11 Multipath Effect

Thus far, no significant multipath effects on velocity determinations have been observed; even though some of the field trials were conducted on water, thought to be a high multipath environment [Zhang et al., 2003b]. Nor have other researchers reported any significant deterioration in velocity determination due to multipath. According to Serrano et al. [2004], ±1cm/s velocity accuracy was achieved despite multipath-rich conditions.

Theoretically the multipath effect reaches a maximum of one quarter of a cycle for the phase observable, and changes periodically [Hofmann-Wellenhof et al., 2001]. This may indicate that the multipath errors associated with the Doppler shift would have been averaged out, since the Doppler frequency shift is the first-derivative of the carrier phase with respect to time. However, further research is needed before drawing a decisive conclusion.

6.3 Numerical Analysis

A numerical analysis of each of the errors that have been discussed has been carried out, and the results are summarised in Table 6-2.

	Error Terms	Apply Correction for Doppler Measurements	Magnitude Estimated
Satellite Orbit	Broadcast Ephemeris	Optional	±1mm/s
Satellite Clock	Satellite Clock Correction	Yes	Negligible
Relativity	L_1-L_2 Correction (Group delay)	No	Negligible
	Orbit Eccentricity	Yes	Several cm/s
	Sagnac	Optional**	Several mm/s
	Receiver Potential Difference	Yes	Sub-mm/s
	Secondary Relativistic Effects	No	Negligible
Atmosphere	Ionospheric Correction	Yes	mm/s ~ cm/s
	Tropospheric Correction	Yes	mm/s ~ cm/s
Receiver	Receiver Site Displacement	No	Negligible
	Multipath Correction	No	Negligible
	Receiver Clock	As an unknown to be estimated	

*(**: Whether or not to apply the Sagnac correction depends on different treatments of the propagation delay due to the Earth's rotation)*

Table 6- 2: Main error sources and their estimated magnitudes in Doppler measurements

It can be seen that the relativity effect is the largest error source for the Doppler shift measurements. Fortunately, it can be well modelled and corrected for. The atmosphere effects are relatively small. However, due to their unpredictable nature they are hard to model and therefore critical to sub-centimetre per second accuracy velocity determination.

6.4 Errors in Doppler Rate "Measurement"

The Doppler rate "observation" equation is rewritten as

$$\lambda_i \dot{D}_i(t) = \ddot{\rho}_r^s = \mathbf{n}_r^s \bullet \ddot{\mathbf{r}}_r^s + \frac{1}{\rho_r^s} \cdot \left[(\dot{\mathbf{r}}_r^s)^2 - (\dot{\rho}_r^s)^2 \right] - d\ddot{I}_{r,i}^s + d\ddot{T}_r^s + c \cdot d\ddot{t}_r(t)$$
$$- c \cdot d\ddot{t}^s(t - \tau_r^s) + d\ddot{M}_r^s - d\ddot{R}_r^s + \varepsilon \qquad \text{6- 23}$$

It can be seen that in the Doppler rate observable, errors in both the satellite acceleration and satellite velocity have some effects. In Chapter Five, by comparing the satellite accelerations derived from the broadcast ephemeris with those derived from the precise IGS SP3 ephemeris, it was concluded that the accuracy of satellite ECEF accelerations is of the order of ±0.1mm/s^2 per axis. This level of accuracy is rather high, though one should be aware that the magnitude of satellite ECEF acceleration tends to be small owing to the stability of GPS satellites in their high orbits.

The relative receiver-satellite velocity is coupled with the Doppler rate observable, thus the inherent errors in both the satellite velocity and the receiver velocity would affect the Doppler rate "measurement". However, these errors are tolerable due to the long distance between the receiver and satellite, since it is in the denominator.

The other errors that affect the Doppler shift rate "measurement" are the second-order derivatives of the corresponding range errors. As has been demonstrated, the change-rates of the range errors are generally small, varying from a few millimetres to several centimetres, and therefore the rate of their change-rates are generally so small that they can be neglected.

The only exception is the satellite clock drift, which is one of the navigation parameters, and should be taken into account. The correction formula is given by differentiating Eq. 6-14 with respect to time

$$d\ddot{t}^s = 2a_{f2} \qquad \text{6- 24}$$

However, in most cases, the broadcast a_{f2} is zero. Thus, this virtual observable is rather "clean".

6.5 Summary

A thorough investigation and analysis of the inherent errors of the Doppler shift measurement have been carried out in this chapter. The error correction formulae and methods have been proposed.

From the error analysis, it is demonstrated that the relativistic effects are the largest error sources in precise velocity determination using GPS. The explicit relativistic correction formulae presented in this chapter are ready to be applied. For those interested in the highest velocity accuracy that GPS can provide, the relativistic corrections must be considered.

The ionospheric delay rate correction is provided for dual-frequency GPS receivers, by either forming the ionosphere-free Doppler 'observable' or obtaining the first-derivative of the geometry-free carrier phase 'observables'. For single-frequency GPS users, the differentiator methods based on both the code-carrier measurements and the empirical standard model are proposed. However, the accuracy of both methods is limited.

The tropospheric delay rate correction has been identified as requiring the appropriate mapping functions that may best represent the change of tropospheric delay along signal profiles. This requires that precise geodetic mapping functions be used.

Since both ionospheric and tropospheric delay rates contribute several millimetres per second level errors to Doppler measurements, and since their changes are hard to predict and to model, the atmosphere becomes the major error source degrading the velocity accuracy using GPS.

If all the inherent errors in Doppler measurement are properly accounted for, millimetres per second levels of velocity accuracy are achievable in real time for standalone GPS users. Much higher accuracy can be achieved in post processing mode using the IGS products. However, this depends on the quality of the available Doppler shift measurements.

It is demonstrated that the virtual "Doppler rate" measurement is rather clean. Only the satellite clock correction is required to be applied in the virtual "Doppler rate" observable. A receiver's acceleration can be determined to a relatively high accuracy as long as the precise "Doppler rate" observables can be derived, either from the Doppler shift measurements or from the accurate carrier phase measurements. This inevitably requires the design of different differentiators, which is the topic of Chapter Seven.

Chapter Seven

DIFFERENTIATOR DESIGN

Previously proposed methods for GPS velocity and acceleration determination fall into two categories: one is to derive velocity and acceleration directly from GPS determined positions, the other is based on the use of Doppler shift measurements. The latter has several advantages: it doesn't rely on the precision of the positions derived from GPS, nor will the accuracy dramatically degrade with an increase of sampling rate (say 10Hz or more). Since there is no direct Doppler rate observable in GPS measurements, as a "virtual" observable it must be derived in order that Eq.4-44 [Jekeli and Garcia., 1997] can be applied directly in the Doppler shift method.

Every GPS receiver measures Doppler shifts, however, primarily as an intermediate process to obtain accurate carrier phase measurements. Thus, the quality of Doppler shift output varies from receiver to receiver. The Trimble 5700™ geodetic receiver, for example, has a measurement precision of ±1mm/s. The observed Doppler is obtained from a tracking loop that is updated at a very high rate. This also enables the receiver to sense the phase acceleration [Harvey, 2004]. Unfortunately the sensed phase acceleration and the Doppler shift on L2 are discarded. Some other GPS receivers, for example the Superstar II™ from NovAtel, have only C/A code and L1 phase outputs [SuperstarII, 2004]; and the Doppler shifts are not output. To obtain accurate velocity and acceleration using these types of receivers, it is necessary to derive the Doppler shifts, i.e. the change-rates of the carrier phase from the measured carrier phase measurements.

Differentiators are required to obtain the Doppler rate "observable" for any type of receiver, or to generate the Doppler shift from the carrier phase measurements. In real-time and dynamic applications it is also desirable that the differentiator should have a wideband frequency response to cover the system dynamics and have a group delay as short as possible in order that the Doppler shift or Doppler rate can be derived instantaneously. For those receivers that output only "raw" Doppler shifts, the derivation of precise Doppler from the carrier phase plays a key role in precise velocity and acceleration determination since the precision of the carrier phase observables can be fully exploited.

Differentiators are also required to derive the tropospheric delay rate and the ionospheric change-rate, as stated previously in Chapter Six. The objective of this chapter is to explore approaches to derive the Doppler rate from GPS measurements, or to derive the precise Doppler shift from the carrier phase in real time and dynamic situations.

Several investigations have been conducted for this purpose in the GPS measurement domain, and the proposed methods can be categorised into:

- Curve fitting [Fenton and Townsend, 1994];
- Kalman smoother/filtering [Hebert et al., 1997];
- Taylor series approximation [Cannon et al., 1998; Bruton et al., 1999];
- Finite Impulse Filter (FIR) filter by using Fourier series with window techniques [Bruton et al., 1999];
- FIR optimal design using the Remez exchange algorithm (ibid). The FIR filtering technique based on the Taylor series approximation was recently adopted by Kennedy [2003] to derive phase accelerations.

This chapter briefly introduces the digital differentiator theory and describes the design problem in real-time dynamic GPS applications. It is followed by a comprehensive review on each method mentioned above. By comparing the various differentiator designs, a series of first-order Infinite Impulse Filters (IIR) are presented which are capable of delivering the derivatives from input signals in real-time dynamic situations. An adaptive scheme is also proposed for noise attenuation.

7.1 The Ideal Differentiator

7.1.1 Digital Filtering

To introduce the digital filter concept, GPS is used as an example. Signals from GPS satellites are measured and recorded at the sampling epochs by a GPS receiver. It is assumed that the sampled date have equal time intervals. The recorded GPS data are discrete measurements of C/A code, L1 carrier phase and D1, etc. In other words, rather than continuous GPS signals for a receiver-to-satellite pair, there is a discrete time series of C/A code ranges, a discrete time sequence of L1 carrier phase measurements, and a discrete sequence of the Doppler shift measurements. These data may be equally spaced; for example, with a time interval T of 0.1s if the GPS receiver has a 10Hz sampling rate.

Suppose that a sequence of x_n are such a set of equally spaced *L1* carrier phase measurements from the continuous *L1* signal of $x(t)$, where n is an integer and t denotes the continuous variable ($t=n\cdot T$), then a digital filter can be defined as follows [Hamming, 1977, p.2]

$$y_n = \sum_{k=-\infty}^{k=\infty} c_k x_{n-k} + \sum_{m=1}^{M} d_k y_{n-m} \qquad 7\text{-}1$$

where c_k and d_k are the coefficients which are referred to as the impulse response of the filter. The impulse response of a filter is denoted as $h(n)$, which is the filter response for a unit input signal pulse (ibid).

The digital filter can be taken as a linear combination of equally spaced samples x_{n-k} of some function $x(t)$ together with the computed values of the output y_{n-k}. By setting different coefficients of the filter, it can be used to selectively suppress or enhance particular parts of the signals.

Note that Eq.7-1 is in a generic form. Without the second term on the right hand side, the filter is referred to as a non-recursive filter. With the second term on the right hand side, the filter is called recursive since the output of y_{n-k} has been recursively used. The filter coefficients c_k and d_k are taken as time-invariant, i.e. constant in the case of classical filter design. However, their values can be varied to achieve a desired filtered result. This is referred to as adaptive filter design. For practical applications, the length of a realisable digital filter is always finite.

7.1.1.1 Transfer Function

The transfer function allows us to describe a filter by means of a convenient and compact expression. The Z-transform has been popularly used to determine the transfer function of filters. With the transfer function, the characteristics of a filter can be analysed, such as its frequency and amplitude responses.

Z-Transform

The Z-transform of a discrete-time signal value of x_n is defined as

$$Z(x_{n-k}) = X_{n-k}(z) = z^{-k} x_n \qquad 7\text{-}2$$

where z is a complex variable. The Z-transform is a linear transformation. For a causal discrete filter n is normally defined to begin at time $n=0$. One important property of the Z-transform is that the Delay Theorem allows z^{-1} to relate the Z-transform of the current input value x_n to the Z-transform of the previous input value x_{n-1} by

$$z^{-1} \cdot Z(x_n) = Z(x_{n-1}) \qquad \Leftrightarrow \qquad X_{n-1}(z) = z^{-1} X_n(z) \qquad 7\text{-}3$$

Therefore z^{-1} serves as a unit delay operator. For a filter, for example having the form of

$$y_n = a_0 x_n + a_1 x_{n-1} + a_2 x_{n-2} - b_1 y_{n-1} - b_2 y_{n-2} \qquad 7\text{-}4$$

the Z-transform of the filter is

$$\begin{aligned} Y(z) &= a_0 Z(x_n) + a_1 Z(x_{n-1}) + a_2 Z(x_{n-2}) - b_1 Z(y_{n-1}) - b_2 Z(y_{n-2}) \\ &= a_0 Z(x_n) + a_1 z^{-1} Z(x_n) + a_2 z^{-2} Z(x_n) - b_1 z^{-1} Z(y_n) - b_2 z^{-2} Z(y_n) \\ &= a_0 X(z) + a_1 z^{-1} X(z) + a_2 z^{-2} X(z) - b_1 z^{-1} Y(z) - b_2 z^{-2} Y(z) \end{aligned} \qquad 7\text{-}5$$

One can just take the Z-transform as the discrete-time cousin of the continuous Laplace transform. z^{-n} is the general form for the solution of the linear difference equation similar to the function e^{-sT} (where s is the complex variable associated with the Laplace Transform, and T is the sampling period of the ideal impulse sampler) which is the general form for the solution of the linear differential equation [Lyons, 2004, p.229].

Transfer Function of Filter

The transfer function of a discrete filter is defined as the rate of the Z-transform of the filter output signal over the Z-transform of the input signal, i.e.

$$H(z) \equiv \frac{Y(z)}{X(z)} \qquad 7\text{-}6$$

Therefore, the transfer function of the filter that has the form of Eq.7-4 is

$$H(z) = \frac{Y(z)}{X(z)} = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}} \qquad 7\text{-}7$$

The transfer function of a digital filter is the most important aspect of filter design and analysis. With the transfer function of a filter determined, one can directly obtain the impulse response of the filter, and analyse the performance of the filter either in the state space or in the frequency domain.

7.1.1.2 Frequency and Amplitude Response

The frequency response of a filter is defined as the spectrum of the output signal divided by the spectrum of the input signal, it can be easily evaluated in the unit circle, i.e. $H(e^{j\omega t})$ where the variable z in the transfer function is replaced by $e^{j\omega t}$.

Accordingly the frequency response of the above example filter is

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} = \frac{a_0 + a_1 e^{-j\omega} + a_2 e^{-2j\omega}}{1 + b_1 e^{-j\omega} + b_2 e^{-2j\omega}} \qquad 7\text{- }8$$

For a 5-point averaging filter in the form

$$y = \frac{1}{5}\sum_{m=-2}^{2} x_{n-m} \qquad 7\text{- }9$$

the frequency response is

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} = \frac{1}{5}(e^{-2j\omega} + e^{-j\omega} + 1 + e^{j\omega} + e^{2j\omega}) = \frac{1}{5}(1 + 2\cos\omega + 2\cos 2\omega) \qquad 7\text{- }10$$

It can be deduced that the frequency response of a digital filter is actually the discrete Fourier transform of the impulse response, i.e. (see Eq.7-1)

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} = \frac{\sum_{k=-\infty}^{k=\infty} c_k \cdot e^{-j\omega k}}{1 - \sum_{m=1}^{M} d_m \cdot e^{-j\omega m}} \qquad 7\text{- }11$$

The amplitude response of the filter is defined as the magnitude of the frequency response, i.e.

$$amplitude\ response = \|H(\omega)\| \qquad 7\text{- }12$$

This allows the amplitude response of the filter to be analysed along with the frequency response. In practice it is more efficient to factor the frequency response into the form of

$$\|H(\omega)\| = G(\omega) \cdot e^{j\Theta(\omega)} \qquad 7\text{- }13$$

where $G(\omega)$ is the gain of the amplitude response and $\Theta(\omega)$ is the phase response. The phase response $\Theta(\omega)$ of a digital filter gives the radian phase shift experienced by each sinusoidal component of the input signal.

It is often more intuitive to consider the phase delay defined by

$$phase \quad delay = \frac{\Theta(\omega)}{\omega} \qquad 7\text{- }14$$

which gives the time delay in seconds experienced by each sinusoidal component of the input signal. There is a more commonly encountered representation of the phase response, which is called the group delay, defined by

$$group \quad delay = \frac{d}{d\omega}\Theta(\omega) \qquad 7\text{- }15$$

In the case of a filter that has a linear phase response, the group delay and the phase delay are identical, i.e. $\Theta(\omega) = \frac{\pi}{2}\omega$.

7.1.1.3 Noise Amplification

A digital filter is a linear combination of the input signals that are usually corrupted by noises. For simplicity, the noises are assumed Gaussian white, and hence the error propagation law applies. This permits the estimation of the noise amplification of the filter. Assume the noise of a series of L1 carrier phase observables of $x_n = x_n^0 + e_n$ to be Gaussian white, i.e.

$$E(e_n, e_m) = \begin{cases} \sigma_x^2 & m = n \\ 0 & m \neq n \end{cases} \quad and \quad E(e_n) = 0 \qquad 7\text{- }16$$

and the outcome of a non-recursive filter is

$$y_n = \sum_{k=-K}^{k=K} c_k x_{n-k} = \sum_{k=-K}^{k=K} c_k x_{n-k}^0 + \sum_{k=-K}^{k=K} c_k \varepsilon_{n-k} \qquad 7\text{- }17$$

then the variance of the filter can be calculated [Hamming, 1977, p.14]

$$E\left\{\sum_{k=-K}^{k=K} c_k \varepsilon_{n-k}\right\}^2 = \sum_{k=-K}^{k=K} c_k^2 E\{\varepsilon_{n-k}\}^2 = \sigma_x^2 \sum_{k=-K}^{k=K} c_k^2 \qquad 7\text{- }18$$

This shows that the sum of the squares of coefficients of the filter determines the noise amplification of the filtering process.

Suppose that the variance of the recursive filter (see Eq.7-1) is $\sigma_{y_n}^2$, similar to the above procedure, it can be expressed as

$$\sigma_{y_n}^2 = \sigma_x^2 \sum_{k=-\infty}^{k=\infty} c_k^2 + \sum_{m=1}^{M} d_{n-k}^2 \sigma_{y_{n-m}}^2 \qquad 7\text{- }19$$

Further assume that $\sigma_{y_{n-1}}^2 = \sigma_{y_{n-2}}^2 \cdots = \sigma_{y_{n-M}}^2 = \sigma_{y_n}^2$, then the variance of the filter can be estimated by

$$\sigma_{y_n}^2 \approx \frac{\sum_{k=-\infty}^{k=\infty} c_k^2}{1 - \sum_{m=1}^{M} d_{n-m}^2} \cdot \sigma_x^2 \qquad 7\text{- }20$$

This suggests that one can either roughly estimate the variance of the recursive filter using Eq.7-20; or "precisely" calculate the filter variance as follows: first compute the initial variance of the recursive filter using Eq.7-20, and then estimate the variance of the filtered signals using Eq.7-19.

