→ GNSS DATA PROCESSING

Volume I: Fundamentals and Algorithms

www.esa.int

European Space Agency

TM-23/1
May 2013

→ GNSS DATA PROCESSING

Volume I: Fundamentals and Algorithms

J. Sanz Subirana, J.M. Juan Zornoza and M. Hernández-Pajares

European Space Agency

Acknowledgements

We would like to thank Dr Javier Ventura Traveset and Professor Dr Günter Hein for suggesting that we write this book and for their enthusiastic support and advice throughout the whole process. We also thank Carlos López de Echazarreta Martín for his help during the editing. We are grateful to Professor Dr Bernhard Hofmann-Wellenhof for his advice and encouragement. Special thanks go to Jordi Català Domínguez for redrawing most of the figures in this book.

Our thanks also go to Adrià Rovira García for his continuous help and the deep implications of his review of this material, to Dagoberto Salazar for initial reviews of the English and to Georgina Sanz for her help and support with the editing of this book. We are also grateful to ESA Communications for their supporting advice and cooperation.

Any comments regarding this book may be addressed to *jaume.sanz@upc.edu*.

Cover: Galileo is the European GNSS (ESA/J. Huart).

To Georgina Sanz Pons
This book is for you

An ESA Communications Production

Publication *GNSS Data Processing, Vol. I: Fundamentals and Algorithms*
(ESA TM-23/1, May 2013)

Production Editor K. Fletcher

Editing Contactivity bv, Leiden, the Netherlands

Publisher ESA Communications
ESTEC, PO Box 299, 2200 AG Noordwijk, the Netherlands
Tel: +31 71 565 3408 Fax: +31 71 565 5433 www.esa.int

ISBN 978-92-9221-886-7 (two volumes plus CD)

ISSN 1013-7076

Preface

This book is inspired by the welcome package that we offer to our doctoral students when they start their activities in the research group. The original package has been updated and compiled into two volumes which contain a self-learning course and software tools aimed at providing the necessary background to start work in an operative way in Global Navigation Satellite System (GNSS). The design and contents are focused on the instrumental use of the concepts and techniques involved in GNSS navigation and it is intended to include all the elements needed to understand how the system works and how to work with it. In this way, after working through the two volumes, the students should be able to develop their own tools for high-accuracy navigation, implementing the algorithms and expanding the skills learned.

The first volume is devoted to the theory, providing a summary of the GNSSs (GPS, Glonass, Galileo and Beidou) fundamentals and algorithms. The second volume is devoted to laboratory exercises, with a wide range of selected practical examples going further into the theoretical concepts and their practical implementation. The exercises have been developed with a specialised software package provided on a CD-ROM together with a set of selected data files for the laboratory sessions. This is an end-to-end GNSS course addressed to all those professionals and students who wish to undertake a deeper study of satellite navigation, targeting the GNSS data processing and analysis issues.

Starting from a review of the GNSS architecture, the contents of Volume I range from the analysis of basic observables (code pseudorange and carrier phase measurements) to setting up and solving the navigation equations for Standard Point Positioning (SPP) and Precise Point Positioning (PPP). It involves, in particular, an accurate modelling of GNSS measurements (up to the centimetre level of accuracy or better) as well as the required mathematical background to achieve the high-accuracy positioning goal. Parameter estimation techniques such as least squares and Kalman filtering are explained from a conceptual point of view, looking towards implementation at an algorithmic level.

For this self-contained educational package, we have tried not only to explain the theoretical concepts and provide the software tools, but also to illustrate the results and give the methodology on GNSS data processing. This is achieved in Volume II through guided examples developed in laboratory sessions using actual GNSS data files in standard formats. Moreover, some additional practical information on public domain servers with GNSS data products (precise orbits and clocks, ionospheric corrections, etc.) is included, among other items.

Most of the algorithms introduced in the theory are implemented in the GNSS-LABoratory (`gLAB`) tool suite. `gLAB` is an interactive and user-friendly educational multipurpose software package for processing and analysing GNSS data to centimetre-level positioning accuracy. The use and functionalities of this tool are thoroughly explained and illustrated through the different guided exercises in the laboratory sessions, together with self-explanatory templates, tool tips and warning messages included in the Graphical User Interface (GUI). `gLAB` has been developed under an European Space Agency (ESA) Education Office contract.

The `gLAB` tool suite is complemented with an additional software package of simple routines implementing different algorithms described in the theory.[1] These elementary routines are included as examples of basic implementations to help students build their own tools. As mentioned above, the target is to provide *effectiveness* in *instrumental use* of the concepts and techniques of GNSS data processing from scratch.

The didactic outline of this book is the result of more than 25 years of university teaching experience. In a similar way, the scientific/technological approach has been enhanced by our experience in developing different R&D projects in the GNSS field.

[1] The `gLAB` source code is also provided as an example of the full implementation of the algorithms in a self-contained tool with a wide range of capabilities.

Contents

How to use this book

This volume contains the fundamentals and algorithms of the GNSS course developed throughout this book. The practical aspects are given in Volume II.

An intuitive approach to GNSS positioning is presented in the first chapter as an introduction to GNSSs. The basic concepts and the geometrical problem are introduced with a simple 2D positioning example. Then, the results are extended to 3D positioning and determination of the time.

The second chapter presents an overview of the GNSSs (i.e. GPS, Glonass, Galileo and Beidou). The GNSS segments are presented in parallel for the four systems emphasising their similarities and pointing out their differences. The signals are explained in independent sections for each system.

The third chapter is devoted to GNSS time reference, coordinate frames and satellite orbits. Only a brief description of the main concepts is presented here, a more detailed explanation being left to Appendix A. Minimum background information on satellite orbits is given to introduce computation of the satellite coordinates. Detailed algorithms to compute GNSS satellite coordinates from broadcast navigation data are given according to the corresponding Interface Control Documents (ICDs). GPS, Galileo or Beidou coordinates are computed from the pseudo-Keplerian elements. Glonass coordinates are computed by integrating the orbit from the initial conditions of position and velocity using a fourth-order Runge–Kutta method. The algorithms for computing coordinates from both almanac data and precise orbit products are also given.

The fourth chapter deals with GNSS measurements and data preprocessing. The code and carrier measurements are introduced and the different combinations of pairs of signals are presented. The interfrequency code biases are introduced after redefining the clocks relative to the code ionosphere-free combination. From these results, the equations describing the measurement content for the different combinations of signals are arranged in a compact and suitable way. After discussing measurement noise and multipath, the chapter concentrates on detection of the carrier phase cycle slip. A broad discussion of this topic is presented and different 'home-made' algorithms are provided as examples of detectors for single-, double- and triple-frequency signals.

The fifth chapter provides the basis for modelling the GNSS measurements. It presents equations and algorithms for an end-to-end model description up to centimetre level or better, to achieve high-accuracy positioning capabilities. Remarks on the model components required for SPP and those additional ones needed for PPP are given along with explanations. As the model terms are common for the four systems (GPS, Glonass, Galileo and Beidou), they are presented in general, although particular specifications for these systems are given when needed. The algorithms and equations are introduced with brief explanations of the physical bases associated with a better understanding of the concepts involved. Different plots of model component assessment with actual data are included to illustrate the effects on the range and positioning domains.

The sixth chapter provides the mathematical background for solving the positioning problem. Code-based positioning is used as the motivation for introducing the linear observation model and the parameter adjustment techniques. A brief review of least and weighted least squares and the minimum variance estimator is given to introduce the Kalman filter. Then the concepts of formal, predicted (Dilution of Precision (DOP)) and measured accuracy are discussed. The second part of this chapter focuses on high-accuracy positioning with code and carrier measurements. The linear observation model is posed for PPP and the Kalman filter configuration and associated parameters are discussed. The chapter ends with a section on carrier phase ambiguity fixing for two- and three-frequency signals. The main concepts and associated equations are given briefly together with a discussion of the related problems. Most of the techniques outlined in this section are used for differential positioning where the ambiguities are fixed in double differences between satellites and receivers. Nevertheless, recent approaches based on undifferenced ambiguity fixing are also discussed.

Six appendices are also included at the end of the book, complementing the contents of Chapter 3 in Appendices A and B, Chapter 4 in Appendix C, Chapter 5 in Appendix F and Chapter 6 in Appendices D and E. A list of acronyms has also been compiled.

1. Introduction

A GNSS involves a constellation of satellites orbiting Earth, continuously transmitting signals that enable users to determine their three-dimensional (3D) position with global coverage.

For many years, the only fully operational GNSS system was the US Global Positioning System (GPS). The Russian GLObal NAvigation Satellite System (Glonass) was restored to full operation in December 2011. The Chinese BeiDou and European Galileo systems are currently under development, although Beidou started an initial operating service (Phase II) in late December 2011.

The positioning principle is based on solving an elemental geometric problem, involving the distances (ranges) of a user to a set of at least four GNSS satellites with known coordinates. These ranges and satellite coordinates are determined by the user's receiver using signals and navigation data transmitted by the satellites; the resulting user coordinates can be computed to an accuracy of several metres. However, centimetre-level positioning can be achieved using more advanced techniques.

1.1 An Intuitive Approach to GNSS Positioning

The basic observable in a GNSS is the time required for a signal to travel from the satellite (transmitter) to the receiver. This travel time, multiplied by the speed of light, provides a measure of the apparent distance (pseudorange) between them.

The following example (inspired by [Kaplan, 1996]) summarises, for a two-dimensional (2D) case, the basic ideas involved in GNSS positioning.

Suppose that a lighthouse is emitting acoustic signals at regular intervals of 10 minutes and intense enough to be heard some kilometres away. Also assume that a ship with a clock perfectly synchronised to the one in the lighthouse is receiving these signals in a time that is not an exact multiple of 10 minutes, for example 20 seconds later ($t = n \times 10^{\mathrm{m}} + 20^{\mathrm{s}}$).

The 20 seconds will correspond to the propagation time of sound from the lighthouse (transmitter) to the ship (receiver). The distance ρ between them can be obtained by multiplying this value by the speed of sound, $v \simeq 340\,\mathrm{m/s}$. That is, $\rho = 20\,\mathrm{s} \times 340\,\mathrm{m/s} = 6.8\,\mathrm{km}$.

Obviously, with a single lighthouse it is only possible to determine a single measure of distance. So, the ship could be at any point on a circle of radius ρ, see Fig. 1.1.

With a second lighthouse, the ship's position will be given by the inter-

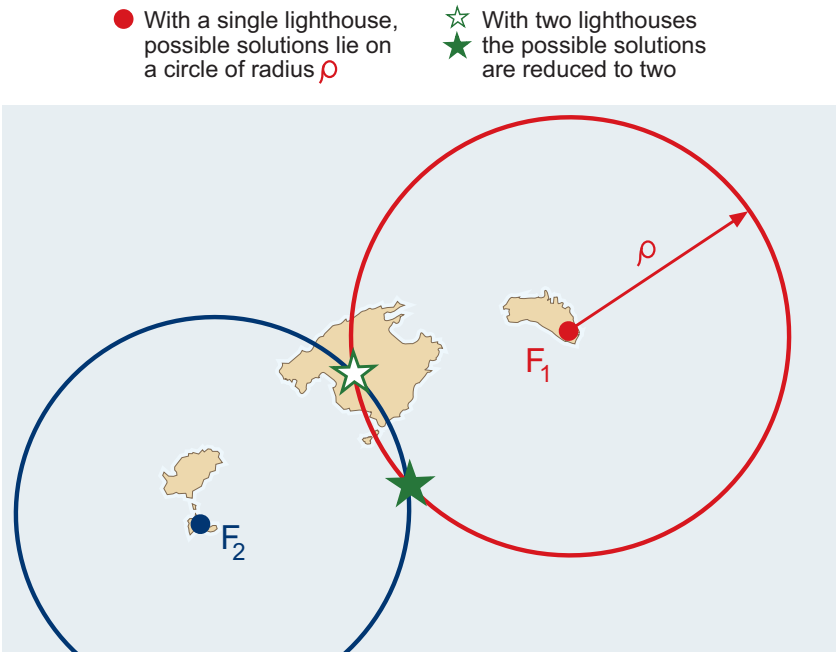


Figure 1.1: In 2D positioning, with a single lighthouse there is a circle of possible locations of the ship. With two lighthouses, the possible solutions are reduced to two. In the figure one of them can be ruled out because it falls on an island.

section of the two circles centred on the two lighthouses and by the radius determined by their distances to the ship (measured using acoustic signals). In this case, the ship could be situated at either of the two points of intersection shown in Fig. 1.1. A third lighthouse will resolve this ambiguity. Nevertheless, a rough knowledge of the ship's position may allow us to proceed without the third lighthouse. For instance, in Fig. 1.1, one of the solutions falls on an island.

1.1.1 A Deeper Analysis of 2D Pseudorange-Based Positioning

So far, perfect synchronism between the lighthouse and ship clocks has been assumed, but in fact this is very difficult to achieve. Note that a synchronisation error between these clocks will produce an erroneous measure of signal propagation time (because it is linked to such clocks) and, as a consequence, an error in the range measurements.

Assume that the ship's clock is biased by an offset $d\tau$ from the lighthouse clocks (which are supposed to be fully synchronised). Then, the measured ranges, R_1 and R_2, will be shifted by an amount $dr = v\,d\tau$:

$$R_1 = \rho_1 + dr, \quad R_2 = \rho_2 + dr \tag{1.1}$$

That is, the radius of the circles in Fig. 1.1 will vary by an unknown amount dr, see Fig. 1.2. Hereafter we will call R_i a *pseudorange*, because it contains an unknown error dr.

At first glance, it might seem that the intersection of these circles (with an undefined radius R_i) could reach any point on the plane (for an arbitrary $d\tau$ value). However, they will only intersect on the branches of a hyperbola, whose foci are located at the two lighthouses, see Fig. 1.2. Indeed, as the clock offset cancels when differencing the pseudoranges, the possible ship

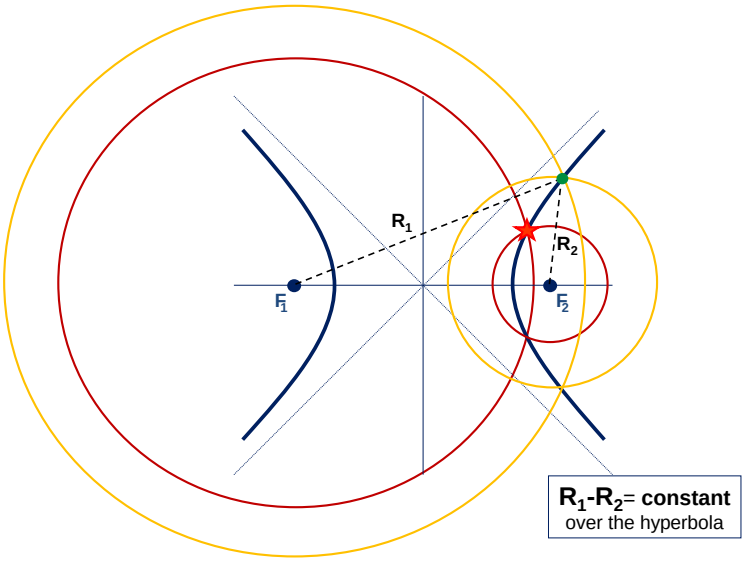


Figure 1.2: An (unknown) offset $d\tau$ in the ship's clock produces a shift in the measured ranges ($R_i = \rho_i + dr$) (or pseudoranges), varying the radius of the circle. But as both pseudoranges R_1 and R_2 have been measured with the same clock, this offset cancels the difference of ranges $R_1 - R_2 = \rho_1 - \rho_2 =$ constant. Therefore, the ship is located at a branch of the hyperbola $R_1 - R_2 =$ constant.

locations must verify the following equation (which defines a hyperbola, see Fig. 1.2):

$$R_1 - R_2 = \rho_1 - \rho_2 = \text{constant} \tag{1.2}$$

A third lighthouse will reduce the uncertainty in the ship's position to just two possible solutions. Such solutions are given by the intersection of two hyperbolas as illustrated in Fig. 1.3. Note that, after estimating the ship's coordinates, its clock offset can be found from equation (1.1).

To complete this analysis, Fig. 1.4 shows another geometric construction (inspired by [Hofmann-Wellenhof et al., 2003]) where the solution is at the centre of a circle with a radius equal to the clock offset $dr = v\, d\tau$, and which is tangent to the three circles of radii ρ_i and centred on the lighthouses.

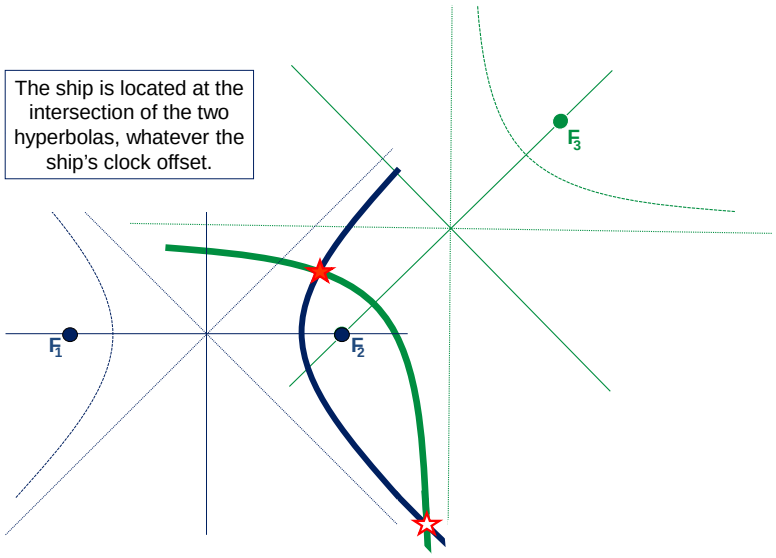


Figure 1.3: The 2D positioning with an unknown user clock offset.

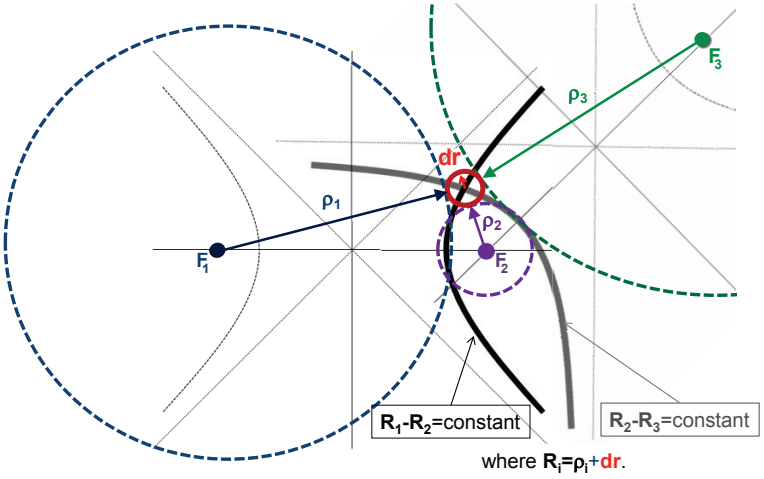


Figure 1.4: Geometrical view 2D positioning, complementing Fig. 1.3.

Finally, and in order to simplify the explanation, consider again the situation in Fig. 1.1 where the ship and lighthouse clocks are assumed fully synchronised. If the range measurements were perfect, a sailor could determine his position as the intersection point of the two circles centred on lighthouses F1 and F2. However, the measurements are not exact, and have some measurement error ε. Figures 1.5 and 1.6 illustrate how this measurement error is translated to the coordinate estimate as an uncertainty region, which depends on the geometry defined by the relative positions of the ship and lighthouses.

1.1.2 Translation to 3D GNSS Positioning

Although the above example corresponds to a 2D case, the basic principle is the same in GNSS:

- **Satellites** (as the lighthouses): In the case of the lighthouses, one assumes that their coordinates are known. In the case of the GNSS satellites, the coordinates are calculated from the navigation data (ephemeris) transmitted by the satellites, see Chapters 2 and 3.
- **Pseudorange** measurements: In GNSS positioning, as in the example, the distances between the receiver and satellites are measured from the travel time of a signal (in GNSS, an electromagnetic wave) from the satellite to the receiver, see Chapters 2 and 4.

Other comments are as follows:

- **Clock synchronisation**: The satellite clocks are among the most critical components of a GNSS. In order to ensure the stability of such clocks, GNSS satellites are equipped with atomic oscillators with high daily stabilities $\Delta f/f \simeq 10^{-13}$ to 10^{-14}. However, despite this high stability, satellite clocks accumulate some offsets over time. These satellite clock offsets are continuously estimated by the ground segment and transmitted to users to correct the measurements[1] (see

[1]Perfect synchronism was assumed between the lighthouse clocks in the previous example.

Chapter 4). The receivers, on the other hand, are equipped with quartz-based clocks, which are much more cheaper but with a poorer stability (about 10^{-9}). This inconvenience is overcome by estimating the clock offset together with the receiver coordinates, as in the previous example.

- **From 2D to 3D positioning**: It is not difficult to extend the previous 2D geometric construction to the 3D case of GNSS positioning, and to show that at least four satellites are needed to compute the three receiver coordinates and clock. In this case, the previous circles and hyperbolas are generalised to spheres and hyperboloids, which intersect in two possible ways. For a ground receiver, one of the solutions is on Earth's surface and the other far away in space. An algebraic method to compute these two solutions (Bancroft's method) is provided in Appendix D. Nevertheless, the usual way to solve this nonlinear problem is to linearise the equations for an approximate user position and solve them iteratively (see Chapter 6).

- **Dilution Of Precision**: The geometry of the satellites (i.e. how the user sees them) affects the positioning error. This is illustrated in Fig. 1.6, where the size and shape of the region change depending on their relative positions. This effect is called Dilution Of Precision (DOP) and is treated in section 6.1.3.2.