7.1.2 Frequency Response of Ideal Differentiator

If *x(t)* is the input signal applied to a differentiator and the output of the signals is *y(t),* then the first-order time derivative of *x(t)* is

$$\frac{dx(t)}{dt} = y(t) \qquad 7\text{- }21$$

This is a differential equation. Applying the Laplace transform to the above equation leads to

$$sX(s) = Y(s) \qquad 7\text{-}22$$

Since the frequency response function $H(\omega)$ is defined as the rate of the Fourier transform of $y(t)$ over the Fourier transform of $x(t)$, i.e. $Y(\omega)$ over $X(\omega)$, where ω is the simplified notation of $e^{j\omega}$, the transform function of the differentiator can be easily obtained from Eq.7-22 by replacing s with $j\omega$ [Stearns, 2003, p.127]

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} = j\omega \qquad 7\text{-}23$$

Thus in digital filtering theory, the ideal differentiator is defined as having the frequency response of

$$H(\omega) = j\omega \qquad 7\text{-}24$$

Analogous to the first-derivative, the ideal N^{th} order derivatives can be deduced and defined as having frequency response of

$$H(\omega) = (j\omega)^N \qquad 7\text{-}25$$

With this, one can obtain the ideal second-order differentiator as having the frequency response of [Ellum and Sheimy, 2002]

$$H(\omega) = -\omega^2 \qquad 7\text{-}26$$

7.1.3 Criteria for Differentiator Design

The process of selecting a filter's length and coefficients is called “filter design”. The objective is to define these parameters such that certain desired stop-band and pass-band parameters will be obtained by the filter.

Differentiator design has been the subject of extensive investigation in digital signal processing. A major issue is that a differentiator amplifies noise at high frequencies. This grows with the order of derivatives to be estimated, and with the required

bandwidth of the filter [Carlsson et al., 1991]. As GPS signals generally have of low-frequency characteristics (see GPS signal spectral figures below in Fig.7-1), it seems that a low-pass filter would be suitable for the design of a differentiator.

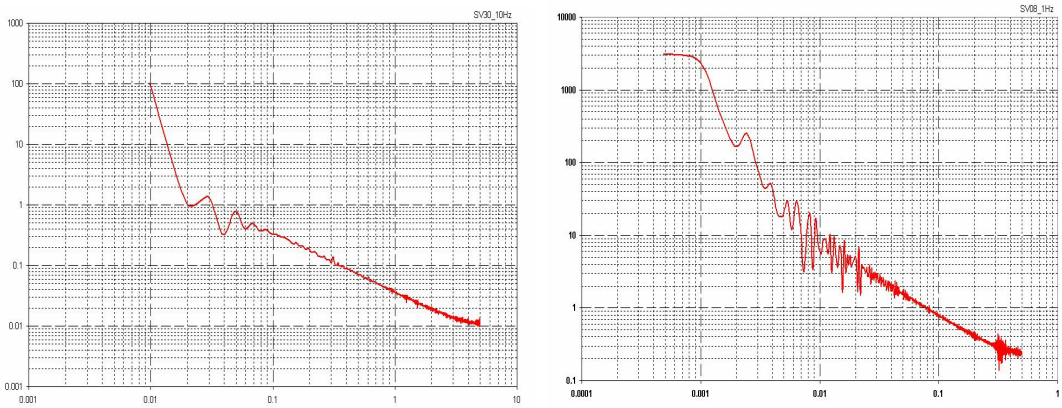

Figure 7- 1: Power spectral densities for the 10Hz (Left) and 1Hz (Right) carrier phase signals in static mode

However, one should be aware that changes of receiver dynamics are normally of high frequency, and therefore it is required to deal with this complexity with a broad/full-band differentiator. Another difficulty arises from the signal correlation. GPS carrier signals can be regarded as Gaussian white only when the sampling rate is lower than 1Hz. When the sampling rate increases, time correlations need to be considered [Bona, 2000; Borre and Tiberius, 2000]. Thirdly, the differentiation may be affected by lack of information on future signals in the case of real-time applications. Finally, there may be aliasing problems due to limited sampling.

Therefore the problem is to obtain the time derivatives from GPS observations when both the signals and the noise are random in character. In the case of corrupting noise being wideband white and the signal being a Gauss-Markov process (suitable for GPS applications), it is apparent that no differentiator is going to perfectly yield the desired time derivative whilst suppressing the noise [Brown and Hwang, 1992, p.172], even though the frequency response of the filter is known exactly. This is a typical Wiener filter problem (ibid), i.e. what should the filter's frequency response be in order to

give the best possible separation of signal from noise? The solution is a compromise between good differentiation and low noise sensitivity to achieve a small total error.

The Kalman filter is a space-state solution of the Wiener filter problem (ibid), by formulating the minimum mean-squares-error estimation criterion into a two-step recursive procedure. By assuming that both the process driving noise and the measurement noise are Gaussian white and there is no correlation between each other, it first predicts the signal state by using the system dynamic equation, and then updates the prediction with the measurements to get the estimates.

A successful Kalman filter requires the appropriate modelling of system dynamics and the associated stochastic random process. Blewitt [1998] elaborated the interrelationship between the functional model and the stochastic model in the sense of equivalence. That is, their functionalities are equivalent so as to achieve the modelling accuracy. Errors not included in the functional model need to be described stochastically, and vice versa. This is applicable to Kalman filtering as well. The less than satisfactory performance of the Kalman filter in the case of Hebert et al. [1997] is not due to the Kalman filter approach itself, but due to the improper modelling of the system state when it is highly dynamic.

When the sampling rate increases, the state equation tends to be adequate to describe the system, even under dynamic conditions. The theoretical difficulties with Kalman filtering, however, are mainly in the determination of the random process of the system driving noise, and dealing with the correlations in measurements and the cross-correlations between the signals and noise. Another associated practical problem is the heavy computational load for a real-time data processor. Finally the outcome of a Kalman filter is a smooth, band-limited solution [Bruton et al., 1999]. Therefore, it is reasonable to find solutions in the frequency domain, rather than in the state space using Kalman filtering.

The digital differentiator design oriented in the frequency domain should still consider the variance of the output. Thus the criteria of our differentiator design can be summarised as follows:

- The magnitude of frequency response is accurate in lower frequencies and is as close to $H(\omega)=j\omega$ as possible in a broad/full-band sense depending on the system dynamics;
- The phase response is linear or approximately linear;
- The group delay is acceptably small;
- The sum of the squares of filter coefficients can be minimised;
- Easy to be implemented in real time, i.e. to be causal and in low-order form since there might be cycle slips and loss-of-lock of signals.

7.2 Taylor Series Approximations

The Taylor series approximations have been widely used to derive differentiators. The differentiators used by Cannon et al. [1998], Hebert [1997] and Hebert et al. [1997], and Kennedy [2002; 2003] are low-order Taylor series. All of them are in the form of central difference approximations as

$$y_n = \sum_{k=-N}^{N} c_k \cdot x_k \qquad 7\text{-}27$$

where N is the order of the filters. Eq.7-27 represents a FIR filter of type III [Chen, 2001, p.299] which is characterised as having zero amplitude response in both ω=0 and ω =1 (normalized frequency), as can be seen in Fig.7-2.

Figure 7- 2: Normalized magnitude response of low order central difference Taylor series approximation

From Fig.7-2 it is apparent that the higher the order is, the closer that a Taylor series approximation is to the ideal differentiator. This indicates that broad-band differentiators can be designed based on the Taylor series, and this can be observed in the work of Khan and Ohba [1999], who gave the explicit coefficients c_i as

$$c_0 = 0$$

$$c_k = c_{-k} = \frac{(-1)^{k+1} N!^2}{k(N-k)!(N+k)!} \qquad 7\text{- }28$$

A FIR filter of type III has an odd length ($2N+1$) and an anti-symmetric impulse response property. Since there is a restriction that the amplitude response must go to zero at the Nyquist frequency, it is impossible to define a full-band differentiator using a finite number of coefficients. This can be seen in Fig.7-3, where a transition of frequency range from 0.85 to 1.0 is associated with the differentiator of the 150th-order Taylor series approximation.

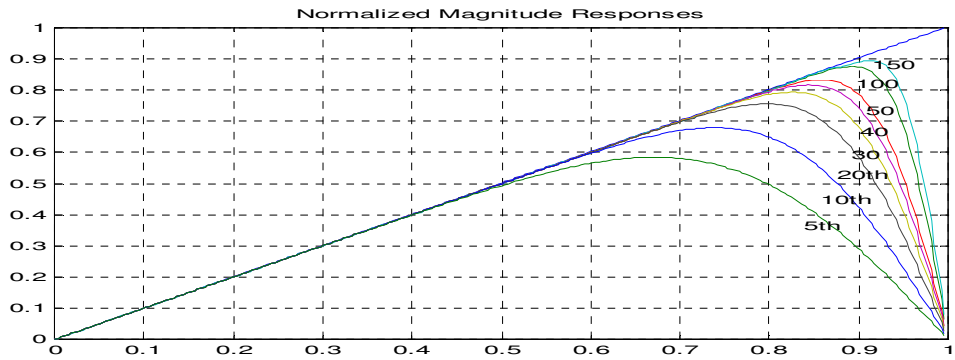

Figure 7- 3: Normalized magnitude responses for arbitrary order Taylor series approximation

Considering that FIR filters using central difference of Taylor series approximations are non-causal and of type III, they are therefore not suitable for real-time applications. However, the Taylor series approximation is still useful for post processing and for low dynamic applications [Hebert et al., 1997]. They might be designed as a full-band differentiator if it can be changed from type III to type IV.

A FIR filter is of type IV if it has an even length and an anti-symmetric impulse response. The type IV FIR is favoured over the Type III as a differentiator in terms of the frequency response. This can be noted from the simplest 2-point FIR differentiator of $y_n = x_n - x_{n-1}$, which has a frequency response of

$$H(z) = \frac{Y(z)}{X(z)} = 1 - z^{-1} \Rightarrow H(\omega) = 1 - e^{-j\omega} = j \cdot 2\sin\frac{\omega}{2} \cdot e^{-j\frac{\omega}{2}} \qquad 7\text{- }29$$

The corresponding amplitude response is shown below Fig.7-4.

It can be seen that even though the differentiator is the simplest in form, it is close to the ideal at low frequencies (<0.2), and has a better amplitude response in the rest of the frequency band than its type IV counterpart of the first-order. It has a linear frequency response and therefore has a constant delay of half the sampling period. The type IV FIR differentiators are superior to the type III in terms of the frequency response since they do not have the undesirable characteristic of being zero at the Nyquist frequency.

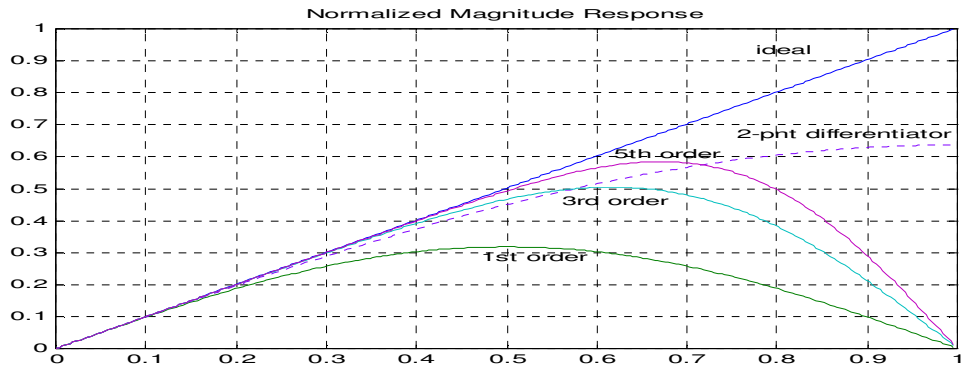

Figure 7- 4: The frequency response of the simplest differentiator versus type III low-order Taylor Series FIR filters

Since the frequency response of type IV FIR filters outperforms the type III at the higher frequency band, this suggests that full-band differentiators can be designed from Taylor series approximations. The explicit formulae for the determination of the full-band filter coefficients were given by Khan and Ohba [1999] and were mathematically proven by Khan et al. [2000]. The differentiators are in the form of

$$y_n = \frac{1}{T}\sum_{k=-N+1}^{N} c_{(2k-1)/2} \cdot x_{(2k-1)/2+n} \qquad 7\text{- }30$$

where

$$c_{(2k-1)/2} = -c_{-(2k-1)/2} = \frac{(-1)^{k+1}(2N-1)!!^2}{2^{2N-2}(2N)!(N+k-1)!(N-k)!(2k-1)^2} \qquad k = 1,2,3\cdots N \qquad 7\text{- }31$$

where the double factorial is defined as $x!! = x(x-2)(x-4)\cdots 1$. The iterative algorithm to improve the computation efficiency was provided and a modification was also proposed to narrow the disclosure at the Nyquist edge (ibid). In summary, the Taylor series approximation can be employed in post processing applications.

7.3 Curve Fitting with Window

Differentiators using the curve fitting with window are popular for velocity and acceleration determination using GPS. Jekeli and Garcia [1997] applied a fifth-order

B-splines approximation to derive phase accelerations, while Fenton and Townsend [1994] adopted a parabolic approximation to obtain the precise Doppler shift. The cited curve fitting techniques use "sliding" windows, where the data are fitted into the polynomials using the least squares approach. The time derivative of the central point of the window is obtained by differentiating the polynomials with respect to time.

For a discrete signal series of $x(t_i)$, if a "sliding" window is set such that the window is centred around time t_0, and the length of the window is from t_{-k} to t_k, then a polynomial approximation of order N can be used to best fit the signals as [Bruton et al., 1999]

$$x(t_i) = \sum_{m=0}^{N} a_m \cdot i^m \qquad i \in [-k, k] \tag{7- 32}$$

The coefficients of the polynomial can then be determined using the standard least squares estimation of the above equation, and the first derivative with respect to time can be obtained by

$$\frac{dx(t_i)}{dt} = \sum_{n=1}^{N-1} a_n \cdot i^n \tag{7- 33}$$

Bruton et al [1999] gave an in-depth review of the curve fitting differentiators. It is concluded that whether a curve fitting uses the approximation of polynomials, parabolas, or cubic splines, it closes to the ideal differentiator only at low frequencies. Since the resultant differentiator is band-limited and low-pass, it is suitable only for low dynamic or static applications. Furthermore, performing the least squares estimation involves intensive computational load since the obtained derivative is only for the central point at t_0. Moreover, to obtain the current derivative at t_0, one has to use the data at t_k, which is an observable in the future. Thus the windowed curve fitting approach is inappropriate for real-time dynamic applications.

7.4 FIR Filters

A Finite Impulse Response (FIR) filter consists of a series of multiplications followed by a summation. This is generally referred to as convolution. The FIR filter operation can be represented by the following equation [Hamming, 1977]

$$y_n = \sum_{i=-K}^{K} c_i \cdot x_{n-i} \qquad 7\text{- }34$$

This is a familiar form as the differentiator designs based on the Taylor series are in such a form. The FIR filter gets its name since its response to an impulse dies away after a finite number of samples. Notice that this form is non-causal and not realisable, and therefore it is not suitable for real-time applications due to the requirement of signal inputs from the future. In order to design a causal FIR differentiator, a change of the form is required. This leads to

$$y_n = \sum_{i=0}^{N} c_i \cdot x_{n-N} \qquad 7\text{- }35$$

7.4.1 Type IV FIR Differentiators

Chen [2001, p.332] showed that for a good differentiator design, the frequency response of an ideal N^{th}-order FIR differentiator should be

$$H_d(\omega) = j\omega \cdot e^{-j0.5N\omega} \qquad 7\text{- }36$$

which has a linear phase of $0.5N\omega$ and thus has a constant group delay of $0.5N$. Since FIR differentiators of type IV have much better frequency response, here a general description of the procedures to determine the filter coefficients is provided following Chen [2001].

The FIR filter of Type IV has an anti-symmetric impulse response with an even length. Firstly, define $M=N/2$ and suppose that the sampling time interval is $T=1$, then

the impulse response can be calculated by the inverse Fourier transform [Antoniou, 1979]

$$c_d[n] = \frac{1}{2\pi}\int_{\omega=-\pi}^{\omega=\pi} j\omega \cdot e^{-j(n-M)\omega} \cdot d\omega = \frac{\cos[(n-M)\pi]}{(n-M)} - \frac{\sin[(n-M)\pi]}{\pi(n-M)^2}$$
$$= -\frac{\sin[(n-M)\pi]}{\pi(n-M)^2} \qquad 7\text{- }37$$

where the infinite length of the Fourier terms has been truncated. The truncation may cause discontinuity at the edges of the window and lead to residual oscillations known as the Gibbs oscillations (ripples in the amplitude response against frequency).

The Fourier transform works on the assumption that the data are periodic, however this is not the case for finite discrete time series. If not quite an integral number of cycles fit into the total duration of the measurements, then the end of one signal segment does not connect smoothly with the beginning of the next, and therefore there are small glitches at regular intervals. Different window methods can be used to smooth the glitches, truncate the filter coefficients, and sharpen up the filter's frequency response. By adopting an appropriate window function, for example the Hamming window function, the FIR coefficients can be determined as [Antoniou, 1979]

$$c[n] = c_d[n] \cdot w_H[n] \qquad 7\text{- }38$$

where w_H is the Hamming window function given as

$$w_H = \begin{cases} 0.54 + 0.46\cos\dfrac{2n\pi}{N-1} & for|n| \le \dfrac{N-1}{2} \\ 0 & otherwise \end{cases} \qquad 7\text{- }39$$

Such a type IV filter of the 7th- order, for example, minimises

$$E = \int_{-\pi}^{\pi} \left| H(e^{j\omega}) - j\omega e^{jN\omega/2} \right|^2 d\omega \qquad 7\text{- }40$$

and has a transfer function

$$H(z) = -0.0260 + 0.0509z^{-1} - 0.1415z^{-2} + 1.2732z^{-3} - 1.2732z^{-4} + 0.1415z^{-5} - 0.0509z^{-6} + 0.0260z^{-7} \qquad 7\text{- }41$$

This is an optimal differentiator in the sense of least squares, with an excellent frequency response at high frequency band. The noise amplification can be calculated from Eq.7-18 to be $\sigma^2(y_n)$=3.2887, which is acceptable.

It is expected that a type IV FIR obtained from the Remez exchange algorithm is able to deliver a better performance. This is because the Remez exchange algorithm is a minimax optimal, i.e.

$$\text{minimise}\{\max[H_{ideal}(\omega) - H_{designed}(\omega)]\} \qquad \forall : 0 < \omega < 1 \qquad 7\text{- }42$$

and is more difficult to mathematically compute, but guarantees that the worst case error has been reduced to a quantifiable value. To verify this, two graphs of the frequency response have been presented for the 7th and 25th order filters using the Remez algorithm in Fig.7-5

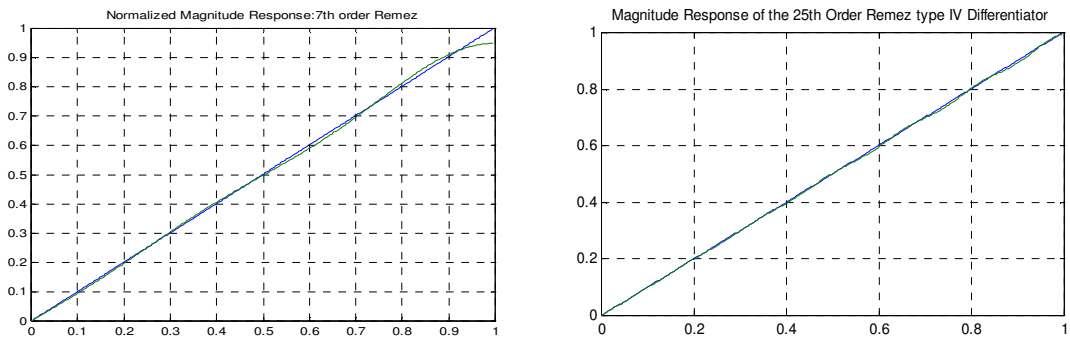

Figure 7- 5: Frequency responses of the type IV FIR filters by Remez exchange algorithm

The FIR filter design using the Remez algorithm is referred to as the "equal ripple design" since this method can suppress the ripples from the Gibbs phenomenon to a certain level, and convert them into equal ripples in both the pass-band and stop-band.

It seems that type IV FIR filters using the Remez exchange algorithm will give us a closed solution, however it has been found that wide-band type IV differentiators are

associated with heavy noise amplification and a large bias. This has been reconfirmed with the performance tests on the type IV FIR using the Taylor series approximations (see Eq.7-31). It is found that only the simplest two-point differentiator of type IV can produce acceptable first-order derivatives. It may require more investigations on type IV FIR filters for differentiator design.

7.4.2 Type III FIR Differentiators

Since FIR filters of type IV perform poorly in the state space, it is worthwhile to investigate type III filters. A FIR filter of type III has an odd length and an anti-symmetric impulse response. In this case, the coefficients of the differentiator are [Chen, 2001]

$$c_{desired}(n) = \begin{cases} \dfrac{\cos[(n-M)\pi]}{(n-M)} & for \quad n \neq M \\ 0 & for \quad n = M \end{cases} \qquad 7\text{-}43$$

To eliminate the Gibbs phenomenon due to the finite truncation, a window function is required (as mentioned earlier) .

Among many windows available, the Kaiser window [Farlex, 2004]

$$w_k(n) = \begin{cases} \dfrac{I_0(\pi\alpha\sqrt{1-(2k/N-1)^2})}{I_0(\pi\alpha)} & for \quad 0 < k < N \\ 0 & otherwise \end{cases} \qquad 7\text{-}44$$

is the most popular since it can be evaluated to any desired degree of accuracy using the rapidly converging series of the zero-order Bessel function of the first kind

$$I_0(x) = 1 + \sum_{k=1}^{\infty}\left[\frac{1}{k!}(\frac{x}{2})^k\right]^2 \qquad 7\text{-}45$$

where the first 15 to 20 terms are sufficient for the filter to converge. Fig.7-6 depicts FIR differentiators of length 31 with and without the Kaiser window respectively.

The ripple of the stop-band is controlled by the adjustable variable α. In our case, it is found that α=7.8

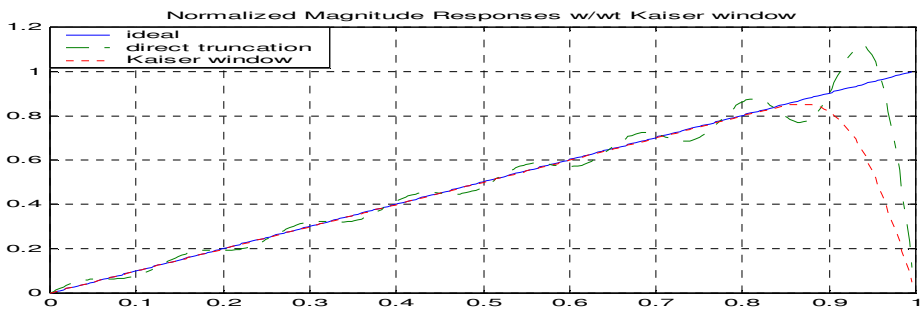

Figure 7- 6: FIR differentiator (filter length=31) with/without Kaiser window

is optimal for the Doppler/Doppler rate derivation when the FIR filter length is 31 to meet the optimality criteria given by Kumar and Roy [1988b], Selesnick [2002]

$$\begin{aligned} &\frac{d}{d\omega}H(\omega)=1 \qquad for \qquad \omega=0,\pi \\ &\frac{d}{d\omega}H(\omega)=0 \qquad for \qquad \omega=0,\omega=\omega_{cut} \\ &\frac{d^{v}}{d\omega^{v}}H(\omega)=0 \qquad for \qquad \omega=0, and \quad v=2,3,\cdots \end{aligned} \tag{7- 46}$$

With the above procedures one can design FIR differentiators with different cut-off frequencies, as can be seen from Fig.7-7. A closer look at Fig.7-7 indicates that the frequency responses of the filters are not linear at both the low and high frequencies. This can be improved, however, see Calsson [1991], Kumar and Roy [1988b].

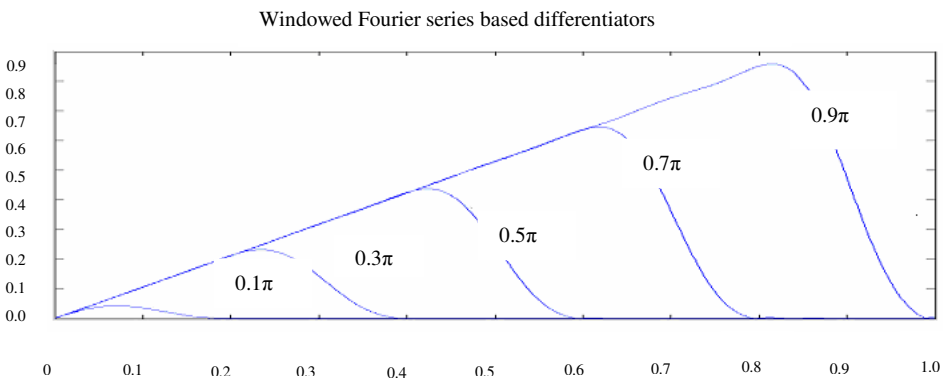

Figure 7- 7: Magnitude responses of FIR differentiators based on the window technique

Theoretically, FIR differentiators of type III can be designed to meet the requirement at nearly all frequencies, as long as N in Eq.7-35 can be increased. Although such filters are causal and linear in phase, the actual derivatives that yielded are with respect to time $t-(N/2)T$. This means that the more taps in a FIR filter, the longer the group delay will be. This property of the FIR filter is detrimental to real-time requirements. However, it can be eased if the sampling period T is small. The difficulty is that increasing the sampling frequency will result in more noisy derivatives. Therefore, a trade-off and compromise must be made when introducing FIR filters in such applications.

In a series of publications, Kumar and Roy [1988a; 1988b; 1989a; 1989b] presented optimal and maximally linear FIR differentiators for low-frequency range, mid-frequency range, and around spot frequencies. They presented explicit formulae and efficient recursive algorithms to calculate the impulse response of the filters. Their contributions have been highly acknowledged, for example, as state of the art differentiators by Al-Alaoui [1993]. In the case where signals have low frequency components that are contaminated by wide-band noise, FIR differentiators of optimum white-noise attenuation are desirable. Kavanagh [2001] investigated the impact of quantisation noise of signals from systems with low-frequency rates of change, and reported that the differentiator proposed by Vainio et al. [1997]

$$h_n = \frac{6(N-1-2n)}{N(N^2-1)} \qquad 0 \le n \le N-1 \tag{7-47}$$

has an optimum white-noise attenuation and a constant group delay. A better differentiator is proposed when the rate experiences a slow change [Kavanagh, 2001]

$$h_n = \begin{cases} \dfrac{1}{N-1} & n = 0 \\ 0 & 0 \le n \le N-2 \\ -\dfrac{1}{N-1} & n = N-1 \end{cases} \qquad 7\text{- }48$$

which has the characteristic of minimising the worst-case error. Clearly, it is the simplest two-point differentiator when N=2 and the FIR filter of the three-point first-order 'central' difference of a Taylor series approximation when N is 3.

7.5 IIR Filters

There is another category of filters referred to as the Infinite Impulse Response filters (IIR). A causal IIR filter is represented by

$$y_n = \sum_{k=0}^{N-1} c_k x_{n-k} + \sum_{m=1}^{M-1} d_m y_{n-m} \qquad 7\text{- }49$$

where the output signal at a given instant is obtained as the weighted sum of the current and past inputs x_n, and the past outputs of y_n. As suggested by its name, an impulse signal input has a response that lasts forever since the output will be recursively used. Notice that Eq.7-49 also represents the Auto Regressive Moving Average (ARMA) model. Since the previous outputs are recursively used, IIR filters can be implemented with a lower order, which gives better performance when compared with FIR filters. Thus they are attractive for real-time applications.

The transfer function of the above IIR filter is

$$H(z) = \frac{\sum_{k=0}^{N} c_k \cdot z^{-k}}{1 - \sum_{n=1}^{M} d_k \cdot z^{-n}} \qquad 7\text{- }50$$

and the magnitude of the frequency response is

$$\left|H(\omega)\right| = \left|\frac{\sum_{k=0}^{N} c_k \cdot e^{-j\omega k}}{1-\sum_{n=1}^{M} d_k \cdot z^{-j\omega n}}\right| = \frac{\left|\Theta_{Num}(\omega)\right|}{\left|\Theta_{Den}(\omega)\right|} \qquad 7\text{- }51$$

The phase response of the IIR filter is the phase of the numerator minus the phase of the denominator [Lyons, 2004, p.237], i.e.

$$\Theta(\omega) = \Theta_{Num}(\omega) - \Theta_{Den}(\omega) \qquad 7\text{- }52$$

An IIR filter is unstable if its response to a transient input increases without bound. Poles and zeros are used to analyse the stability of the IIR filter. The poles are the roots of the denominator and the zeros are the roots of the numerator in the transfer function. The IIR filter described by H(z) is stable if and only if all poles of H(z) are inside the unit circle on the z-plane [Stearns, 2003, p.83].

IIR filters cannot be designed by calculating the impulse response from the known frequency response as was done for FIR designs. Many IIR filters can be derived from the analogue filter designs and then transformed into the sampled z-plane. Another popular method is the bilinear transform. This method relies on the existence of a known *s*-domain transfer function (Laplace transform) of the filter to be designed. The *s*-domain filter coefficients are then transformed into z-domain coefficients.

The IIR differentiator design has been of considerable interest for long [Rabiner and Steiglitz, 1970]. Among various recursive differentiator designs, the Al-Alaoui's IIR family [1992; 1993; 1994; 1995] has been highly acknowledged and widely used [Chen and Lee, 1995]. The novel approach of designing digital differentiators by Al-Alaoui is an extension of the method used for designing analogue differentiators by using integrators. That is, in analogue signal processing differentiators are often obtained by inverting the transfer functions of analogue integrators.

The general procedures to derive the Al-Alaoui family of IIR filters are as follows

- Design an integrator that has the same range and accuracy as the desired differentiator;
- Invert the obtained transfer function of the integrator;
- Reflect the poles that lie outside the unit circle to inside, in order to stabilise the resultant transfer function;
- Compensate the magnitude using the reciprocals of the poles that lie outside the unit circle.

7.5.1 Differentiator from Simpson's Integrator

The transfer function of the Simpson's integrator is [Al-Alaoui, 1994]

$$H_{I_s}(z) = \frac{T(z^2 + 4z + 1)}{3(z^2 - 1)} \qquad 7\text{- }53$$

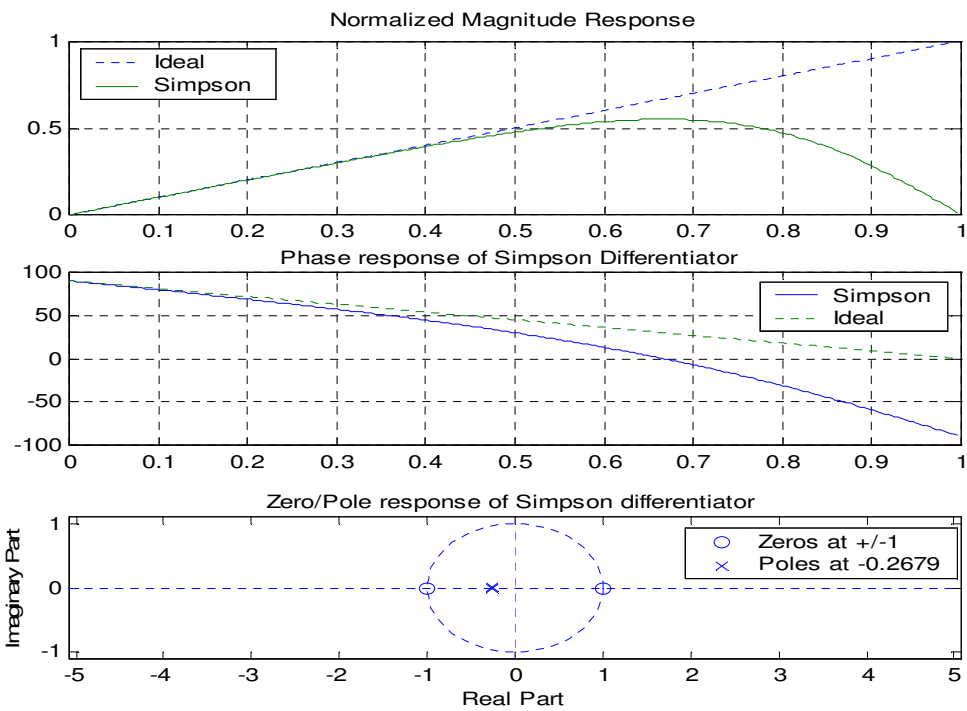

Figure 7-8: Frequency response, zeros, and poles of the differentiator from the Simpson's integrator

Applying Al-Alaoui's procedures, the transfer function of the differentiator is given by

$$H_{D_s}(z) = \frac{3(z^2 - 1)}{3.7321 \cdot T(z^2 + 0.5358z + 0.07181)} \qquad 7\text{-}54$$

As can be seen in Fig.7-8, this second-order IIR differentiator approximates the magnitude of the ideal differentiator up to 0.4 of the full frequency band and has a linear phase close to the ideal for low frequencies. With the pole well inside the circle, the proposed IIR filter is stable.

7.5.2 Differentiator Families

7.5.2.1 Second-Order Family

Having observed that the ideal integrator response lies between the response of the traditional trapezoidal and Simpson integrators, Al-Alaoui proposed that weighting interpolations could be used to reach the ideal from the above integrators. This can be expressed by

$$\begin{aligned} H(z) &= \alpha H_{I_S}(z) + (1-\alpha) H_{I_T}(z) \qquad 0 \le \alpha \le 1 \\ &= \alpha \cdot \frac{T(z^2 + 4z + 1)}{3(z^2 - 1)} + (1-\alpha)\frac{T(z+1)}{2(z-1)} \\ &= \frac{T(3-\alpha)(z + r_1)(z + r_2)}{6(z^2 - 1)} \end{aligned} \qquad 7\text{-}55$$

where

$$r_1 = \frac{3 + \alpha + 2 \cdot \sqrt{3\alpha}}{3 - \alpha} \qquad \textit{and} \quad r_2 = 1/r_1 = \frac{3 + \alpha - 2 \cdot \sqrt{3\alpha}}{3 - \alpha} \qquad 7\text{-}56$$

This is a class of integrators characterised by the zeros being reciprocal pairs around the unit circle in the z-plane. Varying α may result in different integrators such as the Tick's when α=0.8495 for example [Al-Alaoui, 1993]. Applying the same procedures developed by Al-Alaoui, the corresponding class of differentiators is derived with the transfer function of

$$H_D(z) = \frac{6(z^2 - 1)}{T \cdot r_1(3-\alpha)(z + r_2)^2} \qquad 0 \le \alpha \le 1 \qquad \text{7- 57}$$

where r_2 is set to locate the poles inside the unit circle from the reciprocal pair. This results in the designed IIR filters being stable.

Al-Alaoui [1993] reported that the derived low-pass differentiator class has smaller delay and superior performance than the differentiators of Kumar and Roy [1988b]. In the range of low frequencies of interest to this research, the differentiators have a near linear property and thus are attractive for many real-time applications.

7.5.2.1 First-Order Family

A first-order IIR differentiator was developed by Al-Alaoui [1995] with an effective range of 0.8 of the Nyquist frequency based on a non-minimum phase digital integrator. The integrator is a synthesis of the rectangular integrator and the trapezoidal integrator. By assigning weighting factors of ¾ and ¼ to the transfer functions of the integrators, the ideal integrator that has the following transfer function is approximated

$$\begin{aligned} H_I(z) &= \frac{3}{4} H_R(z) + \frac{1}{4} H_T(z) \\ &= \frac{3}{4} \cdot \frac{T}{z-1} + \frac{1}{4} \frac{T(z+1)}{2(z-1)} = \frac{T}{8} \cdot \frac{z+7}{z-1} \end{aligned} \qquad \text{7- 58}$$

Reflecting the zero z=-7 with its reciprocal z=-1/7 and compensating the magnitude by multiplying r=7 results in a minimum phase digital integrator with transfer function

$$H_I(z) = \frac{7 \cdot T}{8} \cdot \frac{z + 1/7}{z - 1} \qquad 7\text{- }59$$

Inverting the above transfer function yields Al-Alaoui's stabilised IIR differentiator of the first-order

$$H_D(z) = \frac{8/7}{T} \cdot \frac{z - 1}{z + 1/7} \qquad 7\text{- }60$$

which has the characteristics as depicted in Fig.7-9.

The new differentiator is able to approximate the ideal differentiator up to 0.78 of the full frequency band, and has an outstanding "linear phase" response. Al-Alaoui reported that within the effective frequency range, it has a less than 2.0% magnitude error. Since the pole is -1/7, it is rather stable; and since it is of first-order, the delay of the filter is just half of the sampling interval and thus it meets every requirement for use in real-time.