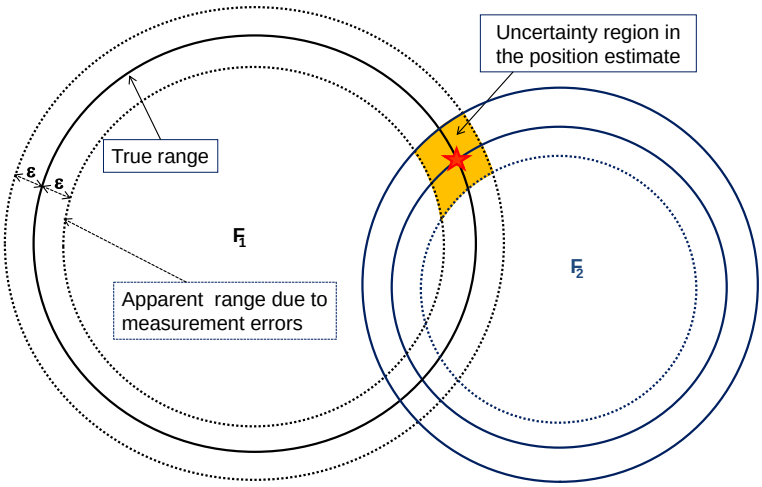


Figure 1.5: The measurement noise ε is translated to the position estimate as an uncertainty region.

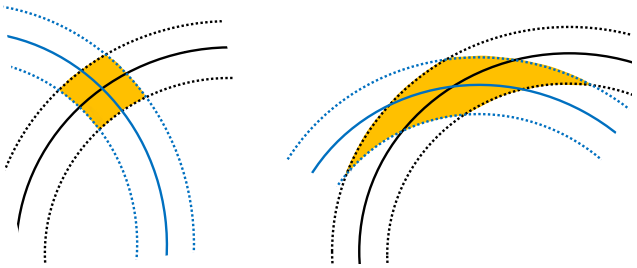


Figure 1.6: The DOP effect in positioning: 2D illustration of the variation of the uncertainty region with geometry.

2. GNSS Architecture

Global Navigation Satellite System (GNSS) is a generic term denoting a satellite navigation system (e.g. GPS, Glonass, Galileo and Beidou) that provides continuous positioning over the globe.[1]

A GNSS basically consists of three main segments: the space segment, which comprises the satellites; the control segment (also referred to as the ground segment), which is responsible for the proper operation of the system; and the user segment, which includes the GNSS receivers providing positioning, velocity and precise timing to users (see Fig. 2.1).

2.1 GNSS Segments

2.1.1 Space Segment

The main functions of the space segment are to generate and transmit code and carrier phase signals, and to store and broadcast the navigation message uploaded by the control segment. These transmissions are controlled by highly stable atomic clocks onboard the satellites.

The GNSS space segments are formed by satellite constellations with enough satellites to ensure that users will have at least four satellites in view simultaneously from any point on Earth's surface at any time.

2.1.1.1 Satellite Constellations

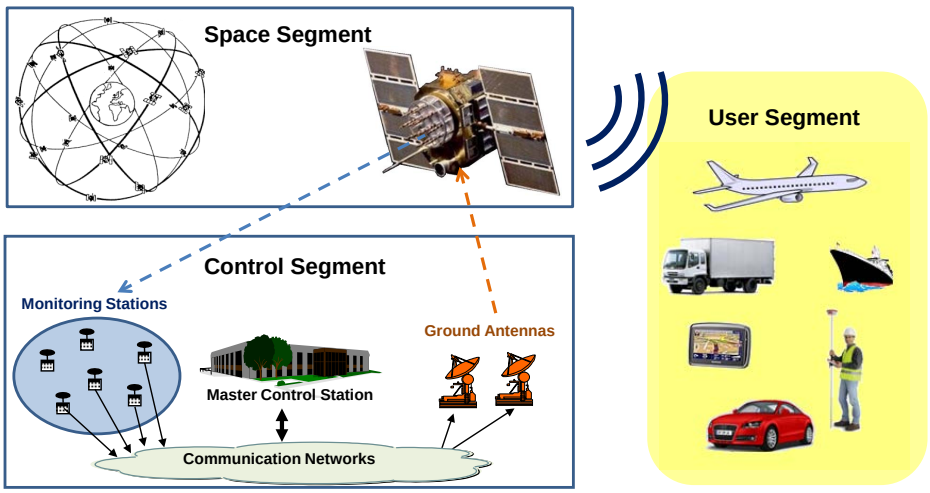
The GPS satellites are arranged in six equally spaced orbital planes surrounding Earth, each with four 'slots' occupied by baseline satellites. This 24-slot arrangement ensures there are at least four satellites in view from virtually any point on the planet.[2] The satellites are placed in a Medium Earth Orbit (MEO) orbit, at an altitude of 20 200 km and an inclination of 55° relative to the equator. Orbits are nearly circular, with an eccentricity of less than 0.02, a semi-major axis of 26 560 km and a nominal period of 11 hours, 58 minutes and 2 seconds (12 sidereal[3] hours), repeating the geometry each sidereal day.

[1] This chapter was written in collaboration with Carlos López de Echazarreta Martín, from ESA/ESAC Communication & Education Office.

[2] The US Air Force normally flies more than 24 GPS satellites to maintain coverage whenever the baseline satellites are being serviced or decommissioned. The extra satellites may increase GPS performance but are not considered part of the core constellation (see *http://www.gps.gov/systems/gps/space/* and *http://www.navcen.uscg.gov* for the status of the GPS constellation).

[3] See section A.1.1.2 in Appendix A.

Figure 2.1: GNSS architecture.



The nominal Glonass constellation consists of 24 MEO satellites deployed in three orbital planes with eight satellites equally spaced in each plane. The orbits are roughly circular, with an inclination of about 64.8°, and at an altitude of 19 100 km with a nominal period of 11 hours, 15 minutes and 44 seconds, repeating the geometry every eight sidereal days.[4] Due to funding problems, the number of satellites decreased from the 24 available in 1996 to only 6 in 2001. In August 2001, the Russian government committed to recover the constellation and to modernise the system, approving new funding. A total of 24 operational satellites plus 2 in maintenance were again available in December 2011, restoring the full constellation.[5]

The planned Galileo constellation in Full Operational Capability (FOC) phase consists of 27 operational and 3 spare MEO satellites at an altitude of 23 222 km and with an orbit eccentricity of 0.002. Ten satellites will occupy each of three orbital planes inclined at an angle of 56° with respect to the equator. The satellites will be spread around each plane and will take about 14 hours, 4 minutes and 45 seconds to orbit Earth, repeating the geometry each 17 revolutions, which involves 10 sidereal days. This constellation guarantees, under nominal operation, a minimum of six satellites in view from any point on Earth's surface at any time, with an elevation above the horizon of more than 10°.

Comment: The Galileo Deployment Plan has two main phases: (1) the In-Orbit Validation (IOV) phase with a reduced constellation of four operational satellites and their related ground infrastructure (2012); and (2) the FOC that involves the deployment of the remaining ground and space infrastructure, including an intermediate initial operational capability phase (by 2014–2016) with 18 satellites in operation (the 4 IOV satellites plus 14 others). Completion of the FOC phase is expected by 2019–2020.

[4] However, as each orbital plane contains eight satellites, a satellite will pass the same place every sidereal day.

[5] See *http://www.glonass-center.ru/en/* for the status of the Glonass constellation.

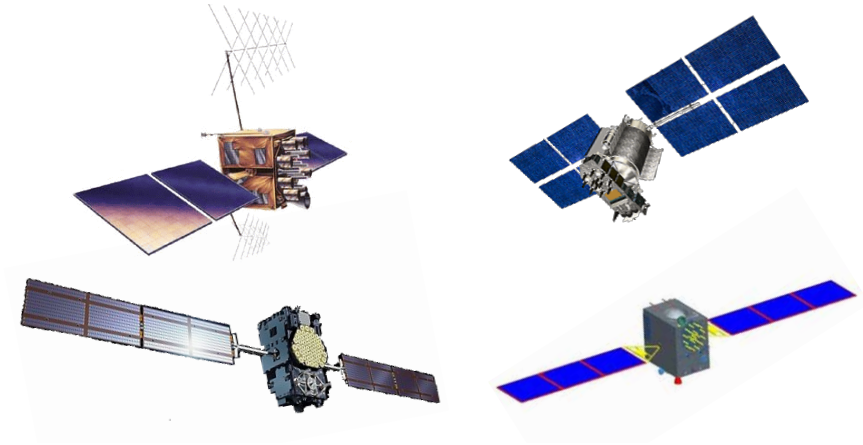


Figure 2.2: GNSS satellites: GPS IIR-M (*top left*), Glonass-M (*top right*), Galileo IOV (*bottom left*) and Beidou-M (*bottom right*).

The Beidou (Compass)[6] constellation (Phase III)[7] will consist of 35 satellites, including 5 Geostationary Orbit (GEO) satellites and 30 non-GEO satellites in a nearly circular orbit. The non-GEO satellites include 3 Inclined Geosynchronous Satellite Orbit (IGSO) ones, with an inclination of about 55°, and 27 MEO satellites orbiting at an altitude of 21 528 km in three orbital planes with an inclination of about 55° and with an orbital period of about 12 hours and 53 minutes, repeating the ground track every seven sidereal days. The GEO satellites, orbiting at an altitude of about 35 786 km, are positioned at 58.75°E, 80°E, 110.5°E, 140°E and 160°E, respectively, and are expected to provide global navigation service by 2020. The previous Phase II involves a reduced constellation of four MEO, five GEO and five IGSO satellites to provide regional coverage of China and surrounding areas. The initial Phase II operating service with 10 satellites started on 27 December 2011.

2.1.1.2 The Satellites

Satellites have various structures and mechanisms to keep them in orbit, communicate with the control segment and broadcast signals to receivers. The satellite clocks are one of the critical components of GNSSs. For this reason, satellites are equipped with very high-stability atomic clocks (rubidium, caesium, hydrogen).

A specific description of satellites for the GPS, Glonass, Galileo and Beidou systems is presented below (see also Fig. 2.2).[8]

[6] *BeiDou* is the Chinese translation of 'Big Dipper' and also means 'Compass' in a metaphoric sense.

[7] The first step taken by China in building a satellite navigation system under its own control was the regional system BeiDou-1, which became operational in May 2003. But BeiDou-2/3, i.e. Phase II or III (also known as Compass), is not an extension of BeiDou-1, whose location measurement scheme requires two-way transmissions between a user and the central control station via the satellite. It is a new GNSS like GPS, Glonass or Galileo. Nevertheless, backward compatibility with BeiDou-1 is assured by the inclusion of five GEO satellites.

[8] Satellite pictures courtesy of US Air Force *http://www.af.mil* (GPS), JSC 'Information Satellite System Reshetnev Company' *http://www.iss-reshetnev.com* (Glonass) and ESA *http://www.esa.int* (Galileo). Beidou-M is from [Cao et al., 2008].

2.1.1.2.1 GPS Satellites

GPS satellites are divided into blocks. Each block comprises a set of satellites usually launched within a certain time interval. A brief description of the different blocks follows.[9]

Block I, Navigation Development Satellites. Eleven satellites of this kind were launched between 1978 and 1985. The Selective Availability (S/A) capability[10] was not yet implemented then. They weighed about 845 kg and had a planned average lifetime of 4.5 years, although some of them lasted for 10 years. They were capable of providing positioning services for 3 or 4 days without any contact with the control centre.