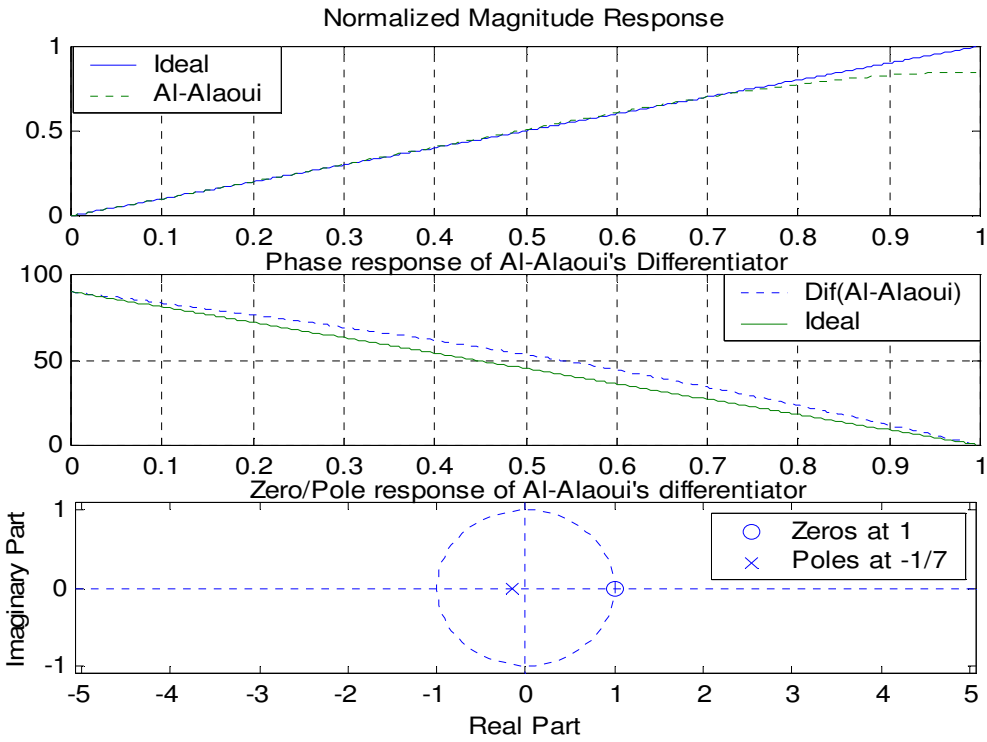

Figure 7- 9: Characteristics of Al-Alaoui's first-order IIR differentiator

Al-Alaoui contributed this differentiator as an individual; however, a family of such first-order differentiators can be derived following the same methodology. That is, while Al-Alaoui designated the weighting factors of ¾ and ¼ empirically, one may obtain the optimal weights experimentally. To achieve this, a variable α is introduced to adjust the weighting factor in the following way

$$H_I(z) = \alpha H_R(z) + (1-\alpha) H_T(z) \qquad 0 < \alpha \approx 0.75 < 1 \qquad \text{7- 61}$$

where α servers as a "tuner" to adjust the integrator so that it better approximates the ideal. α=¾ can be used as a reference to refine the integrator in the desired range of frequencies. This leads to the following transfer function

$$H_I(z) = \alpha H_R(z) + (1-\alpha) H_T(z)$$

$$= \alpha \cdot \frac{T}{z-1} + (1-\alpha)\frac{T(z+1)}{2(z-1)} = \frac{T(1-\alpha)\left[z + \frac{1+\alpha}{1-\alpha}\right]}{2(z-1)} \qquad 7\text{-}62$$

where $0 < \alpha < 1$. Obviously, it has a zero outside the unit circle. Applying Al-Alaoui's procedure to reflect the zero with its reciprocal and compensating for the magnitude, a variable integrator is obtained as

$$H_I(z) = \frac{T(1+\alpha)\left[z + \frac{1-\alpha}{1+\alpha}\right]}{2(z-1)} \qquad 7\text{-}63$$

which represents a set of a minimum phase digital integrators. Inverting the transfer function gives a new set of differentiators with transfer functions as

$$H_D(z) = \frac{2(z-1)}{T(1+\alpha)\left[z + \frac{1-\alpha}{1+\alpha}\right]} \qquad 7\text{-}64$$

since the poles are at $-\frac{1-\alpha}{1+\alpha}$, which is inside the unit circle and therefore the resultant differentiators are stable. Setting α with 0.75 gives the transfer function proposed by Al-Alaoui, and slightly changing α around 0.75 results in differentiators that perform well in the target bandwidth. The noise amplification of this kind of differentiators can be evaluated using Eq.7-20, which is a little bit nosier than the simplest two-point differentiator is.

7.6 Summary

General theories on digital filter design have been introduced, with the aim of finding an appropriate differentiator that can be used to derive Doppler shifts or Doppler rates from GPS observables in real-time and dynamic applications. The differentiators from

both the curve fitting and Kalman filtering approaches require intensive computation and are low-pass, and thus not suitable for this purpose.

Type III FIR differentiators have the inherent characteristic of the frequency response going to zero at the Nyquist frequency. To extend the performance of type III FIR filters to high frequency bands one has to increase the filter taps. This results in a longer time delay, which is detrimental for real-time applications that require the group delay of the differentiator to be as short as possible.

Type III FIR filters can be used to derive Doppler/Doppler rate 'observables' in the post-processing mode, and higher order central difference of Taylor series approximations might outperform those based on windowed Fourier series, since there are no truncations and no associated Gibbs phenomenon.

Type IV FIR differentiators using Fourier series have outstanding frequency response, however they are usually noisy and biased. In this research, it is observed that only the Kavanagh [2001] differentiators of type IV (including the simplest two-point differentiator) deliver good first-order time derivatives. However, they approximate the ideal differentiator only for the range lower than 0.2 of the Nyquist frequency.

IIR filters are more suitable for real-time operation. Since the outputs of the filter are recursively used, they have much lower orders than the FIR filters. The first-order IIR differentiator from Al-Alaoui is ideal in terms of frequency response, phase linearity and half sample group delay. The proposed class of first-order IIR differentiators allows us to define the optimal parameters in the desired frequency range.

It is suggested that the Kavanagh [2001] differentiators can be used for static or in constant velocity modes, and the proposed IIR differentiators of the first-order be adaptively used when systems experience higher dynamics.

Chapter Eight

MISCELLANEA OF VELOCITY AND ACCELERATION DETERMINATION

The physical and mathematical principles of ground velocity and acceleration determination have been elaborated in the previous chapters. With a set of differentiators designed in Chapter Seven, the ground velocity and acceleration can be determined from either the carrier phase measurements or the Doppler shift measurements in real time.

However, there are many operational issues in utilising such a set of differentiators with a GPS receiver operating in real time. Amongst them, the effects of receiver clock reset, cycle slips and loss-of-lock of signals require further discussions.

The output of a differentiator inevitably becomes noisy when increasing the sampling rate, which would result in deterioration of the ground velocity and acceleration estimates. In particular, the acceleration determined based on the Doppler rate method would be severely affected due to the amplification of noise if the incoming signals are carrier phase measurements.

This chapter discusses the above issues. §8.1 describes the effects caused by the receiver clock reset, loss-of-lock of signals and cycle slips, and investigates the corresponding treatments in differentiation processes. §8.2 discusses the second-order differentiator design and proposes the use of a cascade scheme. §8.3 presents alternative methods to derive ground accelerations in both the measurement and

velocity domains, which are expected to have better performance in high sampling rate scenarios.

8.1 Issues in DSP of GPS Receiver

The differentiator design discussed in Chapter Seven is directly related to Discrete Signal Processing (DSP), where both FIR and IIR filters generally work on the basis that incoming digitised discrete signals have equal time intervals. It is assumed that there is no abnormal signal in the incoming time series. These conditions, however, are too ideal for GPS observations.

Time-tags of GPS measurements are dependent on the Numerically Controlled Oscillator (NCO) and subsequently corrected for the estimated receiver clock biases, which are desired from the navigation solution. Due to the instability of the oscillator, and the bias and drift of the receiver clock, the GPS measurements are not equally time-spaced in a strict sense.

Quite often loss-of-lock of signals occurs in a receiver when tracking a GPS satellite. As a result, there will be blank records in the time series of measurements. Neglecting the zeros would lead to an incorrect output from the differentiator, producing erroneous Doppler shifts and Doppler rates, and subsequentially poor velocity and acceleration estimates.

In the case that the carrier phase measurements are used as the input of a differentiator, it is also necessary to take into account the carrier phase cycle slips.

8.1.1 Receiver Clock Reset and Edge Effect

It has been stated in §3.2.3.1 that there are two different schemes for manufacturers to handle GPS receiver clock drifts: one is receiver clock steering, and the other is receiver clock reset.

Under the scheme of receiver clock steering, receiver clock drifts are adjusted epoch by epoch by the receiver internal software so as to synchronise the receiver clock to the GPS system time. In this circumstance, the sampled measurements (L1, C/A, D1, etc.) can be regarded as being output at equal time intervals, and thus poses no problems.

For those receivers utilising the receiver clock reset, the clock is allowed to drift until a threshold of one millisecond is reached, and then there will be an increment of the clock to correct the time at that epoch [Farrell and Barth, 1999, p.151]. When this occurs, the sampling time intervals will be non-uniform. As a result there are edge effects on the derivatives when the signal at the reset epoch is used as an input. If the differentiator were an IIR filter, theoretically it would have effects on each derivative thereafter. Hence signal re-sampling is necessary.

Figure 8-1 illustrates the concept of receiver clock reset and signal re-sampling. Suppose that the differentiator used for Doppler shift or Doppler rate derivation is an IIR filter of first-order or a 3-point FIR filter of the central difference of a Taylor approximation. In the implementation of the filter, there should be a queue acting as a container to manage the time series of measurements that are the filter input. Note that the container is usually set longer than the length of the actual filter.

In the case of a clock reset as shown in Fig.8-1, the sampling space is one millisecond shorter than the normal time interval. This is rendered in yellow in the first two states. If the time series are the carrier phase measurements and the previously derived Doppler shift is 6,000Hz for example, then neglecting this 1ms time shortage would cause ±6Hz error in the derived Doppler shift at this epoch. The edge effect would affect each derivative of the measurements at the reset epoch and the neighbouring epochs as well unless appropriately treated.

To eliminate the edge effect caused by the receiver clock reset, the measurement of the reset epoch should be re-sampled to the normal time, i.e. from $X_n(t_r)$ to

$X'_n(t_r \pm 1ms)$ where t_r indicates the time of the receiver clock reset. This is shown in the first and second states in Fig.8-1, where an extrapolation of the measurements is required. The slim red bar represents the compensation time interval needed.

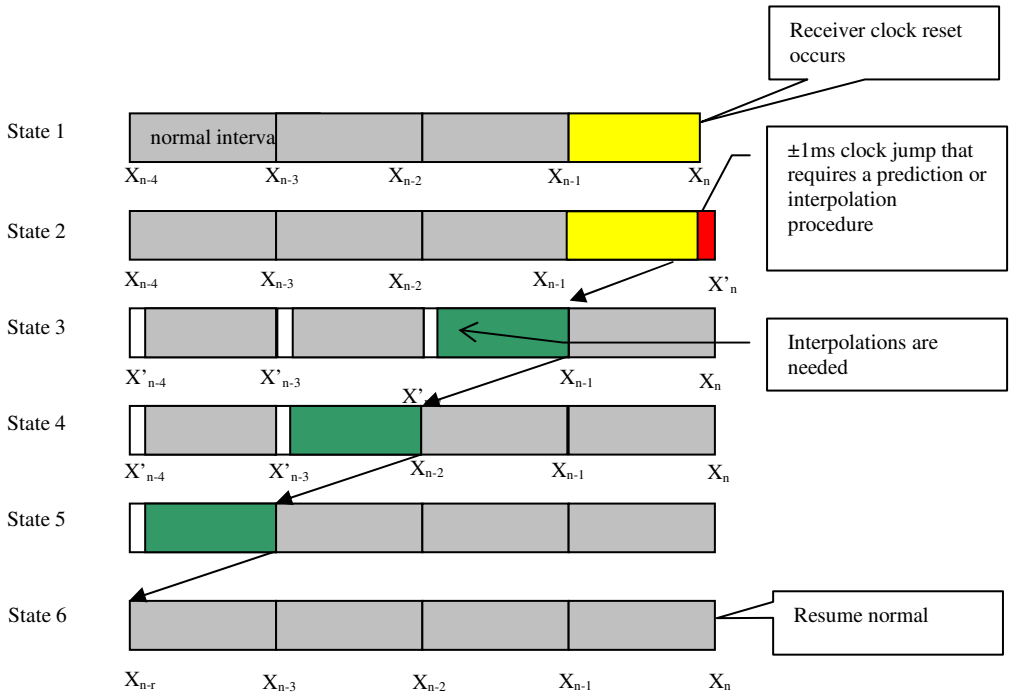

Figure 8- 1: Receiver clock reset and states of signal re-sampling

The prediction can be conducted using a polynomial approximation. In this case, there are five measurements (X_n, X_{n-1}, …X_{n-4}) in the queue, so a polynomial of order 4 can be formed and the expected value can be extrapolated at $t=t_r+1ms$.

The transition of states of the receiver clock reset is also illustrated in Fig.8-1. In State 3, with the push-in of a new measurement, the extrapolated value should be replaced by a recalculation using the polynomial interpolation. This improves the accuracy since extrapolations are inferior to interpolations in nature. It can also be seen that the reset epoch dies out in State 6 when the uniform sampling space is resumed.

The method described above uses raw observations only. An alternative method to deal with the receiver clock reset may use aids from the derived derivatives, where a simple linear prediction scheme might be sufficient. A brief description is provided.

When a new observation is made and the queue is shifted to State 3, from left to right, the data set (X_{n-2} ,X_{n-3}, X_{n-4}) needs to be interpolated as an equally spaced set (X'_{n-2}, X'_{n-3}, X'_{n-4}), where the primes stand for the re-sampled input values. Many interpolation methods can be used for this purpose. However, the Hermite interpolation may be most suitable since there are not only the measurements at each epoch, but also the derived first-derivatives as additional information. Moreover, it has the advantage of working on un-tabulated points.

For n un-tabulated points $y_i = f(t_i)$ ($t_0 < t_1 < \ldots < t < \ldots < t_{n-2} < t_{n-1}$) with known first time derivatives $y_i^{'}$

$$y_i = f(t_i)$$
$$y_i^{'} = f^{'}(t_i) = \frac{df(t)}{dt}\Big|_{t=t_i} \qquad i = 0,1,\cdots,n-1 \qquad \text{8- 1}$$

the Hermite interpolation is [Xu, 1996, p.115]

$$f(t) = \sum_{i=0}^{n-1}[y_i + (t - t_i)(y_i^{'} - 2y_i l_i^{'}(t_i))] \cdot l_i^2(t) \qquad \text{8- 2}$$

where

$$l_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n-1}\left[(t - t_j)/(t_i - t_j)\right]$$
$$l_i^{'}(x) = \sum_{\substack{j=0 \\ j\neq i}}^{n-1}\frac{1}{t_i - t_j} \qquad \text{8- 3}$$

8.1.2 Signal Loss-of-Lock

A GPS receiver may lose track of signals from a particular satellite due to reasons such as a failure in the receiver internal tracking loop, or a signal blockage. When a signal loss-of-lock occurs, there will be no observation at this epoch. Consequently, an abrupt derivative will be generated through the differentiation process unless there is an appropriate treatment of this effect.

When a blank observation is detected, one should determine whether it is due to the satellite falling below the elevation cut-off angle mask. If this is the case, then no extra manipulation is required since the tracking of the satellite is over. If it is determined that the blank observation is a real loss-of-lock, then a prediction of the observation is required.

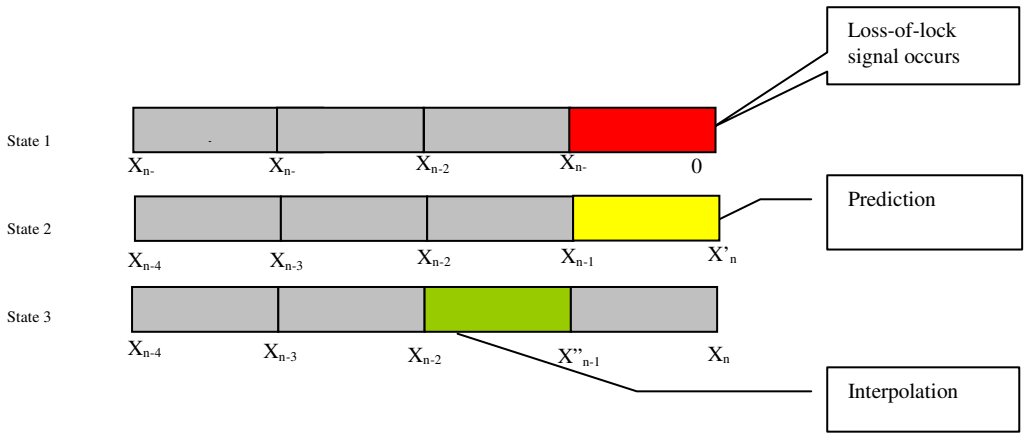

Figure 8- 2: Loss-of-Lock Signal Handling

Figure 8-2 shows the transition of the three states of a signal loss-of-lock. State 1 indicates a blank observation occurring where the measurement is zero. The prediction of the signal is then carried out in State 2, which is rendered in yellow (In fact, State 1 and State 2 are still the same state since no new signal comes into the system, the transition involves replacing the blank observation). When the successive signal is measured at the next epoch, the predicted value should be replaced by an

interpolation in order to improve the accuracy. This is represented through the shift from State 2 to State 3. The same procedures as in §8.1.1 apply in both the prediction and interpolation processes.

If consecutive blank observations occur it is recommended that the differentiator should be re-initialised. In addition, the differentiator needs to be reinitialised if there is a concurrence of a receiver clock reset and a signal loss-of-lock. These treatments of the receiver clock reset and signal loss-of-lock demonstrate why it is desirable that the differentiator design should be of low order. For a real-time system, managing many taps of the filter is a nontrivial task.

8.1.3 Cycle Slips

If carrier phase measurements are used as the input of differentiators, it is important to have a cycle slip detection scheme as the carrier phase measurements occasionally experience cycle slips. When a cycle slip occurs, similar to the receiver clock reset, the time derivatives will be affected. Since the magnitude of cycle slip varies, it is hard to quantify the effects on the derivatives.