Blocks II and IIA, Operational Satellites. These consist of 28 satellites in total, launched from 1989 on, and many still operating. They weigh about 1 500 kg each and have a planned average lifetime of 7.5 years. Since 1990, an improved version has been used, namely Block IIA (advanced), with the capability of mutual communication. They are able to supply positioning services for 180 days without contact with the control segment. However, under normal operating mode, they must communicate daily.

Block IIR, Replacement Operational Satellites. These satellites were produced to replace the II/IIA series as they gradually degraded or exceeded their intended design lifetime. The 'R' in Block IIR stands for replenishment. They weigh about 2 000 kg each and have a planned average lifespan of 10 years. These satellites are capable of autonomously determining their orbits and generating their own navigation messages. They are able to measure the distances between them and to transmit observations to other satellites or to the control segment. A satellite of this category, when completely developed, must be capable of operating for about half a year without any support from the control segment and with no degradation in ephemeris accuracy.

Block IIR-M, Modernised Satellites. The IIR-M series of satellites are upgraded versions of the IIR series, completing the backbone of the GPS constellation. They include a new military signal and the more robust civil signal L2C. The first Block IIR-M satellite was launched on 26 September 2005.

Block IIF, Follow-on Operational Satellites. The first satellite (SVN62) was launched on 28 May 2010.[11] The IIF series expand on the capabilities of the IIR-M series with the addition of a third civil signal (L5) in a frequency band protected for safety-of-life applications. Their theoretical average lifetime is about 15 years, and they have inertial navigation systems.

Block III Satellites. These satellites will introduce significant enhancements in navigation capabilities, by improving interoperability and jam resistance. They will provide the fourth civil signal on the L1 band (L1C).

[9] More information can be found at *http://tycho.usno.navy.mil/gps.html*.

[10] S/A is an intentional degradation of public GPS signals implemented for US national security reasons. S/A was switched off in May 2000 and permanently removed in 2008.

[11] See *http://gge.unb.ca/Resources/GPSConstellationStatus.txt*.

GPS satellites are identified in different ways: by their position in the orbital plane,[12] their NASA reference number, their assigned Pseudo-Random Noise (PRN) code, their Space Vehicle Number (SVN), etc.

Full Operational Capability (FOC):
The GPS constellation achieved FOC in March 1994, when 24 Block II/IIA satellites were operational in their assigned orbits. Nevertheless, FOC was not declared until 17 July 1995 [Hofmann-Wellenhof et al., 2008].

2.1.1.2.2 Glonass Satellites

The following generations of satellites have been developed.[13]

Prototypes (Generation Zero). The first prototypes of the Glonass (Uragan) satellites were placed into orbit in October 1982, with 18 launched between 1982 and 1985. These prototype satellites are referred to as Block I vehicles.

First Generation. The first true Glonass satellites were launched between 1985 and 1990. They are divided into different block vehicles (Block IIa, IIb and IIv), with different design lifetimes between the blocks. These lifetimes ranged from the two-year design of Block IIb to the three-year design of Block IIv, although many spacecraft exceeded this (up to 4.5 years). These satellites are all three-axis stabilised with a mass of about 1 250 kg, and are equipped with a basic propulsion system to allow relocation within the constellation. They have improved time and frequency standards compared with the previous prototypes, and increased frequency stability.

Second Generation. Glonass-M (or Uragan-M) is the second generation of satellites, where the '-M' indicates modernised or modified. They were developed from 1990 on, with the first one sent into orbit in 2001.[14] Glonass-M satellites have a longer design lifetime of seven years as a result of improvements to the propulsion system and clock stability (caesium clocks). This is a large increase in lifetime compared with the 2–3 years of previous first-generation spacecraft, but it is still below the 10-year mean lifetime of the NAVSTAR-GPS.

These satellites have a mass of about 1 480 kg. They are 2.4 m in diameter and 3.7 high, with dual solar arrays of 7.2 m. They also carry corner-cube laser reflectors for precise orbit determination and geodetic research. A remarkable feature of these satellites is the addition of a second civil signal on the G2 band, which allows civil users to cancel out ionospheric refraction.

Third Generation. These satellites are known as the Glonass-K (or Uragan-K) spacecraft. They have an increased lifetime design of 10–12 years and a reduced weight of only 750 kg (allowing their launch in pairs from the Plesetsk Cosmodrome on Soyuz-2-1a rockets). As with the previous generation of satellites, they are three-axis stabilised and nadir pointing with dual solar arrays. They provide the new Code Division Multiple

[12] Every satellite has a place number (1, 2, 3, etc.) within each of the six orbital planes – A, B, C, D, E or F.
[13] For the status of the Glonass constellation see *http://www.glonass-ianc.rsa.ru*.
[14] See *http://gge.unb.ca/Resources/GLONASSConstellationStatus.txt*.

Access (CDMA) signal for civilian applications on the new G3 band, among the civil signals in the G1 and G2 bands using the Frequency Division Multiple Access (FDMA) technique (see section 2.2), and also include a Search and Rescue (SAR) payload. The first Glonass-K satellite was launched on 26 February 2011.

Full Operational Capability (FOC). The Glonass constellation reached FOC on 8 December 2011, after the satellite launched on 4 November (into Slot 3) was declared operational at 11:42 Moscow Time. With 24 operating satellites, the Glonass constellation provides complete coverage of Earth's surface.

2.1.1.2.3 Galileo Satellites

Galileo Satellites Experimental Phase

Two experimental satellites were launched between 2005 and 2008: the Galileo In-Orbit Validation (GIOVE) satellites GIOVE-A and GIOVE-B. Their mission served several purposes:[15] to secure and maintain the Galileo frequency filling with the International Telecommunications Union (ITU); to validate technologies to be used in the Galileo operational constellation; to monitor the environment at MEO; and to generate Galileo signals so that user equipment can be developed.

Both satellites were built in parallel to provide in-orbit redundancy and secure the mission objectives, and both provide complementary capabilities.

GIOVE-A was built by Surrey Satellite Technology Ltd. (SSTL) and was successfully launched on 28 December 2005. It has a mass of about 600 kg and two redundant small-size rubidium atomic clocks onboard, with a stability of 10 ns per day. Its nominal lifetime of 27 months was extended in order to maintain continuity of Galileo's in-orbit validation prior to the launch of GIOVE-B and beyond.

GIOVE-B was built by Astrium and Thales Alenia Space, and was successfully launched on 27 April 2008. Its mass and estimated lifetime are similar to those of GIOVE-A, but it has a more advanced payload. There are four redundant clocks onboard, two small rubidium atomic clocks, with a stability of 10 ns per day, and two passive hydrogen maser clocks, which are a different type of atomic clock with a higher stability (1 ns per day).

Galileo IOV Phase

This phase is aimed at qualifying the Galileo space, ground and user segments through extensive in-orbit/on-ground tests and operations of a core spacecraft constellation and of the associated ground segment. During this phase, four operational satellites complement the two experimental satellites (GIOVE-A and GIOVE-B) already in orbit.

The first two operational Galileo satellites were launched on 21 October 2011 by a Soyuz rocket, and placed in the first orbital plane. The next two

[15] More information is available at *http://www.esa.int/Our_Activities/Navigation/GIOVE-A_navigation_signal_available_to_users*.

satellites were launched on 12 October 2012 and placed in the second orbital plane. They are fully representative of the others that will follow them into orbit. Each satellite has a mass of about 700 kg and combines two rubidium and two passive hydrogen maser clocks with a powerful transmitter to broadcast precise navigation data.

Galileo FOC Phase

By 2014-2016, the constellation is expected to have grown to 18 satellites, including the 4 satellites from the IOV phase, and it is hoped that FOC will be reached by 2019-2020, with a constellation of 27 operational satellites, plus 3 spares.

2.1.1.2.4 Beidou Satellites

There are three variants of Beidou satellites: the geostationary Beidou-G, the geosynchronous Beidou-IGSO and the MEO Beidou-M. They have a design lifetime of eight years and are based on the three-axis-stabilised DFH-3 (DongFangHong) platform,[16] built by CASC (China Aerospace Science and Technology Corporation).

The first Beidou Phase II satellite in orbit was the experimental Beidou-M satellite, launched on 14 April 2007 (local time) to test the onboard payload. This satellite represented the first step in the new Chinese GNSS. By late December 2011 China had put four Beidou-G and five Beidou-IGSO satellites into orbit.

On 27 December 2011, with 10 orbiting satellites, the Beidou system formally started to provide an initial Phase II operating service to China and its surrounding areas (regional service). By the end of 2012, there are five GEO, four MEO and five IGSO navigation satellites in orbit. Completion of the 35-satellite constellation (5 GEO, 27 MEO and 3 IGSO) is scheduled for 2020 (Phase III, global service) [BeiDou-SIS-ICD, 2012].

2.1.2 Control Segment

The control segment (also referred to as the ground segment) is responsible for the proper operation of the GNSS. Its basic functions are:

- to control and maintain the status and configuration of the satellite constellation;
- to predict ephemeris and satellite clock evolution;
- to keep the corresponding GNSS time scale (through atomic clocks); and
- to update the navigation messages for all the satellites.

[16] See *http://space.skyrocket.de/doc_sdat/dfh-3.htm*.

2.1.2.1 GPS Control Segment

The *GPS control segment* is composed of a network of Monitoring Stations (MS), a Master Control Station (MCS) and the Ground Antennas (GA) [ICD-GPS-200]. See the layout in Fig. 2.1.

The *Master Control Station*, located in Colorado Springs, USA, is the core of the control segment. It is responsible for operating the system and providing command, control and maintenance services to the space segment. A new, fully functional backup station, known as the Alternate Master Control Station (AMCS), was set up as part of a modernisation plan[17] at Vandenberg Air Force Base, see Fig. 2.3.

The *Monitoring Stations* are distributed around the world, see Fig. 2.3. They are equipped with atomic clock standards and GPS receivers to collect GPS data continuously for all the satellites in view from their locations. The collected data are sent to the *Master Control Station* where they are processed to estimate satellite orbits (ephemerides) and clock errors, among other parameters, and to generate the navigation message.

Prior to the modernisation programme, the MS network comprised five sites from the United States Air Force (USAF), located in Hawaii, Colorado Springs (Colorado), Ascension Island (South Atlantic), Diego Garcia (Indian Ocean) and Kwajalein (North Pacific). Cape Canaveral (Florida) was incorporated in 2001, also from the USAF.

In 2005, the modernisation programme expanded this network to include six MS operated by the National Geospatial-Intelligence Agency (NGA) of the US Department of Defense (DoD): Adelaide (Australia), Buenos Aires (Argentina), Hermitage (UK), Manama (Bahrain), Quito (Ecuador) and Washington, DC (USA). Five more stations were added afterwards in 2006: Fairbanks (Alaska), Osan (South Korea), Papeete (Tahiti), Pretoria (South Africa) and Wellington (New Zealand). With this configuration, each satellite is seen from at least three MS, which allows more precise orbits and ephemeris data to be computed, therefore improving system accuracy.

The *Ground Antennas* uplink data to the satellites via S-band radio signals. These data include ephemerides and clock correction information transmitted within the navigation message, as well as command telemetry from the MCS. Every satellite can be 'uploaded' three times per day, that is every eight hours; nevertheless, it is usually updated just once a day. The GA are collocated in four of the MS (Ascension Island, Cape Canaveral, Diego Garcia and Kwajalein).

[17] See *http://www.gps.gov/systems/gps/control/* for further information on the GPS Control Segment Modernisation programme.

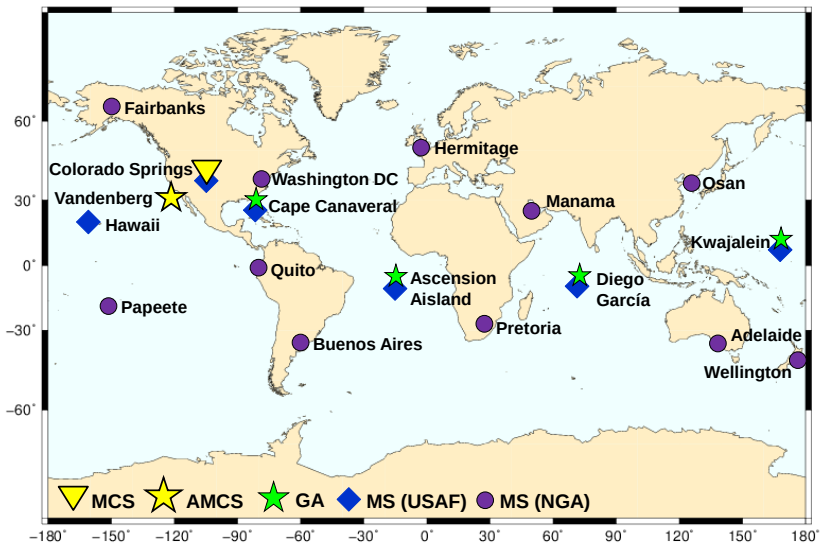


Figure 2.3: GPS ground infrastructure: Master Control Station (MCS), Alternate Master Control Station (AMCS), Monitoring Stations from United States Air Force [MS(USAF)], Monitoring Stations from National Geospatial-Intelligence Agency [MS(NGA)] and Ground Antennas (GA).

2.1.2.2 Glonass Control Segment

As in GPS, the Glonass control segment (or Ground-based Control Facilities) monitors the status of satellites, determines the ephemerides and satellite clock offsets with respect to Glonass time and Coordinated Universal Time (UTC), and, twice a day, uploads the navigation data to the satellites.

The Glonass control segment [GLONASS ICD, 2008] comprises the System Control Centre (SCC), located in Krasnoznamensk (Moscow region), the Network of Command and Tracking Stations, entirely located within former Soviet Union territory, and the Central Synchroniser (CC-M) – the system clock – in Schelkovo (Moscow region).

The SCC coordinates all functions and operations at the system level. It processes information from the command and tracking stations to determine the satellite clock and orbit states, and it updates the navigation message for each satellite.

The command and tracking stations comprise a main network of five Telemetry, Tracking and Control (TT&C) stations distributed throughout the Russian territory (in St Petersburg, Schelkovo, Yenisseisk, Komsomolsk and Ussuriysk), some of them equipped with laser ranging (Laser Station (LS)) and other monitoring facilities (MS), see Fig. 2.4. This network is complemented in the modernisation plan by additional MS in the former Soviet Union territory (Zelenchuck, Nurek, Ulan-Ude and Yakutsk).

The Central Synchroniser is responsible for the Glonass time scale. It is connected with the 'phase control system' that monitors satellite clock time and phase signals.

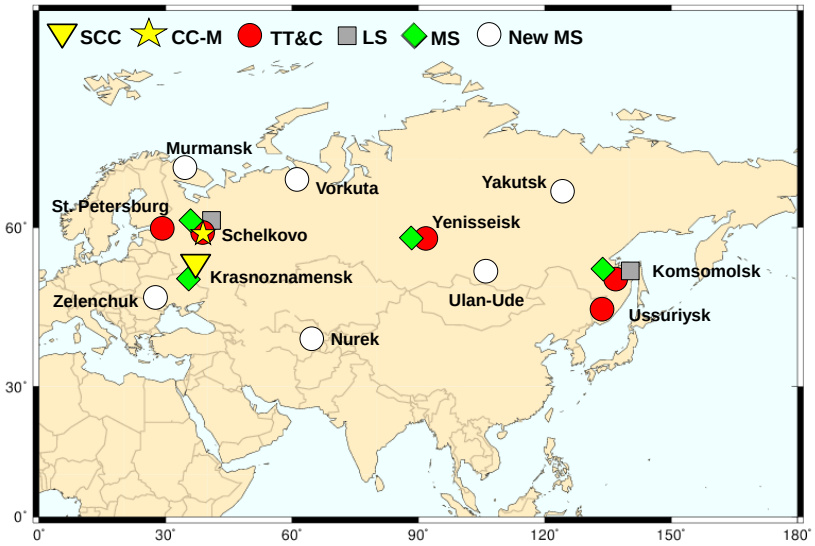


Figure 2.4: Glonass ground infrastructure: System Control Centre (SCC), Central Synchroniser (CC-M), Telemetry, Tracking and Control (TT&C) stations, Laser Station (LS), Monitoring Stations (MS).

2.1.2.3 Galileo Ground Segment

The Galileo ground segment controls the entire satellite constellation, the navigation system facilities and the dissemination services. The FOC involves two Ground Control Centre (GCC), five Telemetry, Tracking and Control (TT&C) stations, nine Mission Uplink Stations (ULS) and a worldwide network of Galileo Sensor Stations (GSS), see Fig. 2.5.

The Ground Control Segment (GCS) is responsible for the constellation control and management of Galileo satellites. It provides the TT&C function for the whole satellite constellation. Its functional elements are deployed within the GCCs and the globally distributed TT&C stations.[18] The TT&C stations use S-band frequency antennas 13 m in diameter to provide a secure exchange of data between the control centres and satellites.

The Ground Mission Segment (GMS) is responsible for the determination and uplink of the navigation and integrity data messages[19] needed to provide the navigation and UTC time transfer service. The GMS includes a worldwide network of GSS, continuously collecting data to be processed by GCC for determining Galileo navigation and integrity data messages. Each of the GSS is equipped with three parallel reception channels: one channel for the determination of orbit data and clock synchronisation, a second channel for integrity determination, and a third redundant channel. The global geographical distribution of such stations has been selected to ensure permanent access to any satellite of the constellation at any time.

Navigation and integrity data are uplinked from the GCC to the satellites by nine ULS. These stations are equipped with C-band parabolic antennas about 3 m in diameter. The uplinks are designed to ensure the

[18] That is, the majority of the functions are contained within the operator-controlled GCC, with only the autonomous TT&C functions located at remote sites.

[19] To warn users whether the transmitted signals can be trusted, and to receive timely alert messages in case of failures.

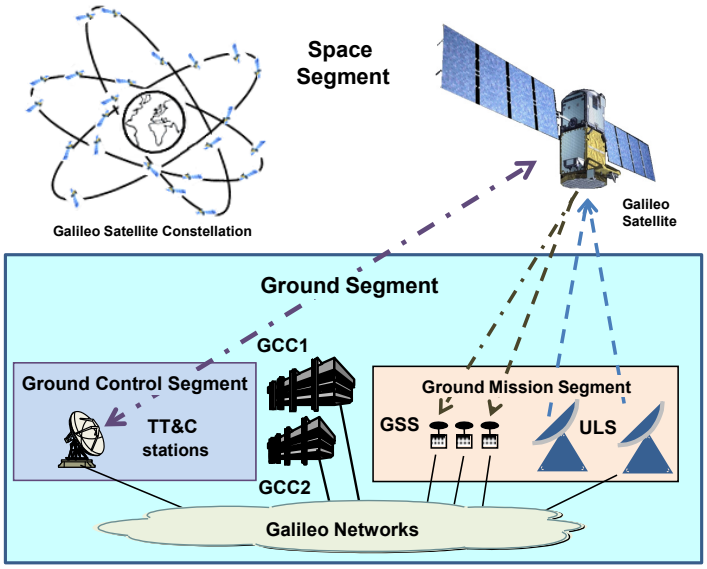


Figure 2.5: Galileo ground segment layout.

provision of navigation and integrity information worldwide for all users, even in the case of a potential failure due to the loss of a single antenna.

A hybrid communication network interconnects the remote stations (ULS, GSS and TT&C stations) with the GCC by different means of standard and special radio, wired data and voice communication links, ensuring communication between all the sites.

The two GCC constitute the core of the ground segment. There are two redundant elements located in Fucino (Italy) and Oberpfaffenhofen (Germany). Some of their main functions are:

- orbit determination and synchronisation;
- control of all Galileo satellites and uploading navigation data messages;
- monitoring and control, performance monitoring, performance prediction and maintenance management functions of the ground segment elements;
- monitoring and control of all ground segment elements in real time;
- generation of navigation messages; and
- computation of Galileo System Time (GST) and provision of a reliable and stable coordinated time reference for the Galileo system.

2.1.2.4 Beidou Control Segment

There is little information on the Beidou ground segment at the time of writing. It is basically composed of 1 MCS, 2 ULS and 30 MS.

2.1.3 User Segment

The user segment is composed of GNSS receivers. Their main function is to receive GNSS signals, determine pseudoranges (and other observables) and solve the navigation equations in order to obtain the coordinates and provide a very accurate time.

The basic elements of a generic GNSS receiver are: an antenna with preamplification, a radio frequency section, a microprocessor, an intermediate-precision oscillator, a feeding source, some memory for data storage and an interface with the user. The calculated position is referred to the antenna phase centre. For more details see [Fantino et al., 2011].

2.2 GNSS Signals

GNSS satellites continuously transmit navigation signals at two or more frequencies in L band. These signals contain ranging codes and navigation data to allow users to compute both the travel time from the satellite to the receiver and the satellite coordinates at any epoch. The main signal components are described as follows:

Carrier: Radio frequency sinusoidal signal at a given frequency.

Ranging code: Sequences of zeros and ones which allow the receiver to determine the travel time of the radio signal from the satellite to the receiver. They are called PRN sequences or PRN codes.

Navigation data: A binary-coded message providing information on the satellite ephemeris (pseudo-Keplerian elements or satellite position and velocity), clock bias parameters, almanac (with a reduced-accuracy ephemeris data set), satellite health status and other complementary information.

Frequency Allocation

The allocation of frequency bands is a complex process because multiple services and users can fall within the same range. That is, the same frequencies can be allocated for different purposes in different countries. The ITU is a United Nations agency coordinating the shared global use of the radio spectrum. It involves, for instance, television, radio, cell (mobile) phone, radar satellite broadcasting, etc., and even microwave ovens. The ITU divides the electromagnetic spectrum into frequency bands, with different radio services assigned to particular bands.