Cycle slip detection has been a topic that has attracted much research attention. This is because the accuracy of conventional GPS baseline solutions relies on “cycle slip free” carrier phase measurements and “fixed integer ambiguities”. Many cycle slip detection approaches for static baseline processing and RTK have been developed, which can be grouped into the following schemes:

- Comparing the difference between consecutive carrier phase and code values (range residual);
- Comparing successive ionospheric residuals;
- Comparing the residuals with a curve fitting;
- Comparing with the Doppler shift values.

In the measurement domain, a dual-frequency GPS receiver has more options to detect and even estimate cycle slips than a single-frequency receiver. The changes of L_g, the geometry-free linear combination of L1 and L2 observations, can be used to detect cycle slips very effectively. The only limitation of using L_g is that it does not have the ability to detect concurrent cycle slips with the same cycles on both L1 and L2. Since the possibility of this cycle slip happening is very low, it is recommended to use L_g time series to detect cycle slips for dual-frequency GPS receivers.

Although it is rarely used, the principle of utilising the Doppler shifts to facilitate cycle slip detections can be demonstrated by the fact that any deviation between the predicted carrier phase and the actual carrier phase measurements is due to either the measurement noise or a cycle slip, i.e.

$$\varphi_{residual} = \varphi_n - \left[\varphi_{n-1} + 0.5(D_n + D_{n-1})\Delta t\right] \qquad 8\text{-}4$$

Suppose measurements are recorded at an interval of one minute, the carrier phase residuals calculated by the above equation are suspect since there might be large Doppler shift changes during this long time interval. It is therefore not viable for use in cycle slip detection for static GPS applications with low sampling rates.

However, in the case when the sampling rate is high, say 10Hz or even higher, the Doppler aiding becomes an effective tool to detect and estimate cycle slips since the dramatic change of $\varphi_{residual}$ can be attributed to a cycle slip. This method does not depend on the measurements from another frequency band, and thus is more favourable for single-frequency GPS applications.

As this research targets single-frequency GPS applications, Doppler aiding is adopted as the sole means for detecting and estimating the cycle slip values of carrier phase measurements. By managing the queue of the carrier phase measurements and the derived Doppler shifts, a carrier phase measurement is firstly predicted using Eq.8-4 where the current Doppler shift is estimated using an extrapolation of the relative

velocity between the receiver and the satellite. A cycle slip is flagged if the calculated residual $\varphi_{residual}$ exceeds a preset threshold.

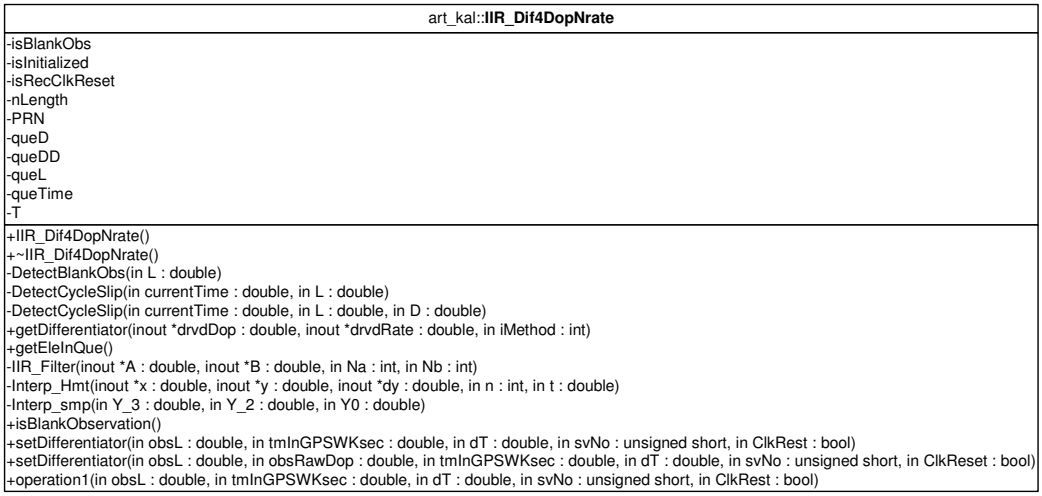

Figure 8- 3: The IIR differentiator class design in UML

Figure 8-3 shows the IIR differentiator class designed in this research, which is expressed in the Unified Modelling Language (UML) [Booch et al., 1999]. This figure is automatically generated from the C++ source code by the Microsoft Visio™ software.

Whenever a new observation arrives, the public member function *IIR_Dif4DopNrate:: setDifferentiator()* is called. It detects the receiver clock reset at the epoch, issuing a Boolean true to *IsRecClkReset* when there is a real receiver clock reset. *IsRecClkReset* is one of the private member variables of the IIR class, taking a Boolean false in most of the time. When it becomes true, the signal re-sampling procedures described in §8.1.1 will be activated. A blank observation will be detected and corrected using the scheme described in §8.1.2. Interpolation and prediction are also carried out when required in the case of a receiver clock reset or a signal loss-of-lock. Meanwhile, a cycle slip is monitored through calling *DetectCycleSlip(…)*, a

private member function of the class. A detected cycle slip will trigger a re-initialisation of the differentiator (meaning that the carrier phase measurements will not be used to derive Doppler shifts until the differentiator is re-initialised). All these are accomplished by the private member functions that have minus signs in front of their names in Fig.8-3.

For brevity, no further coding details of the implementation are presented here since they may vary with different systems and platforms. However, the point is that when designing such systems, one must be aware of the operational problems due to the presence of the receiver clock reset, blank observations and cycle slips. A successful implementation requires these effects to be considered and appropriately treated.

8.2 Differentiator for Second-Order Time Derivatives

A differentiator for second-order time derivatives is able to directly derive the Doppler rate from the carrier phase observables. The simplest way of designing such a second-order differentiator is to use the Taylor series approximations. A FIR filter of any length can be designed in this way. Low-order filter coefficients can be found in many mathematical handbooks [cf. Beyer, 1980].

An alternative method of FIR filter design uses the Fourier transform. The ideal frequency response of the second-order differentiator is (see § 7.1.2)

$$H(\omega) = -\omega^2 \qquad 8\text{-}5$$

With the frequency response, the impulse response of the filter (filter coefficients) can then be determined using the inverse Fourier transform. The procedures are similar to those described in § 7.4.1, and readers are referred to Antoniou [1979] for more details.

Nevertheless, it is found that the derived second-time derivatives are very noisy at 10Hz rate, regardless of the chosen method. The errors in the carrier phase

measurements have been greatly amplified. Although both the Taylor series approximation and the Fourier transform methods have their own successful applications in other areas, they are not suitable for Doppler rate derivation for acceleration determination using GPS at high-sampling rates.

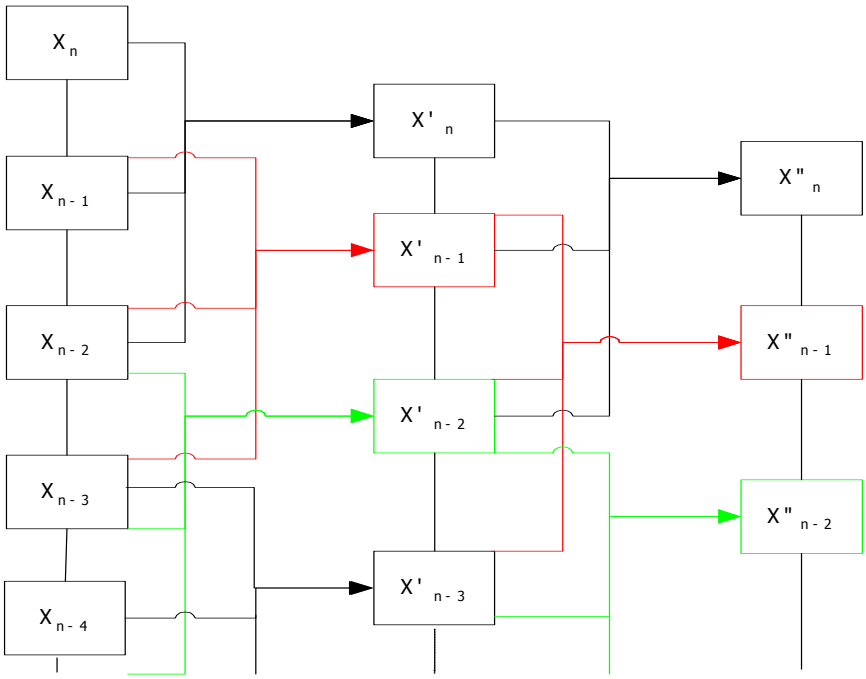

Figure 8- 4: Cascade differentiators for the second-order time derivatives

Cascade differentiators are therefore designed whereby there are two columns of the first-derivative differentiators running in cascade. The outputs of the differentiators in the first column, which are the first-order time derivatives (Doppler shifts), are redirected as the input of the second differentiator column. Figure 8-4 illustrates the implementation of such cascade differentiators.

The second-order time derivatives from the cascade differentiators generally have better performance in noise suppression; however, they are still too noisy when the sampling rate increases.

8.3 Alternative Methods of Ground Acceleration Determination

The acceleration determination described earlier is based on the virtual observables of the Doppler rate that are derived through specially designed differentiators. It is necessary that there are differentiators running for each tracked satellite. This is not trivial, especially when it is in real time and the sampling rate is high, since the system would experience very heavy computational load and the accuracy of the derivatives could degrade. So it is useful to develop alternative methods of ground acceleration determination.

Another reason to develop such alternatives is to maximally exploit the high precision of both the GPS Doppler shift and the carrier phase measurements. One may expect an improved accuracy of acceleration by derivation in the measurement domain than in the coordinate domain.

In this section, two alternatives of ground acceleration determination are presented.

8.3.1 Ground Acceleration from Doppler Observations

To introduce the concept one may write the Doppler observation equations at two consecutive epochs

$$\begin{aligned}\lambda D_r^s(t) = &\left[\mathbf{n}_r^s(t) + \frac{\dot{\mathbf{r}}^{\mathbf{s}}}{c}\right]\cdot[\dot{\mathbf{r}}^{\mathbf{s}}(t-\tau_r^s(t)) - \dot{\mathbf{r}}_r(t)] - d\dot{I}_r^s + d\dot{T}_r^s + c\cdot d\dot{t}_r(t) \\ &- c\cdot d\dot{t}^s(t-\tau_r^s(t)) + \frac{\Phi_0 - \Phi(\mathbf{r}(t))}{c} + \frac{2GM}{c}\left(\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|}\right) + d\dot{R}_{Sagnac} + \varepsilon_r^s\end{aligned} \qquad 8\text{-}6$$

$$\begin{aligned}\lambda D_r^s(t-\Delta t) = &\left[\mathbf{n}_r^s(t-\Delta t) + \frac{\dot{\mathbf{r}}^{\mathbf{s}}}{c}\right]\cdot[\dot{\mathbf{r}}^{\mathbf{s}}(t-\Delta t-\tau_r^s(t-\Delta t)) - \dot{\mathbf{r}}_r(t-\Delta t)] - d\dot{I}_r^s \\ &+ d\dot{T}_r^s + c\cdot d\dot{t}_r(t-\Delta t) - c\cdot d\dot{t}^s(t-\Delta t-\tau_r^s(t-\Delta t)) + \frac{\Phi_0 - \Phi(\mathbf{r}(t-\Delta t))}{c} \\ &+ \frac{2GM}{c}\left(\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|}\right) + d\dot{R}_{Sagnac} + \varepsilon_r^s\end{aligned} \qquad 8\text{-}7$$

The change of Doppler shift between these two epochs is due to the relative motion of receiver r and satellite s, and other differences such as the change-rates of the ionosphere and troposphere, and the measurement noise. Intuitively, the change of receiver velocity can be modelled by

$$\dot{\mathbf{r}}_r(t) = \dot{\mathbf{r}}_r(t-\Delta t) + \boldsymbol{\alpha}(t)\cdot\Delta t \qquad \text{8- 8}$$

where $\boldsymbol{\alpha}(t)$ is the receiver acceleration. More precisely, $\boldsymbol{\alpha}(t)$ is the average acceleration of receiver r during the time interval Δt. Differencing between the successive epochs and neglecting the relativistic terms and propagation errors

$$\begin{aligned}\nabla D_r^s\lambda = &\left[\mathbf{n}_r^s(t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot[\dot{\mathbf{r}}^s(t-\tau_r^s(t)] - \left[\mathbf{n}_r^s(t-\Delta t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot[\dot{\mathbf{r}}^s(t-\Delta t-\tau_r^s(t-\Delta t)] \\ &-\left[\mathbf{n}_r^s(t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot\dot{\mathbf{r}}_r(t) + \left[\mathbf{n}_r^s(t-\Delta t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot\left[\dot{\mathbf{r}}_r(t) - \boldsymbol{\alpha}(t)\cdot\Delta t\right] \\ &+c\cdot d\dot{t}^s(t-\Delta t-\tau_r^s(t-\Delta t)) - c\cdot d\dot{t}^s(t-\tau_r^s(t)) + c\cdot d\dot{t}_r(t) - c\cdot d\dot{t}_r(t-\Delta t) + \varepsilon\end{aligned} \qquad \text{8- 9}$$

This equation can then be rewritten into

$$\begin{aligned}\nabla D_r^s\lambda = &\left[\mathbf{n}_r^s(t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot[\dot{\mathbf{r}}^s(t-\tau_r^s(t)] - \left[\mathbf{n}_r^s(t-\Delta t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot[\dot{\mathbf{r}}^s(t-\Delta t-\tau_r^s(t-\Delta t)] \\ &-\left[\mathbf{n}_r^s(t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot\dot{\mathbf{r}}_r(t) + \left[\mathbf{n}_r^s(t-\Delta t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot\dot{\mathbf{r}}_r(t) - \left[\mathbf{n}_r^s(t-\Delta t) + \frac{\dot{\mathbf{r}}^s}{c}\right]\cdot\boldsymbol{\alpha}(t)\cdot\Delta t \\ &+c\cdot d\dot{t}^s(t-\Delta t-\tau_r^s(t-\Delta t)) - c\cdot d\dot{t}^s(t-\tau_r^s(t)) + c\cdot d\dot{t}_r(t) - c\cdot d\dot{t}_r(t-\Delta t) + \varepsilon\end{aligned} \qquad \text{8- 10}$$

In Eq.8-10 all the terms on the right hand side are known except the acceleration vector. The unknowns are the three components of the average ground acceleration. Therefore one can solve the receiver acceleration after the velocity determination by using a least-squares estimation scheme.

Since the contributions from the line-of-sight corrections for two consecutive epochs are numerically identical when the sampling rate is high, they vanish through the

differencing process. Thus one may simplify Eq.8-10 by neglecting the line-of-sight correction terms. This leads to

$$\begin{aligned}\nabla D_r^s\lambda = {} & \mathbf{n}_r^s(t)\cdot\dot{\mathbf{r}}^s[(t-\tau_r^s(t)] - \mathbf{n}_r^s(t-\Delta t)\cdot\dot{\mathbf{r}}^s[(t-\Delta t-\tau_r^s(t)] \\ & -\mathbf{n}_r^s(t)\cdot\dot{\mathbf{r}}_r(t) + \mathbf{n}_r^s(t-\Delta t)\cdot\dot{\mathbf{r}}_r(t) - \mathbf{n}_r^s(t-\Delta t)\boldsymbol{\alpha}(t)\cdot\Delta t \\ & + c\cdot d\dot{t}^s(t-\Delta t-\tau_r^s(t-\Delta t)) - c\cdot d\dot{t}^s(t-\tau_r^s(t)) \\ & + c\cdot d\dot{t}_r(t) - c\cdot d\dot{t}_r(t-\Delta t) + \varepsilon\end{aligned} \qquad 8\text{-}11$$

The satellite clock terms can be further simplified using the broadcast correction model, however they may be left in the above form since their values can be re-used if saved.

The resultant acceleration is the average for the entire period of *t-Δt* to *t*. The benefits of this method are obvious. It is free of cycle slips and the effect of the receiver clock reset. When the sapling rate increases, an averaging filter would be helpful to obtain a smoothed acceleration.

8.3.2 Ground Acceleration from the Carrier Phase Observations

In this section a method for ground acceleration determination based on carrier phase observations is described. The ground velocity is derived from the between-epoch differencing of the carrier phase observables. This velocity is the average speed of the receiver during the period between the two epochs. Acceleration is then derived from differentiation of the velocities with respect to time. This method suits receivers with carrier phase measurements, but without Doppler shift output.

8.3.2.1 Ground Velocity Determination

Similar to the approach of determining the acceleration from Doppler measurements, it is assumed that the change of carrier phase between two consecutive epochs is caused by the change of system states, i.e.

$$\mathbf{r}_r(t) = \mathbf{r}_r(t-\Delta t) + \mathbf{V}(t)\cdot\Delta t \qquad 8\text{-}12$$

where the receiver is regarded as travelling at a constant velocity *V(t)*. Note that significant modelling errors are present when the sampling space is too long or the system experiences high dynamics.

The observation equation for a carrier phase measurement is

$$\begin{aligned}\lambda\varphi_r^s &= \mathbf{n}_r^s(t)\cdot[\mathbf{r}^s(t-\tau_r^s(t)-\mathbf{r}_r(t)] - dI_r^s(t) + dT_r^s(t) + c\cdot dt_r(t)\\ &\quad - c\cdot dt^s(t-\tau_r^s(t)) + dR_r^s(t) - N\cdot\lambda + \varepsilon\end{aligned} \qquad 8\text{-}13$$

where dR_r^s stands for the relativistic corrections of the carrier phase measurement. The between-epoch difference of the carrier phase observables can be formed as

$$\begin{aligned}\nabla\varphi_r^s\lambda &= \mathbf{n}_r^s(t)\cdot[\mathbf{r}^s(t-\tau_r^s(t)-\mathbf{r}_r(t)] - \mathbf{n}_r^s(t-\Delta t)\cdot[\mathbf{r}^s(t-\Delta t-\tau_r^s(t-\Delta t)) - \mathbf{r}_r(t-\Delta t)]\\ &\quad -[dI_r^s(t) - dI_r^s(t-\Delta t)] + [dT_r^s(t) - dT_r^s(t-\Delta t)] + c\cdot[dt_r(t) - dt_r(t-\Delta t)]\\ &\quad - c\cdot[dt^s(t-\tau_r^s(t)) - dt^s(t-\Delta t-\tau_r^s(t-\Delta t))] + \varepsilon\end{aligned} \qquad 8\text{-}14$$

if no cycle slip occurs. This equation can be simplified by approximating the line-of-sight vectors associated with the receiver positions to that at the central time, and substituting the proceeding receiver position with the current receiver position, as follows

$$\begin{aligned}\nabla\varphi_r^s\lambda &\approx \mathbf{n}_r^s(t)\cdot\mathbf{r}^s(t-\tau_r^s(t)) - \mathbf{n}_r^s(t-\Delta t)\cdot\mathbf{r}^s(t-\Delta t-\tau_r^s(t-\Delta t)) + \mathbf{n}_r^s(t-\frac{1}{2}\Delta t)\mathbf{V}\Delta t\\ &\quad + c\cdot[dt_r(t) - dt_r(t-\Delta t)] - c\cdot[dt^s(t-\tau_r^s(t)) - dt^s(t-\Delta t-\tau_r^s(t-\Delta t))] + \varepsilon\end{aligned} \qquad 8\text{-}15$$

where $\boldsymbol{n}_r^s(t-\frac{1}{2}\Delta t)$ is the corrected line-of-sight vector at the central time. The receiver velocity can be therefore resolved by observing four or more satellites since there are only four unknowns.