Figure 2.6 shows the frequency bands for the Radionavigation Satellite Service (RNSS). There are two bands in the region allocated to the Aeronautical Radio Navigation Service (ARNS) on a primary basis worldwide. These bands are especially suitable for Safety-of-Life (SoL) applications because no other user of this band is allowed to interfere with the GNSS signals. These correspond to the upper L band (1 559–1 610 MHz), containing the GPS L1, Galileo E1, Glonass G1 and Beidou B1 bands, and to the bottom of the lower L band (1 151–1 214 MHz) where the GPS L5, Glonass G3, Galileo E5 and Beidou B2 bands are located.

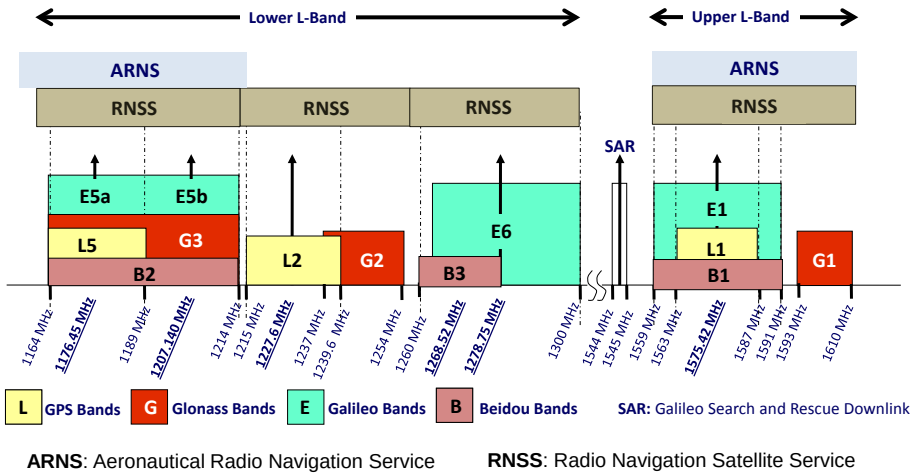


Figure 2.6: GPS, Glonass, Galileo and Beidou navigational frequency bands.

The remaining GPS L2, Glonass G2, Galileo E6 and Beidou B3 signals are in the 1 215.6–1 350 MHz bands. These bands were allocated to radio location services (ground radars) and RNSS on a primary basis, so the signals in these bands are more vulnerable to interference than the previous ones.

2.2.1 GPS Signals

Legacy GPS signals are transmitted on two radio frequencies in the L band, referred to as Link 1 (L1) and Link 2 (L2),[20] or L1 and L2 bands. They are right-hand circularly polarised and their frequencies are derived from a fundamental frequency $f_0 = 10.23\,\text{MHz}$, generated by onboard atomic clocks.

$$\begin{aligned} \text{L1} &= 154 \times 10.23\,\text{MHz} = 1\,575.420\,\text{MHz} \\ \text{L2} &= 120 \times 10.23\,\text{MHz} = 1\,227.600\,\text{MHz} \end{aligned}$$

Two services are available in the current GPS system:

SPS: The Standard Positioning Service is an open service, free of charge for worldwide users. It is a single-frequency service in the frequency band L1.

PPS: The Precise Positioning Service is restricted by cryptographic techniques to military and authorised users. Two navigation signals are provided in two different frequency[21] bands, L1 and L2.

The GPS uses the CDMA technique to send different signals on the same radio frequency, and the modulation method used is Binary Phase Shift Keying (BPSK) (for more details see [Enge and Misra, 1999] or [Avila-Rodriguez, 2008]).

[20] They also transmit two additional signals at frequencies referred to as L3 (associated with the Nuclear Detonations Detection System) and L4 (for other military purposes).

[21] Transmission at two frequencies allows dual-frequency user receivers to cancel out ionospheric refraction, which is one of the main sources of error (see section 5.4.1.1).

The following types of PRN codes and messages are modulated over the two carriers (see Fig. 2.7):

- *Coarse/Acquisition (C/A) code*, also known as civilian code $C(t)$: This sequence contains 1 023 bits and is repeated every millisecond (i.e. a chipping rate of 1.023 Mbps). Then, the duration of each C/A code *chip* is 1 μs, which means a *chip width* or wavelength of 293.1 m. This code is modulated only on L1. The C/A code defines the SPS.

- *Precision code*, $P(t)$: This is reserved for military use and authorised civilian users. The sequence is repeated every 266 days (38 weeks) and a weekly portion of this code is assigned to every satellite, called the PRN sequence. Its chipping rate is 10 Mbps, which leads to a wavelength of 29.31 m. It is modulated over both carriers L1 and L2. This code defines the PPS.

- *Navigation message*, $D(t)$: This is modulated over both carriers at 50 bps, reporting on ephemeris and satellite clock drifts, ionospheric model coefficients and constellation status, among other information.

$$\begin{aligned} s_{L1}(t) &= a_P P_i(t) D_i(t) \sin(\omega_1 t + \phi_{L_1}) + a_C C_i(t) D_i(t) \cos(\omega_1 t + \phi_{L_1}) \\ s_{L2}(t) &= b_P P_i(t) D_i(t) \sin(\omega_2 t + \phi_{L_2}) \end{aligned}$$

  The index i stands for the i-th satellite.

In order to restrict access of civilian users to full system accuracy, the following protection was introduced:

- *S/A or Selective Availability*: This covers intentional satellite clock degradation (process δ) and ephemeris manipulation (process ϵ). The effect on horizontal positioning ranges from about 10 m (S/A off) to 100 m (S/A on) (2σ error). The process δ acts directly on the satellite clock fundamental frequency, which has a direct impact on pseudoranges to be calculated by users' receivers. The process ϵ consists of truncating information related to orbits. US President Bill Clinton ordered the cessation of GPS S/A on 2 May 2000. It was permanently removed in 2008, and not included in the next generations of GPS satellites.

- *A/S or Anti-Spoofing*: This consists of P code encryption by combining it with a secret W code, resulting in the Y code, which is modulated over the two carriers L1 and L2. The purpose is to protect military receivers against an adversary transmitting a faulty copy of the GPS signal to mislead the receiver, and to deny access of non-authorised users to the precise ranging code P at L1 and L2 frequencies (as the only C/A code available on L1).

The signal structure is summarised in Fig. 2.7 and Table 2.1.[22]

[22]Note that Fig. 2.7 and Table 2.1 only show the legacy signals on the L1 and L2 bands.

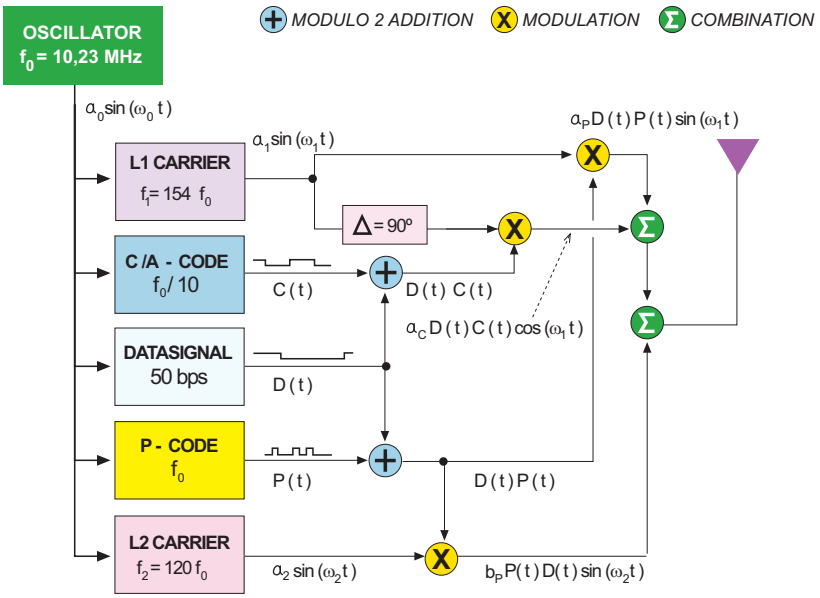


Figure 2.7: Legacy GPS signal structure (source [Seeber, 1993]).

Table 2.1: Legacy GPS signal structure (source [Seeber, 1993]).

Atomic clock frequency	$f_0 = 10.23$ MHz
Frequency L1	$154 \times f_0$
	1575.420 MHz
Wavelength L1	19.03 cm
Frequency L2	$120 \times f_0$
	1227.600 MHz
Wavelength L2	24.42 cm
P code frequency (chipping rate)	$f_0 = 10.23$ MHz (Mbps)
P code wavelength	29.31 m
P code period	266 days, 7 days/satellite
C/A code frequency (chipping rate)	$f_0/10 = 1.023$ MHz
C/A code wavelength	293.1 m
C/A code period	1 ms
Navigation message frequency	50 bps
Frame length	30 s
Total message length	12.5 min

GPS Signal Modernisation: Introduction of New Signals

The GPS signal modernisation includes an additional Link 5 (L5) frequency and several new ranging codes on the different carrier frequencies. They are referred to as the civil signals L2C, L5C and L1C and the military M code. All of them are right-hand circularly polarised.

Modernisation of the GPS system began in 2005 with the launch of the first IIR-M satellite. This satellite supported the new military M signal and the second civil signal L2C. This latter signal is specifically designed to meet commercial needs, allowing the development of low-cost, dual-frequency civil GPS receivers.

The L2C code is composed of two ranging codes multiplexed in time: the L2 Civil Moderate (L2CM) code and the L2 Civil Long (L2CL) code (for more details see [Enge and Misra, 1999]). The L2C code is BPSK modulated onto the L2 carrier frequency and broadcast at a higher effective power level than the original L1 C/A signal. This, together with its powerful cross-correlation properties, facilitates tracking with large signal-level variations from satellite to satellite,[23] making reception easier under trees and even indoors. This signal will also be interoperable with the Chinese Beidou system. In December 2010 there were nine GPS satellites broadcasting L2C (PRN01, 05, 07, 12, 15, 17, 25, 29, 31).[24]

The military M code signals are designed to use the edges of the band with only a minor signal overlap with the pre-existing C/A and P(Y) signals (see Fig. 2.8). This military M code is modulated into L1 and L2 carriers using the Binary Offset Carrier (BOC) scheme (for more details see [Avila-Rodriguez, 2008]). It has been designed for autonomous acquisition, so that a receiver is able to acquire the M code signal without access to C/A or P(Y) code signals.

The GPS modernisation plan continued with the launch of the Block IIF satellites that include, for the first time, the third civil signal on the L5 band (i.e. within the highly protected ARNS band).[25] This new L5C signal has a new modulation type and was designed for users requiring SoL applications. There are two signal components: the in-phase component (L5-I) with data and ranging code, both modulated via BPSK onto the carrier; and the quadrature component (L5-Q), with no data but also having a ranging code BPSK modulated onto the carrier. This signal has an improved code/carrier tracking loop and its high power and signal design provide robustness against interference. Moreover, its higher chipping rate than the C/A code (see Table 2.2) provides superior multipath performance.