Note that the velocity obtained through the Doppler shift measurements is an instantaneous velocity, while the velocity obtained from the carrier phase is with respect to the central time between two epochs, or in other words, an average velocity.

8.3.2 2 Ground Acceleration Determination

The ground acceleration may then be derived using a differentiator with the derived ground velocities as the filter input. It is preferred that the differentiator in our application be able to mitigate the effects of high-frequency noise. The output of the differentiator is a smoothed ground acceleration with a fixed time lag.

8.3.3 Kalman Filter Design for Acceleration Determination

Running Kalman filters for each tracked satellite to derive Doppler and Doppler rate involves a significant computational overload. However, if the receiver velocity has been obtained, then a Kalman filter may be used for ground acceleration determination in the velocity domain.

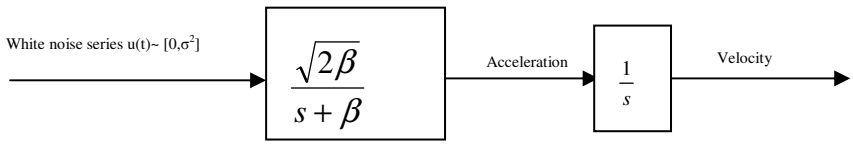

Figure 8- 5: Gauss-Markov process as the driving noise of acceleration

As the receiver may experience high dynamics, rather than just using a white noise or random walk process as the system driving noise, the Gauss-Markov random process model can be used to represent the system driving noise. Figure 8-5 illustrates the shaping process of white noise to become the driving noise of acceleration through the Gauss-Markov random process and the driving noise of velocity through a further integration. In this case the noise of the observed velocity is modelled by an

integrated Gauss-Markov random process, which is originally generated from a Gaussian white process.

The Gauss-Markov random process model uses an exponential autocorrelation function that has the ability to represent the correlations of system dynamics. That is, a close correlation represents the system in static or at constant speed, while a loose correlation suggests that the system is experiencing high dynamics.

The adopted Kalman filter has the following state equation in a continuous differential form

$$\begin{bmatrix} \dot{\mathbf{V}} \\ \dot{\boldsymbol{\alpha}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\beta \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \boldsymbol{\alpha} \end{bmatrix} + \begin{bmatrix} 0 \\ \sqrt{2\beta} \cdot \sigma \end{bmatrix} \qquad 8\text{-}16$$

where β is referred to as the correlation time constant which determines the degree of correlation of the representing random process, σ represents the Gaussian white noise.

The corresponding transition matrix for the state is given as [Brown and Hwang, 1992]

$$\Phi = \begin{bmatrix} 1 & \frac{1}{\beta}(1 - e^{-\beta \cdot \Delta t}) \\ 0 & e^{-\beta \cdot \Delta t} \end{bmatrix} \qquad 8\text{-}17$$

This Kalman filter is characterised with the integrated Gauss-Markov process as the system driving noise, which can be expressed by

$$\begin{aligned} \begin{bmatrix} \mathbf{V} \\ \boldsymbol{\alpha} \end{bmatrix}_{k+1} &= \begin{bmatrix} 1 & \frac{1}{\beta}(1 - \mathrm{e}^{-\beta \cdot \Delta t}) \\ 0 & \mathrm{e}^{-\beta \cdot \Delta t} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{V} \\ \boldsymbol{\alpha} \end{bmatrix}_{k} + \mathrm{w}_{k} \\ \tilde{\mathbf{V}} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \boldsymbol{\alpha} \end{bmatrix} + \mathrm{v}_{k} \end{aligned} \qquad 8\text{-}18$$

where the tilde bar denotes the vector of measurements, and the corresponding state driving noise covariance matrix $\boldsymbol{Q}$=E[w_k, w_k^T] is given by

$$\begin{bmatrix} \frac{2\sigma^2}{\beta}[\Delta t - \frac{2}{\beta}(1-e^{-\beta\Delta t}) + \frac{1}{2\beta}(1-e^{-2\beta\Delta t})] & 2\sigma^2(\frac{1}{\beta}(1-e^{-\beta\Delta t}) - \frac{1}{2\beta}(1-e^{-2\beta\Delta t})) \\ 2\sigma^2(\frac{1}{\beta}(1-e^{-\beta\Delta t}) - \frac{1}{2\beta}(1-e^{-2\beta\Delta t}) & \sigma^2(1-e^{-2\beta\Delta t}) \end{bmatrix} \quad 8\text{- }19$$

In the above model, the acceleration of the current epoch is related to the acceleration at the previous epoch by

$$\boldsymbol{\alpha}_{k+1} = e^{-\beta\cdot\Delta t} \cdot \boldsymbol{\alpha}_k + w_k \quad 8\text{- }20$$

It can be seen that by steering β to different values, various receiver dynamic states could be accommodated by this integrated Gauss-Markov process model. That is, for example, a large β indicates less correlation in the acceleration, which suggests that the system may change acceleration dramatically and thus be in a highly dynamic mode; a small β indicates high correlations in acceleration which implies that the system is in a stable mode. If β approaches infinity, the Gauss-Markov model will approach the random walk process. It is in this sense that one may anticipate that the Gauss-Markov model could be superior to the random walk model, and therefore could be used to accommodate different dynamics.

To use this Kalman filter one needs to predefine all the parameters, especially to assign a fixed value to β so as to suit the system dynamics. For a real-time application, it is most desirable that β can be adaptively determined to best fit the changes of state. This is referred to as adaptive Kalman filter design.

The adaptation can be accomplished by augmenting the correlation constant β into the state using extended Kalman filter techniques [Grewal et al., 2001]. In the implementation, it should be noted that even though β is a constant, it must be treated as a random variable [Chui and Chen, 1987, p.117]

$$\beta_{k+1} = \beta_k + \sigma_\beta \quad 8\text{- }21$$

where σ_β is a Gaussian white noise.

Then the augmented Kalman filter is given by

$$\begin{bmatrix} \mathbf{V} \\ \mathbf{A} \\ \beta \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & \frac{1}{\hat{\beta}}(1-e^{-\hat{\beta}\cdot\Delta t}) & (-\frac{1}{\hat{\beta}^2}+\frac{\Delta t}{\hat{\beta}}e^{-\hat{\beta}\cdot\Delta t}+\frac{1}{\hat{\beta}^2}e^{-\hat{\beta}\cdot\Delta t})\cdot\hat{\mathbf{V}} \\ 0 & e^{-\hat{\beta}\cdot\Delta t} & -\hat{\beta}\Delta t e^{-\hat{\beta}\cdot\Delta t}\cdot\hat{\mathbf{V}} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{A} \\ \beta \end{bmatrix}_k + \begin{bmatrix} w_k \\ w_\beta \end{bmatrix} \qquad 8\text{- }22$$

Equation 8-22 is obtained by calculating the dynamic Jacobian, which agrees with the dynamic model given by Grewal et al. [2001, p.211] when β is large. The corresponding Q-matrix is

$$\begin{bmatrix} \frac{2\sigma^2}{\beta}[\Delta t-\frac{2}{\beta}(1-e^{-\beta\Delta t})+\frac{1}{2\beta}(1-e^{-2\beta\Delta t})] & 2\sigma^2(\frac{1}{\beta}(1-e^{-\beta\Delta t})-\frac{1}{2\beta}(1-e^{-2\beta\Delta t})) & 0 \\ 2\sigma^2(\frac{1}{\beta}(1-e^{-\beta\Delta t})-\frac{1}{2\beta}(1-e^{-2\beta\Delta t}) & \sigma^2(1-e^{-2\beta\Delta t}) & 0 \\ 0 & 0 & \sigma_\beta^2 \end{bmatrix} \qquad 8\text{- }23$$

In conjunction with the velocity measurements, this Kalman filter can be used to derive acceleration in dynamic situations.

8.4 Summary

This chapter presents miscellaneous topics associated with velocity and acceleration determination. It outlines the causes of abnormalities in the Doppler shift or Doppler rate derivation. General procedures for handling receiver clock resets, blank observables and cycle slips were discussed. Doppler aiding for real-time cycle slip detection and estimation was presented.

Alternatives for ground acceleration determination are introduced. These methods have the capability of exploiting the high precision of the GPS system to generate ground acceleration without using differentiators in the measurement domain. Rather than deriving acceleration from the GPS-determined positions, which is generally a second-order differentiation process, the acceleration derived directly from precise velocities has a better accuracy. For this purpose, Kalman filtering techniques were

presented and an adaptive Kalman filter was proposed to suit changes of receiver dynamics.

Chapter Nine

CONCLUSIONS AND RECOMMENDATIONS

Position and velocity determination are associated with the scientific definition of navigation. An improvement in the accuracy of either may lead to an improvement of accuracy of the other. In this research, an intensive investigation into precise velocity determination using GPS has been conducted, with an aim to ensure positional accuracy improvement.

Issues in relation to real-time precise velocity and acceleration determination using GPS in standalone mode have been discussed in the context of this thesis. The following is a summary of the conclusions of this research and recommendations for future work.

9.1 Conclusions

As part of the navigation solution, a user's velocity can be determined using GPS in real time. This velocity, thanks to the latest hardware development, can be measured at a very high sampling rate. From a known starting point, the relative positions of the user can be determined through an integration process with high sampling rate velocities as the input.

The positions from such a scheme may have a better accuracy than those derived from the SPS within a prescribed period. The more accurate these velocities can be determined, the longer such a prescribed time will last. This research serves this

pursuit through a comprehensive investigation into high accuracy velocity determination using GPS.

It is found that among various methods of GPS velocity determination, the Doppler shift method has the advantage of producing ground velocity with the highest accuracy, and the resolved velocity is the instant velocity at the measurement epoch (if the delay due to signal processing time is negligible).

All the error sources that affect the positional accuracy in Precise Point Positioning have corresponding effects on velocity determination. The magnitude of a specific error source on Doppler shift measurements is the first time derivative of the source error in the range measurements.

It is demonstrated that the GPS satellite velocities from the broadcast ephemeris have better than 0.001m/s accuracy for each axis. Such an accuracy level of satellite velocity can be obtained using the rotational matrix method, differentiation of positions or position polynomials.

Through a comprehensive error analysis, it has been concluded that the relativistic effects are the largest error sources in a receiver to satellite Doppler shift measurement. An in-depth analysis of the relativistic Doppler effects provides intuitive views in the frequency domain, which is analogous to the Fourier transform. From the Doppler shift observation equation developed here, it can be seen that the high velocity of a satellite in orbit bends the receiver to satellite line-of-sight towards the satellite travel direction, causing a small frequency shift. The readily known orbit eccentricity correction for satellite clock is a periodic function in terms of orbit eccentricity, orbit semi-major axis, and the orbital eccentric anomaly, i.e. $\frac{2\sqrt{GM}}{c} e \cdot \sqrt{a} \cdot \sin E_k$ (see Eq.3-18) where the role of the orbit eccentricity is not obvious. However, in the frequency domain the corresponding term is in the form of

$\frac{2GM}{c}(\frac{1}{a_{orb}} - \frac{1}{\|\mathbf{r}^s\|})$. It is apparent that this relativistic term is caused by orbit eccentricity, since in a circular orbit it becomes zero. It also can be seen that the difference of the receiver gravitational potential to that at the geoid, $\frac{\Phi_0 - \Phi(\boldsymbol{r}_r)}{c}$, contributes an extra frequency shift. This frequency shift is characterised as being the same for all the Doppler shift measurements of satellites in view. Therefore the receiver potential difference only causes a biased receiver clock rate estimate, and it won't affect a user's velocity. Although the Earth's rotation correction is discussed and its correction formula is developed, a user can neglect it if an iteration scheme to account for the propagation time is adopted.

The ionospheric delay rate can be easily handled for dual-frequency GPS receivers. It can be eliminated by forming the "ionospheric free" Doppler measurements, or derived by a low-pass differentiator with the geometry-free carrier phase linear combination L_g as filter input.

It is difficult for single-frequency GPS users to correct for the ionospheric delay rate. Although the delay rate can be derived from the time series of "code–carrier phase", the accuracy is degraded due to the poor accuracy of code range measurements. The change-rate from the Klobuchar model also suffers from poor accuracy problems.

In this research, the tropospheric delay rate correction has been identified as requiring appropriate mapping functions that may best represent the change of the tropospheric delay along the signal profile. This requires precise geodetic mapping functions to be used. A simulation using Chao's model found that the delay rate on Doppler measurements is several centimetres per second in elevation range from 10° to 30°. Up to 60°, half centimetre per second delay rate remains. As such, it is better to use the tropospheric model with the highest accuracy available, and apply a differentiator to get the delay rate correction.

Since both the ionospheric and tropospheric delay rates contribute a few millimetres to centimetres per second errors to Doppler measurements, and since their changes are hard to predict and model, it is concluded that the atmosphere is the major error source degrading the velocity accuracy.

Ground velocity at sub-centimetre per second level is now achievable from GPS Doppler shift measurements. However, the velocities derived manifest a small bias and coloured noise due to the presence of unmodelled errors and an uneven distribution of GPS satellites.

The virtual Doppler rate "measurement" is rather clean since major error sources have no numerical effects on it except the real-time satellite orbital accelerations from the broadcast ephemeris, which is still better than ±0.1mm/s. Thus, ground accelerations can be potentially determined at a relatively high accuracy. Although the derived accelerations may be slightly biased, they have no drift compared to using an accelerometer.

The main problem associated with the Doppler rate method is that there are no Doppler rate observables in GPS receivers. It must be derived either from the carrier phase measurements or from the Doppler shift measurements. As the sampling rate increases, the noise corrupts the accuracy of the derived Doppler rates, and consequently degrades the ground acceleration results. In addition, a heavy computational load is required when using this method, demanding more powerful CPUs and memory. Therefore, it may be more economic and effective to employ an accelerometer when high sampling rate accelerations are required.

The methods used to determine acceleration using GPS are computationally cumbersome. The inclusion of acceleration for low dynamic motion modelling has minor contributions to the resolution of the state when high sampling position and velocity can be obtained. Since inertial acceleration can be directly sensed with accelerometers, it is more convenient to rely on the inertial sensors to obtain the

kinematic accelerations. GPS in this case, is more suitable for calibrating the sensed accelerations.

The IIR differentiators derived in this research are not subject to Doppler/Doppler rate derivation. They are ideal for many real-time applications owing to their characteristics, such as the broadband frequency response, linear phase response and short group delay, and easy implementation.

It is sensible to acknowledge that the errors of the GPS derived velocity and acceleration are not Gaussian white. Consequently, the adaptive Kalman filter designed in §8.3.3 applies only for the case that the driving noise of velocity complies with an integrated Gauss-Markov random process.

In summary, this research has continued and extended the velocity and acceleration determination using GPS to a higher level. All the tools and algorithms required for real-time precise velocity and acceleration determination using GPS have been developed with enough thoroughness and at the highest accuracy. Together they form the theoretical basis for such applications.

9.2 Recommendations for Future Work

It is recommended that further study of the multipath effect on the Doppler shift measurements be undertaken in order to draw a decisive conclusion.

Further work needs to be carried out into the implementation of the theories and approaches developed in this research. In particular, it is worthwhile to implement Neill's mapping function and the adaptive Kalman filtering developed in Chapter Eight into the C++ package and carry out real-world tests.

Future tests need to be carried out on the calibration of velocity bias and the recognition of velocity driving noises of any GPS system of interest. This is because any bias in the derived velocities would be accumulated through integration, causing

rapid divergence of the position solution. On the other hand, failure in recognition of the driving noise process of derived velocity will lead to a failure in Kalman filtering. However, such system recognition tends to be costly and time-consuming.

The investigation into the differentiator of type IV FIR filters is still at an early stage. Differentiator and integrator designs have been open topics for decades; one may easily find hundreds of papers in IEEE journals. Further research may be needed since this type of differentiator gives an almost ideal frequency response.

Further research into the relativistic "line-of-sight correction" term of the Doppler shift measurement is recommended. It is demonstrated in this thesis that the line-of-sight direction is tilted toward the satellite motion direction in the frequency domain. However, it is still unclear whether the satellite velocity term accidentally maps into the line-of-sight direction or its presence has more physical meaning. Since the frequency, distance and propagation time are inter-correlated with each other, it is sensible to investigate whether or not the line-of-sight direction is changed in a GPS measured distance.

It is recommended to investigate $\frac{\Phi_0 - \Phi(\mathbf{r}_r)}{c}$, the relativistic Doppler shift correction term induced by the receiver gravitational potential difference to the geoid. As the Earth's gravity field determination is one of the main tasks in geodesy, it is interesting that the Doppler frequency shift contains the Earth gravitational potential information. Due to the presence of c in the denominator, it is not possible to solve $\Phi(\mathbf{r}_r)$ directly with enough accuracy given the current GPS frequencies and measurement accuracy. The author believes that further research on how to achieve competitive accuracy of $\Phi(\mathbf{r}_r)$ using Doppler shift measurements from a satellite system could be useful.

Finally, further work can be done on determining the velocities of CHAMP and GRACE satellites by applying the theory and algorithms developed in this research. Although the high sampling rate Doppler observations of the Black Jack GPS

receivers are not provided in the public accessible archives, contact with the ISDC of CHAMP/GRACE at the GFZ may solicit Doppler observations that are worthwhile for scientific purposes.

APPENDIX

In this appendix, the algorithm of Klobuchar model is presented following Klobuchar [1996] and ARINC [2000], since it is used to analyse the correction of the ionospheric change-rate.

$$dI_{L1} = \begin{cases} c \cdot F \cdot \left(5.0 * 10^{-9} + (AMP)(1 - \frac{x^2}{2} + \frac{x^4}{24}) \right) & if \quad |x| < 1.57 \\ 5.0 * 10^{-9} \cdot c \cdot F & if \quad |x| \geq 1.57 \end{cases} \qquad \text{A- 1}$$

where:

$$AMP = \begin{cases} \sum_{n=0}^{3} \alpha_n \phi_m^n & if \quad AMP \geq 0 \\ 0 & if \quad AMP < 0 \end{cases} \qquad \text{A- 2}$$

$$x = \frac{2\pi(t - 50400)}{PER} \qquad (radians) \qquad \text{A- 3}$$

$$PER = \begin{cases} \sum_{n=0}^{3} \beta_{n\,m}^{\;n} & if \quad PER \geq 72{,}000 \\ 72{,}000 & if \quad PER < 72000 \end{cases} \qquad \text{A- 4}$$

$$F = 1.0 + 16.0 \cdot [0.53 - E]^3 \qquad \text{A- 5}$$

In the above equations, *AMP* and *PER* represent the amplitude and the period of the half-cosine wave respectively, α_i, and β_i are the broadcasting ionosphere parameters, t is the local time at the sub-ionospheric point, ϕ_m is the geomagnetic latitude of the sub-ionospheric point, E is the elevation of the satellite, and F is the slant factor which converts a vertical ionospheric delay to the line-of-sight direction.

The calculations of the variables that must be solved for in the above equations are as follows

- Calculate the Earth-centred angle, ψ

$$\psi = 0.0137/(E+0.11) - 0.022 \qquad (semicircle) \qquad \text{A- 6}$$

- Compute the sub-ionospheric latitude, ϕ_I

$$\phi_I = \left\{ \begin{array}{lll} \phi_u + \psi \cdot \cos A & & \\ if & \phi_I > +0.416 \quad then & \phi_I = +0.416 \\ if & \phi_I < -0.416 \quad then & \phi_I = -0.416 \end{array} \right\} \quad (semicircle) \qquad \text{A- 7}$$

where ϕ_u is the user's latitude.

- Then, compute the sub-ionospheric longitude, λ_I

$$\lambda_I = \lambda_u + (\psi \sin A / \cos \phi_I) \qquad (semicircle) \qquad \text{A- 8}$$

where λ_u is the user's longitude

- Find the geomagnetic latitude, ϕ_m, of the sub-ionospheric location looking toward each GPS satellite

$$\phi_m = \phi_u + 0.064 \cos(\lambda_I - 1.617) \qquad (semicircle) \qquad \text{A- 9}$$

- Calculate the local time, *t*, at the sub-ionospheric point

$$t = 4.32 \times 10^4 \lambda_I + GPStime \qquad seconds$$

$$t = \begin{cases} t = t - 86{,}400 & if \quad t > 86{,}400 \\ t = t + 86{,}400 & if \quad t < 0 \end{cases} \qquad \text{A- 10}$$

NOTATIONS, SYMBOLS AND ACRONYMS

Notations

- Matrices are in upper case and bold typeface
- Vectors are in lower case and bold typeface
- Variables are specified in italic in the context
- The following notations specify an arbitrary quantity *X*:
  - X true value
  - X_0 initial value
  - X_k value at epoch *k*
  - $X(t)$ value at epoch *t*
  - $\|\mathbf{X}\|$ normal or length of vector $\boldsymbol{X}$
  - $\hat{X}$ estimated value of *X*
  - $\dot{X}$ first derivative of *X* with respect to time
  - $\ddot{X}$ second derivative of *X* with respect to time
  - X^T transpose of *X*
  - X^{-1} inverse of *X*
- The following notations specify a GPS observation variable
  - X_r^s superscript *s* denotes satellite *s*, subscript *r* stands for receiver *r*, X_r^s specifies the variable associated with receiver *r* and satellite *s*

- $X^{s}_{r,i}$ subscript *i* specifies that the variable is at L_i frequency band: *1* for L1 (1575.42MHz) and 2 for L2 (1227.60MHz)
- X^{s} variable associated with satellite *s*
- X_{r} variable associated with receiver *r*

Symbols

Symbols are defined when used within the text. Commonly used symbols are listed below for quick reference

- α acceleration
- B geodetic latitude
- c speed of light in vacuum
- D Doppler frequency shift
- d representation of an error
- dI ionospheric error
- dT tropospheric error
- dτ clock error
- dM(dm) multipath error
- ε measurement noise
- E elevation angle
- f frequency
- G universal gravitation constant
- GM product of the universal gravitation constant G *and* the

- Earth's mass M
- Ω angular rate of the Earth's rotation
- P code (phase) range
- $\mathbf{Q}$ process driving noise matrix
- L geodetic longitude
- λ wavelength, or astronomic longitude
- $\mathbf{n}$ normal vector used to represent a line-of-sight unit vector
- N integer ambiguity
- φ carrier phase
- $\mathbf{r}$ position vector or range vector
- $\mathbf{R}$ measurement noise matrix in Kalman filtering
- $\boldsymbol{v}$ velocity
- ρ geometric range between receiver and satellite
- dR relativistic error
- Z variable for Z-transform

Acronyms

The following acronyms are used frequently. Their corresponding meanings are given as

AS	Anti-Spoofing
BPS	Bits per Second
BPSK	Bi-Phase Shift Key
C/A	Coarse/Acquisition

CEP	Circular Error of Probability
DD	Double-Difference
DGPS	Differential Global Positioning System
DLL	Delay Lock Loop
DoD	Department of Defense
DOP	Dilution of Precision
ECEF	Earth-Centred-Earth-Fixed
ECI	Earth-Centred-Inertial
FIR	Finite Impulse Response
FOC	Full Operational Capability
GAST	Greenwich Apparent Sidereal Time
GDOP	Geometric Dilution of Precision
GPS	Global Positioning System
HDOP	Horizontal Dilution of Precision
IERS	International Earth Rotation and Reference Service
IF	Intermediate Frequency
IGS	International GNSS (former GPS) service
IIR	Infinite Impulse Response
INS	Inertial Navigation System
L1	Primary GPS carrier signal frequency at 1575.42 MHz
L2	Secondary GPS carrier signal frequency at 1227.60MHz
NAVSTAR	Navigation System with Timing and Ranging
NCO	Numerically Controlled Oscillator
NGS	National Geodetic Service
PLL	Phase Lock Loop
PPK	Post-Processing Kinematic
PPS	Precise Positioning Service
PPP	Precise Point Positioning
PRN	Pseudo Random Number, used to identify GPS satellites
PVA	Positioning, Velocity and Acceleration

PVT	Positioning, Velocity and Timing
R&D	Research and Development
RF	Radio Frequency
RMS	Root Mean Square
RTD	Real-Time Differential
RTK	Real-Time Kinematic
SA	Selective Availability
SD	Single-Difference
SPS	Standard Positioning Service
SV	Space Vehicle
TEC	Total Electron Content
US	United States
UTC	Universal Time Coordinated
VLBI	Very Long Baseline Interferometry
WAAS	Wide Area Augmentation System
WGS-84	World Geodetic System 1984

REFERENCES

Al-Alaoui, M. A. (1992). "Novel Approach to Designing Digital Differentiators". *Electronics Letters,* 28(15):pp.1376-1378.

Al-Alaoui, M. A. (1993). "Novel Digital Integrator and Differentiator". *Electronics Letters,* 29(4):pp.376-378.

Al-Alaoui, M. A. (1994). "Novel IIR Differentiator from the Simpson Integration Rule". *IEEE Transactions on Circuits and System I: Fundamental, Theory and Applications,* 41(2):pp.186-187.

Al-Alaoui, M. A. (1995). "A Class of Second-Order Integrators and Low-Pass Differentiators". *IEEE Transactions on Circuits and System I :Fundamental , Theory and Applications,* 42(4):pp.220-223.

Antoniou, A. (1979). *Digital Filters: Analysis and Design.* McGraw-Hill Book Company, ISBN: 0-07-002117-1.

ARINC (2000). "Interface Control Document, Navstar GPS Space Segment/Navigation User Interfaces, ICD-GPS-200, Revision C". Arinc Research Corporation, IRN-200C-004. http://www.navcen.uscg.gov/gps/geninfo/ICD-GPS-200C%20with%20IRNs%2012345.pdf, Accessed on 15/05/2006.

Ashby, N. (2003). "Relativity in the Global Positioning System". Living Review, Relativity, Online Article (6). http://www.livingreviews.org/Articles/Volume6/2003-ashby/. Accessed on 25/02/2004.

Ashby, N. and J. J. Spilker (1996). "Introduction to Relativistic Effects on the Global Positioning System". In B. W. Parkinson (Ed.) *Global Positioning System: Theory and Applications.* AIAA, Inc, (I):pp.623-698. ISBN:1-56347-106-X.

Ashjaee, J. and R. Lorenz (1992). "Precision GPS Surveying After Y-Code". *ION GPS 92: The Fifth International Technical Meeting of the Satellite Division of the Institute of Navigation. Sept 16-18, 1992,* Albuquerque, New Mexico. pp.657-659.

Askne, J. and H. Nordius (1987). "Estimation of Tropospheric Delay for Microwaves from Surface Weather Data". *Radio Science,* 22(3):pp.379-386.

Balard, N., R. Santerre, M. Cocard and S. Bourgon (2006). "Single GPS Receiver Time-Relative Positioning with Loop Misclosure Corrections". *GPS Solutions,* 10:pp.56-62.

Bancroft, S. (1985). "An Algebraic Solution to the GPS Equations". *IEEE Trans. Aerosp. and Elec. Systems,* 21(7):pp.56-59.

Beutler, G. (1998). "GPS Satellite Orbits". In P. J. G. Teunissen and A. Kleusberg (Ed.) *GPS for Geodesy.* Springer-Verlag Berlin Heidelberg, Berlin, Germany. pp.43-110. ISBN:3-540-63661-7.

Beutler, G., I. I. Mueller and R. E. Neilan (1994). "The International GPS Service for Geodynamics (IGS): Development and Start of Official Service on January 1, 1994". *Bulletin Geodesique,* 68(1):pp.43-46.

Beyer, W. H. (1980). *Handbook of Mathematical Sciences.* The fifth edition.CRC Press, ISBN: 0849306558.

Blewitt, G. (1998). "GPS Data Processing Methodology". In P. J. G. Teunissen and A.Kleusberg (Ed.) *GPS for Geodesy.* Springer, pp.231-270. ISBN:3-540-63661-7.

Bock, Y. (1998). "Reference Systems". In P. J. G. Teunissen and A. Kleusberg (Ed.) *GPS for Geodesy.* Springer-Verlag, Berlin, Germany. pp.1-42. ISBN:3-540-63661-7.

Bona, P. (2000). "Precision, Cross Correlation, and Time Correlation of GPS Phase and Code Observations". *GPS Solutions,* 4(2):pp.3-13.

Booch, G., J. Rumbaugh and I. Jacobson (1999). *The Unified Modeling Language User Guide.* Addison-Wesley, ISBN: 0201571684.

Borre, K. and C. Tiberius (2000). "Time Series Analysis of GPS Observables". *Proceedings of the ION GPS 2000. 19-22, Sept,* Salt Lake City, UT. pp.1885-1894.

Braasch, M. S. (1996). "Multipath Effects". In B. W. Parkinson (Ed.) *Global Positioning System: theory and applications.* Washington DC, USA. (I):pp.547-568. ISBN:1-56347-106-X.

Brown, R. G. and P. Y. C. Hwang (1992). *Introduction to Random Signals and Applied Kalman Filtering.* John Wiley &Sons, INC, ISBN: 0471-52573-1.

Brozena, J. M., G. L. Mader and M. F. Peters (1989). "Interferometric Global Positioning System: Three-Dimensional Positioning Source for Airborne Gravimetry". *Journal of Geophysical Research,* 94:pp.12153-12162.

Bruton, A. M. (2000). "Improving the Accuracy and Resolution of SINS/DGPS Airborne Gravimetry". *PhD Dissertation*. Dept. of Geomatics Engineering, The University of Calgary.

Bruton, A. M., C. L. Glennie and K. P. Schwarz (1999). "Differentiation for High-Precision GPS Velocity and Acceleration Determination". *GPS Solutions,* 2(4):pp.7-21.

Bruton, A. M. and K. R. Schwarz (2002). "Deriving Acceleration from DGPS: Toward Higher Resolution Applications of Airborne Gravimetry". *GPS Solutions,* 5(3):pp.1-14.

Calvert, J. B. (2004). "Doppler Effect: Classical and Relativistic Analysis of the Effects of Motion on Waves". http://www.du.edu/~jcalvert/phys/doppler.htm, Accessed on 20/08/2004.

Cannon, M. E., G. Lachapelle, M. C. Szarmes, J. M. Hebert, J. Keith and S. Jokerst (1998). "DGPS Kinematic Carrier Phase Signal Simulation Analysis for Precise Velocity and Position Determination". *Journal of the Institute of Navigation,* 44(2):pp.231-245.

Carlsson, B. (1991). "Maximum Flat Digital Differentiator". *Electronics Letters,* 27(8):pp.675-677.

Carlsson, B., A. Ahlen and M. Sternad (1991). "Optimal Differentiation Based on Stochastic Signal Models". *IEEE Transactions on Signal Processing,* 39(2):pp.341-353.

Chao, C. C. (1974). "The Tropospheric Calibration Model for Mariner Mars 1971". *Tech. Rep.* Jet Propulsion Laboratory, pp.61-76.

Chen, C. K. and J. H. Lee (1995). "Design of High-Order Digital Differentiators Using L1 Error Criteria". *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing,* 42(4):pp.287-291.

Chen, C. T. (2001). *Digital Signal Processing: Spectral Computation and Filter Design.* Oxford University Press, ISBN: 0-19-513638-1.

Chui, C. K. and G. Chen (1987). *Kalman Filtering with Real-Time Applications.* Springer-Verlag, ISBN: 3-540-18395-7.

Collins, J. P. and R. B. Langley (1996). "Mitigating Tropospheric Propagation Delay Errors in Precise Airborne GPS Navigation". *Proceedings of PLANS'90. IEEE. 22-26, Apr,* Atlanta, CA. pp.582-589.

Collins, J. P. and R. B. Langley (1997). "Estimating the Residual Tropospheric Delay for Airborne Differential GPS Positioning". *10th Int. Tech. Meeting of the*

Satellite Division of the US Inst. of Navigation. 16-19 Sept., Kansas City, Missouri, USA. pp.1197-1206.

Collins, P., R. Langley and J. LaMance (1996). "Limiting Factors in Tropospheric Propagation Delay Error Modelling for GPS Airborne Navigation". *Proceedings of the 52nd ION Annual Meeting. 19-21 June,* Cambridge, MASS, USA. pp.519-528.

Comp, C. J. and P. Axelrad (1996). "An Adaptive SNR-Based Carrier Phase Multipath Mitigation Technique". *Proceeding of the ION GPS-96. 17-20,Sept,* Kansas City, Missouri, USA. pp.683-697.

Davis, J. L., T. A. Herring, I. I. Shapiro, A. E. E. Rogers and G. Elgered (1985). "Geodesy by Radio Interferometry: Effects of Atmospheric Modeling Errors on Estimates of Baseline Length". *Radio Science,* 20:pp.1593-1607.

DoD (1996). "NAVSTAR GPS User Equipment Introduction". *Public Release Version*. Department of Defense, USA, www.spacecom.af.mil/usspace/gps_support/documents/gpsuser.pdf. Accessed on 22/06/2003.

Ellum, C. and N. E. Sheimy (2002). "Inexpensive Kinematic Attitude Determination from MEMS-Based Accelerometers and GPS-Derived Accelerations". *Journal of the Institute of Navigation,* 49(3):pp.117-127.

EL-Rabbany, A. (2002). *Introduction to GPS: the Global Positioning System.* Artech House, ISBN: 1-58053-183-1.

Farlex (2004). "TheFreeDictionaryCom:Kaiser window". http://encyclopedia.thefreedictionary.com/Kaiser%20window, Accessed on 07/07/2006.

Farrell, J. A. and M. Barth (1999). *The Global Positioning System & Inertial Navigation.* McGraw-Hill, ISBN: 0-07-022045-X.

Fenton, P. and B. Townsend (1994). "NovAtel Communications Ltd,-What's New?" *Proceedings of KIS'94. 30, Aug-2,Sept,* Banff, AB, CA. pp.25-29.

Ford, T. J. and J. Hamilton (2003). "A New Positioning Filter: Phase Smoothing in the Position Domain". *Navigation. Journal of the Institute of Navigation,* 50(2):pp.65-78.

Georgiadou, Y. and A. Kleusberg (1988). "On Carrier Signal Multipath Effects in Relative GPS Positioning". *Manuscripta Geodaetica,* Vol.13:pp.172-179.

Goad, C. C. (1998). "Single-Site GPS Models". In P. J. G. Teunissen and A. Kleusberg (Ed.) *GPS for Geodesy.* Springer-Verlag, Berlin, Germany. pp.437-456. ISBN:3-540-63661-7.

Goad, C. C. and L. Goodman (1974). "A Modified Hopfield Tropospheric Refraction Correction Model". *Presented at the Fall Annual Meeting of the American Geophysical Union. 12-17, Dec,* San Francisco, California, USA.

Goodman, M. and S. Jacques (2000). "Fact Sheet: Civilian Benefits of Discontinuing Selective". Commerce News: United States Department of Commerce. Washington, DC. 20230l. http://www.igeb.gov/sa/benefits.shtml, Accessed on 10/07/2002.

Grant, D. B. (1990). "Combination of Terrestrial and GPS Data for Earth Deformation Studies". *PhD thesis*. School of Surveying and Spatial Information Systems, University of New South Wales.

Grant, D. B., A. Stolz, B. Merminod and C.C. Mazur (1990). "Contributions to GPS Studies". In C.Rizos (Ed.). Sydney, School of Surveying, Unisurv S-38, UNSW, ISBN: 0-85839-056-6.

Grewal, M. S., L. R. Weill and A. P. Andrews (2001). *Global Positioning System, Inertial Navigation, and Integration.* John Wiley &Sons, Inc, ISBN: 0471-35032-X.

Gurtner, W. (2001). "RINEX: the Receiver Independent Exchange Format Version 2.10". http://www.ngs.noaa.gov/CORS/Rinex2.html, Accessed on 09/08/2004.

Hamming, R. W. (1977). *Digital Filters.* Prentice-Hall, Inc., ISBN: 0-13-212571-4.

Harvey, N. (2004). "Doppler and Carrier Phase Measurements in Trimble 5700 GPS Receiver. Trimble Support (RQST00000334956)". Private Communication, Christchurch, New Zealand.

Hebert, C. J., J. Keith, S. Ryan, M. Szarmes, G. Lachapelle and M. E. Canon (1997). "DGPS Kinematic Carrier Phase Signal Simulation Analysis for Precise Aircraft Velocity Determination". *Proceedings of the ION GPS-97. 1-11,July,* Albuquerque, NM. pp.335-350.

Herring, T. A. (1992). "Modeling Atmospheric Delays in the Analysis of Space Geodetic Data". *Proceedings of the Symposium on Refraction of Transatmospheric Signals in Geodesy.* Netherlands Geodetic Commission, Publications on Geodesy. pp.157-164.