The next step involves the Block III satellites, which will provide the fourth civil signal on L1 band (L1C). This signal is designed to enable interoperability between GPS and international satellite navigation systems (such as Galileo).[26] Multiplexed Binary Offset Carrier (MBOC) modulation is used to improve mobile reception in cities and other challenging environments. L1C comprises the L1C-I data channel and L1C-Q pilot channel. The implementation proposed for MBOC is the Time Multiplexed BOC (TMBOC). See [Avila-Rodriguez et al., 2006] and [Avila-Rodriguez, 2008] for more details. This signal will be broadcast at the same frequency as the original L1-C/A signal, which will be retained for backward compatibility.

Figure 2.8 shows the layout of the different GPS signals and ranging codes for the different modernisation phases. Figure 2.15, at the end of this chapter, shows a panel comparing the signals of the different GNSSs.

[23] C/A code acquisition may be impossible for very weak signals in the presence of a strong C/A signal.

[24] Satellite PRN01/SVN49 (launched on 24 March 2009) also began transmitting the data-less L5 test signal. This satellite was decommissioned from active service on 6 May 2011. See the status of the GPS constellation at *http://tycho.usno.navy.mil/gpscurr.html*.

[25] The first satellite (PRN25) was launched on 28 May 2010, with full L5 capability; the second (PRN01) on 16 July 2011.

[26] Originally, the signal was developed as a common civil signal for GPS and Galileo, but new satellite navigation providers (Beidou in China, QZSS in Japan) are also adopting L1C as a future standard for international interoperability.

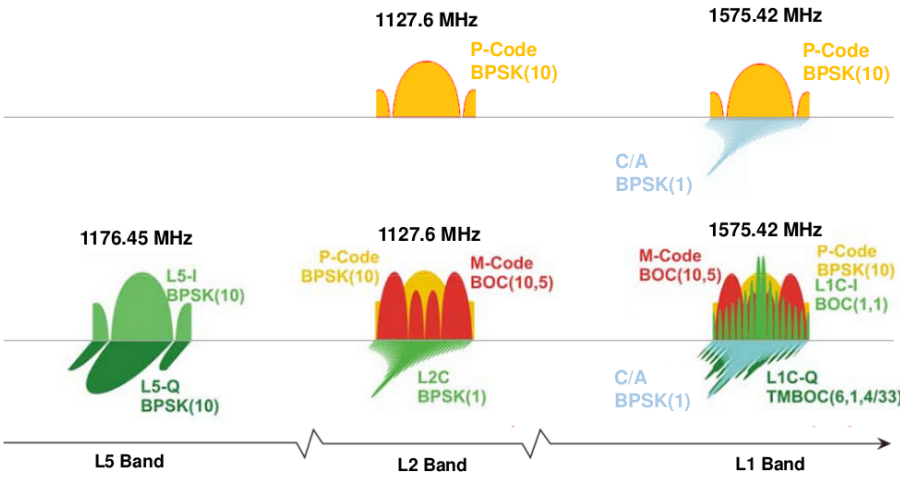


Figure 2.8: Spectra of GPS signals before (*top*) and after modernisation (*bottom*). Courtesy of Stefan Wallner.

Table 2.2: Current and new GPS navigation signals. All civil signals are provided free of charge to all users worldwide (Open Services).

Link	Carrier freq. (MHz)	PRN code	Modulation Type	Code rate (Mcps)	Data rate (bps)	Service
L1	1 575.420	C/A	BPSK(1)	1.023	50	Civil
		P	BPSK(10)	10.23	50	Military
		M	$BOC_{sin}(10,5)$	5.115	N/A	Military
		L1C-I data	MBOC(6,1,1/11)	1.023	50	Civil
		L1C-Q pilot			–	
L2	1 227.600	P	BPSK(10)	10.23	50	Military
		L2C M	BPSK(1)	1.023	25	Civil
		L			–	
		M	$BOC_{sin}(10,5)$		N/A	Military
L5	1 176.450	L5-I data	BPSK(10)	10.23	50	Civil
		L5-Q pilot			-	

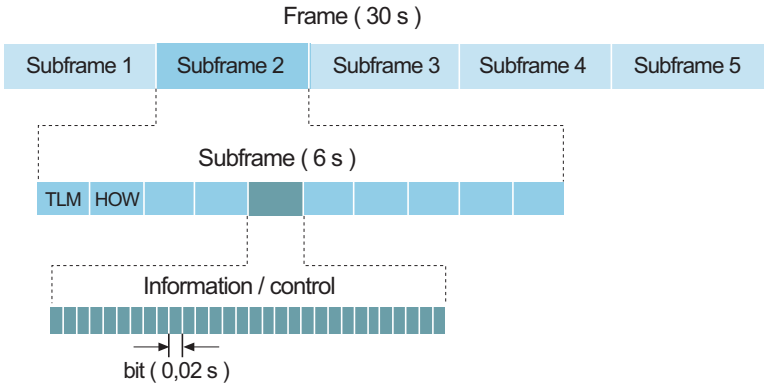
Table 2.2 contains a summary of the current and future GPS signals, frequencies and applied modulations. The ranging code rate and data rate are also given in the table.

2.2.1.1 GPS Navigation Message

Every satellite receives from the ground antennas the navigation data, which are sent back to users through the navigation message.

The *navigation message* contains all the necessary information to allow users to perform the positioning service. This includes the *ephemeris* parameters, needed to compute the satellite coordinates with sufficient accuracy, the *time parameters and clock corrections*, needed to compute satellite clock offsets and time conversions, the *service parameters* with satellite health information, the *ionospheric parameters model*, needed for single-frequency receivers, and the *almanacs*, allowing computation of the position of 'all satellites in the constellation', with a reduced accuracy (1–2 km of 1σ error), which is needed for acquisition of the signal by the receiver. The ephemeris and clock parameters are usually updated every two hours, while the almanac is updated at least every six days.

Figure 2.9: Navigation message.



The current 'legacy' Navigation Message (NAV) is modulated on both carriers at 50 bps. The whole message contains 25 pages (or 'frames') of 30 s each, forming the master frame that takes 12.5 min to be transmitted. Every frame is subdivided into five subframes of 6 s each; in turn, every subframe consists of 10 words, with 30 bits per word (see Fig. 2.9).

Every subframe always starts with the telemetry word TLM, which is necessary for synchronisation. Next, the transference word (HOW) appears. This word provides time information (seconds of the GPS week), allowing the receiver to acquire the week-long P(Y) code segment.

The contents of every subframe are as follows:

- *Subframe 1* contains information about the parameters to be applied to satellite clock status for its correction. These values are polynomial coefficients that allow time onboard to be converted to GPS time. The subframe also contains information on satellite health condition.
- *Subframes 2 and 3* contain satellite ephemerides.
- *Subframe 4* provides ionospheric model parameters (in order to adjust for ionospheric refraction), UTC information, part of the almanac, and indications whether the A/S is activated or not (which transforms the P code into encrypted Y code).
- *Subframe 5* contains data from the almanac and on constellation status. It allows rapid identification of the satellite from which the signal comes. A total of 25 frames are needed to complete the almanac.

Subframes 1, 2 and 3 are transmitted with each frame (i.e. they are repeated every 30 s). Subframes 4 and 5 contain different pages (25 each) of the navigation message (see Fig. 2.9). Thus, transmission of the full navigation message takes $25 \times 30\,\text{s} = 12.5\,\text{min}$.

The contents of subframes 4 and 5 are common to all satellites. Hence the almanac data for all in-orbit satellites can be obtained from a single tracked satellite.

GPS modernisation has introduced four new data messages: CNAV, CNAV-2, MNAV and L5-CNAV. The first three are civil messages, while

MNAV is a military message. They provide more accurate and frequent message data than the legacy NAV.

CNAV and MNAV have a similar structure and modernised data format. The new format allows for more flexibility, better control and improved content. Its design replaces the use of frames and subframes of data (repeated in a fixed pattern) of the original 'legacy' NAV by means of a packetised message-based communications protocol, where individual messages can be broadcast in a flexible order with varying repeat cycles. Moreover, Forward Error Correction (FEC) and advanced error detection (such as a Cyclic Redundancy Check (CRC)) are used to achieve better error rates and reduced data collection times. Furthermore, MNAV includes new improvements for the security and robustness of this military message.

A preliminary description of the CNAV structure can be found in the GPS Interface Specification [IS-GPS-200, 2010]. It includes up to 63 different message types, of which 15 have already been defined.

L5-CNAV is modulated onto the L5I signal component, containing basically the same information as NAV and CNAV, but in a different format. CNAV-2 provides more accurate data, is modulated onto L1C-I and is divided into frames, which are subdivided into three subframes.

2.2.1.2 User Domain GPS Performance

The positioning and timing accuracy for the GPS SPS is given in Table 2.3. These figures are taken from the performance specification documents of the SPS, namely [GPS/SPS-SS, 1995] and [GPS/SPS-PS, 2008], and are based on a 95th percentile level and Signal In Space (SIS) only. The specifications for the PPS are quite similar to those of the SPS with S/A off; see [GPS/PPS-PS, 2007]. These values are very conservative, since the actual performances usually exceed these official values. Indeed, SPS horizontal and vertical accuracies are better than 7 and 12 m, respectively, for 95% of the time that is currently observed. The modernised GPS, with the new L5 signal, is expected to provide civil users with horizontal accuracies of about 2–3 m for 95% of the time, 'as good as PPS users currently enjoy' [McDonald and Hegarty, 2000].

Table 2.3: Predictable positioning and timing accuracy for GPS SPS.

Error (95%)	SPS	
	S/A on	S/A off
Position		
Horizontal	100 m	13 m
Vertical	156 m	22 m
Time	340 ns	40 ns

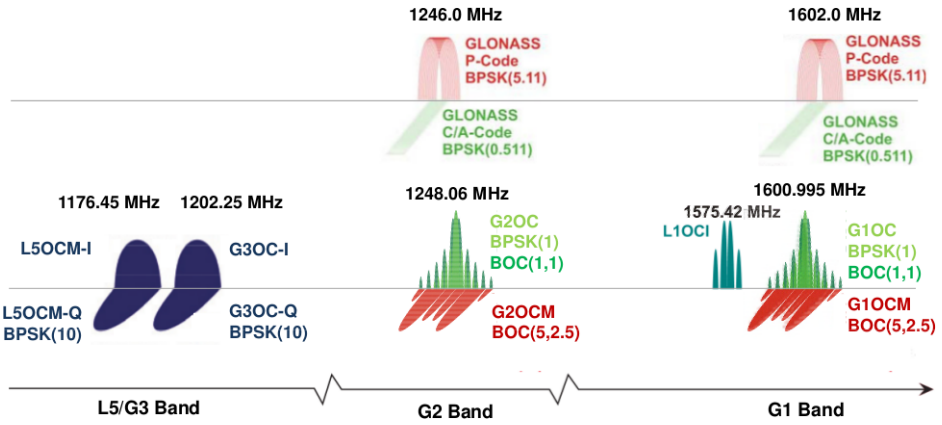


Figure 2.10: Spectra of Glonass signals. Legacy FDMA signals before and after modernisation (*top*), and new CDMA signals after modernisation (*bottom*). Courtesy of Stefan Wallner.