Hofmann-Wellenhof, B., H. Lichtenegger and J. Collins (2001). *Global Positioning System Theory and Practice.* 5th edition.Springer-Verlag/Wien New York, ISBN: 3-211-83534-2.

Hopfield, H. S. (1969). "Two-Quadratic Tropospheric Refractivity Profile for Correcting Satellite Data". *Journal of Geophysical Research,* 74(18):pp.4487-4499.

IAU (2000). "IAU Resolutions Adopted at the 24th General Assembly". http://danof.obspm.fr/IAU_resolutions/Resol-UAI.doc. Accessed on 08/12/2004.

Ifadis, I. M. (2000). "A New Approach to Mapping the Atmospheric Effect for GPS observations". *Letter: Earth Planets Space,* 52:pp.703-709.

Jaldehag, K., C. Thomas and J. Azoubib (1998). "Use of a Dual-Frequency Multi-channel Geodetic GPS Receiver for the Estimation of Ionospheric Delays Applied to Accurate Time Transfer". *12th European Frequency and Time Forum. 10-12,Mar,* Warzawa, Poland. pp.499-504.

Jekeli, C. (1994). "On the Computation of Vehicle Accelerations Using GPS Phase Accelerations". *International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation. 30, Aug-2,Sept,* Banff, Canada. pp.473-481.

Jekeli, C. and R. Garcia. (1997). "GPS Phase Accelerations for Moving-base Vector Gravimetry". *Journal of Geodesy,* 71(10):pp.630-639.

Kavanagh, R. C. (2001). "FIR Differentiators for Quantized Signals". *IEEE Transactions on Signal Processing,* 49(11):pp.2713-2720.

Kennedy, S. (2002). "Precise Acceleration Determination from Carrier Phase Measurements". *Proceeding of the ION GPS-2002. 24-27, Sept,* Portland Oregon, USA. pp.962-972.

Kennedy, S. L. (2003). "Precise Acceleration Determination from Carrier-Phase Measurements". *Navigation, Journal of the Institute of Navigation,* 50(1):pp.9-19.

Khan, I. R. and R. Ohba (1999). "New Design of Full Band Differentiators Based on Taylor Series". *IEE Proceedings-Vision, Image and Signal Processing,* 146(4):pp.185-189.

Khan, I. R., R. Ohba and N. Hozumi (2000). "Mathematical Proof of Explicit Formulas for Tap-Coefficients of Full-Band FIR Digital Differentiators". *IEE Proceedings-Vision, Image and Signal Processing,* 147(6):pp.553-555.

Kleusberg, A., D. Peyton and D. Wells (1990). "Airbornegravimetry and the Global Positioning System". *IEEE PLAN'S90. 20-23,Mar,* Las Vegas, NV. pp.273-278.

Klobuchar, J. A. (1996). "Ionospheric Effects on GPS". In B. W. Parkinson and J. J. Spilker (Ed.) *Global Positioning System: Theory and Applications.* American Institute of Aeronautics and Astronautics Inc., Washington DC. (I):pp.485-514. ISBN:1-56347-106-X.

Kornfeld, R. P., R. J. Hansman and J. J. Deyst (1998). "Single-Antenna GPS-Based Aircraft Attitude Determination". *Navigation, Journal of the Institute of Navigation,* 45(1):pp.51-60.

Kouba, J. (2003). "A Guide to Using International GPS Service (IGS) Products". Natural Resources Canada, Ottawa. http://igscb.jpl.nasa.gov/igscb/resource/pubs/GuidetoUsingIGSProducts.pdf. Accessed on 20/03/2007.

Kouba, J. and P. Heroux (2001). "Precise Point Positioning Using IGS Orbit and Clock Products". *GPS Solutions,* 5(2):pp.12-28.

Kumar, B. and S. C. D. Roy (1988a). "Coefficients of Maximally Linear, FIR Digital Differentiators for Low Frequencies". *Electronics Letters,* 24(9):pp.563-565.

Kumar, B. and S. C. D. Roy (1988b). "Design of Digital Differentiators for Low Frequencies". *Proceedings of the IEEE,* 76(3):pp.287-289.

Kumar, B. and S. C. D. Roy (1989a). "Design of Efficient FIR Digital Differentiators and Hilbert Transformers for Midband Frequency Ranges". *Int.J.Circuit Theory Applic.,* 17(4):pp.483-488.

Kumar, B. and S. C. D. Roy (1989b). "Maximally Linear FIR Digital Differentiators for Midband Frequencies". *Int.J.Circuit Theory Applic.,* 17:pp.21-27.

Langley, R. B. (1998). "Propagation of the GPS Signals". In P. J. G. Teunissen and A. Kleusberg (Ed.) *GPS for Geodesy.* Springer-Verlag Berlin Heidelberg, Berlin, Germany. pp.111-149. ISBN:3-540-63661-7.

Leick, A. (1995). *GPS Satellite Surveying.* 2nd.John Wiley & Sons, ISBN: 0-471-30626-6.

Leva, J. L., M. U. d. Haag and K. V. Dyke (1996). "Performance of Standalone GPS". In E. D. Kaplan (Ed.) *Understanding GPS: Principles and Applications.* Artech House, Boston, London. pp.237-320. ISBN:0890067937.

Lyons, R. G. (2004). *Understanding Digital Signal Processing.* Prentice Hall PTR, ISBN: 0-13-108989-7.

Marini, J. W. (1972). "Correction of Satellite Tracking Data for an Arbitrary Tropospheric Profile". *Radio Science,* 7(2):pp.223-231.

Marshall, J. (2002). "Creating and Viewing Skyplots". *GPS Solutions,* 6(1-2):pp.118-120.

McCarthy, D. (2000). "IERS Conventions 2000". IERS Technical Notes, International Earth Rotation Service.

Mendes, V. B. (1999). "Modeling the Neutral Atmosphere Propagation Delay in Radiometric Space Techniques". *PhD Dissertation*. Department of Geodesy and Geomatics Engineering, University of New Brunswick.

Michaud, S. and R. Santerre (2001). "Time-Relative Positioning with a Single Civil GPS Receiver". *GPS Solutions,* 5(2):pp.71-77.

Misra, P. and P. Enge (2001). *Global Positioning System: Signals, Measurements, and Performance.* 1st edition.Ganga-Jamuna Press, ISBN: 0-9709544-0-9.

Montenbruck, O. and E. Gill (2000). *Satellite Orbits: Models, Methods, and Applications.* Springer-Verlag, ISBN: 3-540-67280-x.

Mueller, J. and G. Seeber (2004). "Re: Questions on Relativistic Effects". Private Communication.

Niell, A. E. (1996). "Global Mapping Functions for the Atmosphere Delay at Radio Wavelengths". Journal of Geophysical Research, B2-101. pp.3227-3246. ftp://web.haystack.edu/pub/aen/nmf/NMF_JGR.pdf. Accessed on 12/03/05.

Niell, A. E. (2000). "Improved Atmospheric Mapping Functions for VLBI and GPS". *Letter: Earth Planets Space,* 52:pp.699-702.

NIMA (2000). "Department of Defence World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems". National Imagery and Mapping Agency, TR8350.2. http://164.214.2.59/GandG/tr8350_2.html.

Parkinson, B. W. (1996). "GPS Error Analysis". In B. W. Parkinson (Ed.) *Global Positioning System: Theory and Applications.* Washington DC, USA. (I):pp.469-483. ISBN:1-56347-106-X.

PolaRx2 (2004). "PolaRx2 GPS Receiver Specifications". Septentrio Satellite Navigation, www.septentrio.com, Accessed on 29/09/2005.

Psiaki, M. L., S. P. Powell and P. M. J. Kinter (1999). "The Accuracy of the GPS-Derived Acceleration Vector, a Novel Attitude Reference". *Proceedings of the 1999 AIAA Guidance, Navigation, and Control Conference. August,* Portland,OR. pp.751-760.

Psiaki, M. L., S. P. Powell and P. M. J. Kintner (2000). "The Accuracy of The GPS-Derived Acceleration Vector". *Journal of Guidance, Control, and Dynamics 2000,* 23(3):pp.532-538.

Rabiner, L. and K. Steiglitz (1970). "The Design of Wide-Band Recursive and Nonrecursive Digital Differentiators". *IEEE Transactions on Audio and Electroacoustics,* 18(2):pp.204-209.

Remondi, B. (1991). "NGS Second Generation ASCII and Binary Orbit Formats and Associated Interpolation Studies". *Paper presented at the XX General Assembly of the IUGG. August 11-24,* Vienna, Austria.

Remondi, B. W. (2004). "Computing Satellite Velocity Using the Broadcast Ephemeris". *GPS Solutions,* 8(3):pp.181-183.

Rizos, C. (1999). "Principles and Practice of GPS Surveying". Sydney, AUS. http://www.gmat.unsw.edu.au/snap/gps/gps_survey/principles_gps.htm, Accessed on 23/03/2007.

Rothacher, M., L. Mervart, G. Beutler, E. Brockmann, S. Fankhauser, W. Gurtner, J. Johnson, S. Schaer, T. Springer and R. Weber (1996). "Bernese GPS Software Version 4.0". Astronomical Institute, University of Berne, pp.418.

Roulston, A. J. (2001). "High Precision Real-Time GPS Positioning Using Prompt Orbital Information". *Masters Dissertation*. Dept. of Geospatial Science, RMIT University.

Rousmaniere, J. (2003). "Navigation Introduction". Microsoft Encarta Reference Library 2003. Accessed on 12/10/2003.

Saastamoinen, J. (1972). "Atmospheric Correction for the Troposphere and Stratosphere in Radio Ranging of Satellites". *Geophysical Monograph 15.* American Geophysical Union, Washington, DC, 1972.

Saastamoinen, J. (1973). "Contribution to the Theory of Atmospheric Refraction". *Bulletin Geodesique,* 107:pp.34.

Seeber, G. (1993). *Satellite Geodesy: Foundations, Methods, and Applications.* Walter de Gruyter, ISBN: 3-11-012753-9.

Seeber, G. (2003). *Satellite Geodesy.* 2nd edition.Walter de Gruyter GmbH & Co, ISBN: 3-11-017549-5.

Selesnick, I. W. (2002). "Maximally Flat Low-pass Digital Differentiator". *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing,* 49(3):pp.219.

Serrano, L., D. Kim, R. B. Langley, K. Itani and M. Ueno (2004). "A GPS Velocity Sensor: How Accurate Can It Be? A First Look". *Proceedings of ION NTM 2004.* San Diego, California, USA. pp.875-885.

Simsky, A. and F. Boon (2003). "Carrier Phase and Doppler-Based Algorithms for Real-Time Standalone Positioning". *Proceedings of GNSS 2003, The European Navigation Conference. 22-25 April,* Graz, Austria. http://www.septentrio.com/papers/DARTS_GNSS2003.pdf. Accessed on 20/03/2007.

SMC/GP, J. (2004). "Control Segment Program Office". NAVSTAR GPS Joint Program Office, https://gps.losangeles.af.mil/control. Accessed on 07/05/2004.

Spilker, J. J. (1996a). "Fundamentals of Signal Tracking Theory". In B. W. Parkinson and J. J. Spilker (Ed.) *Global Positioning System: Theory and Applications.* American Institute of Aeronautics and Astronautics, Inc., Washington DC. (I):pp.245-328. ISBN:1-56347-106-X.

Spilker, J. J. (1996b). "GPS Signal Structure and Theoretical Performance". In B. W. Parkinson and J. J. Spilker (Ed.) *Global Positioning System: Theory and Applications.* American Institute of Aeronautics and Astronautics, Inc., Washington, DC. (I):pp.57-119. ISBN:1-56347-106-X.

Spilker, J. J. (1996c). "Tropospheric Effects on GPS". In B. W. Parkinson and J. J. Spilker (Ed.) *Global Positioning System: Theory and Applications.* American Institute of Aeronautics and Astronautics, Inc., Washington DC. (I):pp.517-546. ISBN:1-56347-106-X.

Springer, T. A. and U. Hugentobler (2001). "IGS Ultra Rapid Products for (Near-) Real-Time Applications". *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy,* 26(6-8):pp.623-628.

Stearns, S. D. (2003). *Digital Signal Processing with Examples in MATLAB.* CRC press, ISBN: 0-8493-1091-1.

Stenbit, J. P. (2001). "Global Positioning System Standard Positioning Service Performance Standard". *Public Release*. Assistant Secretary of Defence for Command, Control, Communications, and Intelligence, http://www.navcen.uscg.gov/gps/geninfo/2001SPSPerformanceStandardFINAL.pdf. Accessed on 23/03/2007.

SuperstarII (2004). "Superstar II Product Brochure". 2004: http://www.seabed.nl/pdf/superstar.pdf.

Szarmes, M., S. Ryan, G. Lachappelle and P. Fenton (1997). "DGPS High Accuracy Aircraft Velocity Determination Using Doppler Measurements". *Proceedings of KIS-97. 3-6,June,* Banff, AB, Canada.

Teunissen, P. J. G. and A. Kleusberg (1998). "GPS Observation Equations and Positioning Concepts". In P. J. G. Teunissen and A. Kleusberg (Ed.) *GPS for Geodesy.* Springer, pp.187-230. ISBN:3-540-63661-7.

Thayer, G. D. (1974). "An Improved Equation for the Radio Refractive Index of Air". *Radio Science,* 9(10):pp.803-807.

Ulmer, K., P. Hwang, B. Disselkoen and M.Wagner (1995). "Accurate Azimuth from a Single PLGR+GLS DoD GPS Receiver Using Time Relative Positioning". *Proceedings of ION GPS-95.* Palm Springs, CA, USA. pp.1733-1741.

Vainio, O., M. Renfors and T. Saramaki (1997). "Recursive Implementation of FIR Differentiators with Optimum Noise Attenuation". *IEEE Transactions on Instrumentation and Measurements,* 46:pp.1202-1207.

Van Dierendonck, A. J. (1996). "GPS Receivers". In B. W. Parkinson and J. J. Spilker (Ed.) *Global Positioning System: Theory and Applications.* American Institute of Aeronautics and Astronautics, Inc., Washington DC. (1):pp.329-408. ISBN:1-56347-106-X.

Van Dierendonck, A. J., M. E. Cannon, M. Wei and K. P. Schwarz (1994). "Error Sources in GPS-Derived Acceleration for Airborne Gvavimetry". *Proceedings of ION NTM-94. 24-26,JAN,* San Diego, USA. pp.811-820.

Van Graas, F. and A. Soloview (2003). "Precise Velocity Estimation Using a Stand-Alone GPS Receiver". *Proceedings of ION NTM 2003. 22-24 January 2003,* Anaheim, California, USA. pp.262-271.

VBOX (2004). "VBOX II 100Hz Speed Sensor". Product Brochure, 2004. http://www.m-techautomotive.co.uk/vbox/downloads/vb2sps100_data.pdf. Accessed on 23/03/2007.

Wei, M. and K. P. Schwarz (1995). "Analysis of GPS-Derived Acceleration from Airborne Tests". *IAG Symposium G4, IUGG XXI General Assembly. 2-14,July,* Boulder, Colorado, USA.

Wells, D. E. (1974). "Doppler Satellite Control". University of New Brunswick,

Wells, D. E., B. N. Beck, D. Delikaraoglou, A. Kleusberg, E. J. Krakiwsky, G. Lachapelle, R. B. Langley, M. Nakiboglu, K. P. Schwarz, J. M. Tranquilla and P. Vanicek (1987). *Guide to GPS Positioning.* ISBN: 0-920-114-73-3.

Wikipedia (2004). "Doppler Effect: from Wikipedia, the Free Encyclopaedia". Wikipedia, http://en.wikipedia.org/wiki/Doppler_effect. Accessed on 20/03/2007.

Witchayangkoon, B. (2000). "Elements of GPS Precise Point Positioning". *PhD Dissertation*. Spatial Information Science and Engineering, The University of Maine.

Wolfe, J. D., W. R. Williamson and J. L. Speyer (2003). "Hypothesis Testing for Resolving Integer Ambiguity in GPS". *Navigation, Journal of the Institute of Navigation,* 50(1):pp.45-56.

Wu, J. T., S. C. Wu, G. A. Hajj, W. I. Bertiger and S. M. Lichten (1993). "Effects of Antenna Orientation on GPS Carrier Phase". *Manuscripta Geodetica,* 18:pp.1647-1660.

Xu, G. (2003). *GPS: Theory, Algorithms and Applications.* Springer, ISBN: 3-540-67812-3.

Xu, S. (1996). *Collection of Commonly used C Algorithms.* Press House of Tsinghua University, ISBN: 7-302-02290-9/TP.1128.

Yang, M. and K.-H. Chen (2001). "Performance Assessment of a Noniterative Algorithm for Global Positioning System (GPS) Absolute Positioning". *Proceedings of National Science Council, ROC(A).* pp.102-106.

Zhang, J., K. Zhang, R. Grenfell and R. Deakin (2006a). "GPS Satellite Velocity and Acceleration Determination Using the Broadcast Ephemeris". *Journal of Navigation,* 59:pp.293-305.

Zhang, J., K. Zhang, R. Grenfell and R. Deakin (2006b). "On the Relativistic Doppler Effect for Precise Velocity Determination Using GPS". *Journal of Geodesy,* 80(2):pp.104-110.

Zhang, J., K. Zhang and R. Grenfell (2005a). "On Real-Time High Precision Velocity Determination for Standalone GPS Users". *Survey Review,* Accepted on 04/03/2005.

Zhang, J., K. Zhang, R. Grenfell, Y. Li and R. Deakin (2005b). "Real-Time Doppler/Doppler Rate Derivation for Dynamic Applications". *Journal of GPS,* 4(1-2): Paper 12. http://www.gmat.unsw.edu.au/wang/jgps/v4n12/index_v4n12.htm. Accessed on 12/06/2006.

Zhang, K., R. Deakin, R. Grenfell, Y. Li, J. Zhang, W. N. Cameron and D. M. Silcock (2004). "GNSS for Sports: Sailing and Rowing Perspectives". *Proceedings of the International Symposium on GPS/GNS. Dec. 6-8, 2004,* Sydney, Australia. Paper(75).

Zhang, K., R. Grenfell, R. Deakin, Y. Li and J. Zhang (2003a). "Current Development of a Low-Cost, High Output Rate RTK GPS Multisensor System for Rowers". *Proceedings of SatNav 2003.* Melbourne, Australia. Paper (75),

Zhang, K., R. Grenfell, R. Deakin, Y. Li, J. Zhang, A. Hahn, C. Gore and T. Rice (2003b). "Towards a Low-cost, High Output Rate, Real Time GPS Rowing Coaching and Training System". *Proceedings of the ION GPS/GNSS 2003.* Portland, Oregon, USA. pp.489-498.