2.2.2 Glonass Signals

Legacy Glonass signals are right-hand circularly polarised and centred on two radio frequencies in the L band, referred to here as the G1 and G2[27] bands, see Fig. 2.10.

Two services are currently available from Glonass:

SPS: The Standard Positioning Service (or Standard Accuracy Signal Service) is an open service, free of charge to worldwide users. The navigation signal was initially provided only in the frequency band G1, but since 2004 the new Glonass-M satellites also transmits a second civil signal in G2.

PPS: The Precise Positioning Service (or High-Accuracy Signal Service) is restricted[28] to military and authorised users. Two navigation signals are provided in the two frequency bands G1 and G2.

In contrast to GPS satellites that share the same frequencies, each Glonass satellite broadcasts at a particular frequency within the band. This frequency determines the frequency channel number of the satellite and allows users' receivers to identify the satellites (with the FDMA technique). Glonass modernisation planning includes the transmission of CDMA signals in the G1, G2 and G3 (L3) bands, and even in the GPS L5 band, in addition to transmitting legacy FDMA signals in the G1 and G2 bands (see Fig. 2.10 and Fig. 2.15 below).

The actual frequency of legacy Glonass signal transmission on G1 and G2 can be derived from the channel number k by applying the following expressions:

Frequency band G1: $f1(k) = 1602 + k \times 9/16 = (2848 + k) \times 9/16\,\text{MHz}$
Frequency band G2: $f2(k) = 1246 + k \times 7/16 = (2848 + k) \times 7/16\,\text{MHz}$

The frequency numbers k were originally envisaged to provide 24 channels, with $k = 1, \ldots, 24$, but according to the guidelines of the International Electric Communication Union, all Glonass satellites launched after 2005

[27] We use G1 and G2 instead of L1 and L2 to better differentiate from GPS. Nevertheless, the ICD uses L1 and L2; see *http://rniikp.ru/en/pages/about/publ/ikd51en.pdf*.

[28] Although code P is not encrypted, its unauthorised use is not recommended by the Russian Ministry of Defence because it may be changed without prior notice.

had to use $k = -7, \ldots, 6$. This change was introduced to avoid interference problems with radio astronomy frequency bands and satellite communication services. The reduction from 24 to 12 channels is compensated for at setup, because two satellites in the same orbit transmit at exactly the same frequency but occupy antipodal locations. Consequently, they will never be simultaneously in view by any user on Earth's surface (nevertheless, space receivers will need to implement discriminating functions to distinguish the satellites).

Two ranging codes, the coarse acquisition C/A (open civil code) and the precise P (military) code, are modulated onto these frequencies together with a navigation message D, using the BPSK technique. The C/A and P codes have periods of 1 ms and 1 s, and chip widths of 586.7 and 58.67 m, respectively, and are about two times noisier than the GPS ones (see Table 2.4).

As in GPS, the C/A code was initially modulated only on G1, while the military code P is modulated on both carrier frequencies, G1 and G2; however, the new Glonass-M satellites (from 2004) also transmit the C/A signal in the G2 frequency band. On the other hand, and unlike GPS, in Glonass the PRN sequences of such codes are common to all satellites, because the receiver identifies the satellite by its frequency.[29]

No S/A (i.e. intentional degradation of the standard accuracy signal) is applied in Glonass, and no P-code encryption has been reported so far. Although the military P code has not been officially published, it has been deciphered by different research groups. Nevertheless, this code may be changed by the Russian Ministry of Defence without prior warning.

Glonass Signal Modernisation: Introduction of New Signals and CDMA Usage

The modernisation of Glonass added a new third frequency G3 to the ARNS band for the Glonass-K satellites. This signal will provide a third civil C/A_2 and military P_2 codes, and is especially suitable for SoL applications. The plans for Glonass signal modernisation are summarised in Fig. 2.10 (further details can be found in [Urlichich et al., 2011] and [Avila-Rodriguez, 2008]).

The addition of CDMA and FDMA signals was initiated first with the Glonass-K launch in February 2011, providing CDMA signals at a frequency $f = 1\,202.025$ MHz in the G3 band (close to the Galileo E5b carrier).[30]

[29] Note that this applies for legacy signals where the FDMA technique is used. For the new Glonass signals, the satellites use the same frequency and are identified with different PRN codes using CDMA.

[30] In 2014 it is planned to launch the first Glonass-K2 satellite, with FDMA signals in the G1 and G2 bands and CDMA signals in G1, G2 and G3 [Urlichich et al., 2011].

Table 2.4: Legacy Glonass signal structure.

Atomic clock frequency	$f_0 = 0.511$ MHz
Frequencies L1	$9/16(2848 + k) =$
	$1602.000 + 0.5625k$ MHz
Wavelength L1	18.7 cm ($k = 0$)
Frequencies L2	$7/16(2848 + k) =$
	$1246.000 + 0.4374k$ MHz
Wavelength L2	24.1 cm ($k = 0$)
P code frequency (chipping rate)	$10f_0 = 5.11$ Mcps
P code wavelength	58.67 m
P code period	1 s
C/A code frequency (chipping rate)	$f_0 = 0.511$ Mcps
C/A code wavelength	586.7 m
C/A code period	1 ms
Navigation message frequency	50 bps
Frame length	30 s (on CA), 10 s (on P)
Total message length	2.5 min (on CA), 12 min (on P)

Table 2.5: Legacy Glonass navigation Signals. All civil signals are provided free of charge to all users worldwide (Open Services). The military codes are for authorised users only. Note: $k = 1, \ldots, 24$ for satellites launched before 2005; subsequently $k = -7, \ldots, 6$.

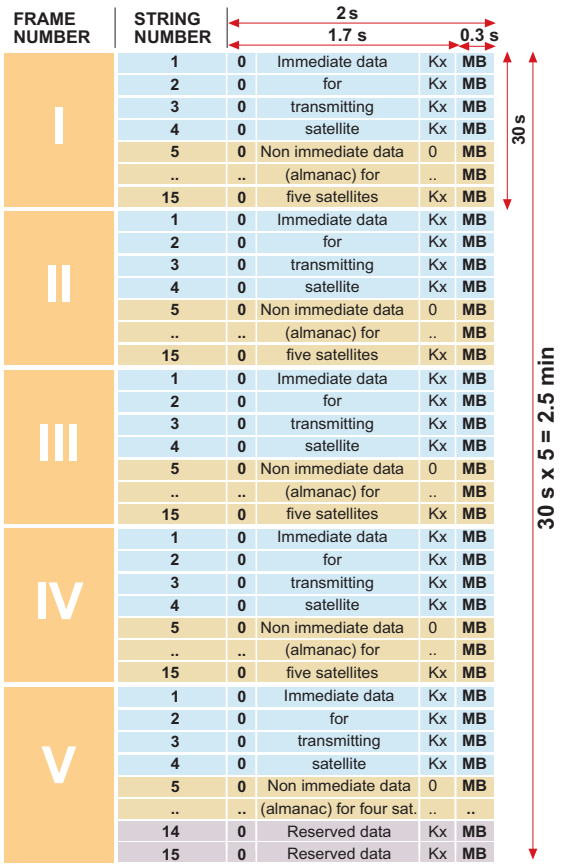
Band	Carrier freq. (MHz)	PRN code	Modulation type	Code rate (Mcps)	Data rate (bps)	Service
G1	1602.0000+	C/A	BPSK(0.511 Mcps)	0.511	50	Civil
	$+0.5625k$	P	BPSK(5.11 Mcps)	5.11	50	Military
G2	1246.0000+	C/A	BPSK(0.511 Mcps)	0.511	50	Civil
	$+0.4375k$	P	BPSK(5.11 Mcps)	5.11	50	Military

2.2.2.1 Glonass Navigation Message

Glonass satellites modulate two navigation messages at 50 bps onto the standard (C/A) and high-accuracy (P) signals (Table 2.5), each message giving users the necessary information for positioning (i.e. parameters to compute the coordinates of the Glonass satellites, their clock offsets and various other system parameters).

The *navigation message of the standard accuracy signal (C/A)* is broadcast as continuously repeating superframes with a duration of 2.5 min. Each superframe consists of five frames of 30 s, and each frame consists of 15 strings of 2 s duration (100 bits long).

The message content divides the data into *immediate data* of the transmitting satellite and *non-immediate data* for the other satellites, see Fig. 2.11. The immediate data are repeated in the first four strings of every frame and comprise the ephemeris parameters, satellite clock offsets, satellite healthy flag and the relative difference between the carrier frequency of the satellite and its nominal value. The non-immediate data are broadcast in strings 5 to 15 of each frame (almanac for 24 satellites). Frames I to IV contain the almanac for 20 satellites (5 per frame), and the fifth frame



FRAME NUMBER	STRING NUMBER	2 s / 1.7 s		0.3 s		
I	1	0	Immediate data	Kx	MB	30 s
	2	0	for	Kx	MB	
	3	0	transmitting	Kx	MB	
	4	0	satellite	Kx	MB	
	5	0	Non immediate data	0	MB	
	..	..	(almanac) for	..	MB	
	15	0	five satellites	Kx	MB	
II	1	0	Immediate data	Kx	MB	
	2	0	for	Kx	MB	
	3	0	transmitting	Kx	MB	
	4	0	satellite	Kx	MB	
	5	0	Non immediate data	0	MB	
	..	..	(almanac) for	..	MB	
	15	0	five satellites	Kx	MB	
III	1	0	Immediate data	Kx	MB	
	2	0	for	Kx	MB	
	3	0	transmitting	Kx	MB	
	4	0	satellite	Kx	MB	
	5	0	Non immediate data	0	MB	
	..	..	(almanac) for	..	MB	
	15	0	five satellites	Kx	MB	
IV	1	0	Immediate data	Kx	MB	
	2	0	for	Kx	MB	
	3	0	transmitting	Kx	MB	
	4	0	satellite	Kx	MB	
	5	0	Non immediate data	0	MB	
	..	..	(almanac) for	..	MB	
	15	0	five satellites	Kx	MB	
V	1	0	Immediate data	Kx	MB	
	2	0	for	Kx	MB	
	3	0	transmitting	Kx	MB	
	4	0	satellite	Kx	MB	
	5	0	Non immediate data	0	MB	
	..	..	(almanac) for four sat.	..	..	
	14	0	Reserved data	Kx	MB	
	15	0	Reserved data	Kx	MB	

30 s x 5 = 2.5 min

Figure 2.11: Legacy Glonass navigation message structure (source [GLONASS ICD, 2008]).

almanac for 4 satellites. The last two strings of frame 5 are reserved bits (the almanac of each satellite uses two strings).

The ephemerides values are predicted from the SCC for a 24 h period, and the satellite transmits a new set of ephemerides every 30 min. These data differ from GPS data. Instead of Keplerian orbital elements, they are provided as Earth-Centred, Earth-Fixed (ECEF)[31] Cartesian coordinates in position and velocity, with lunisolar acceleration perturbation parameters. Glonass-ICD [GLONASS ICD, 2008] provides integral equations based on the fourth-order Runge–Kutta method, which includes the second zonal geopotential harmonic coefficient. The almanac is quite similar to the GPS one, in terms of modified Keplerian parameters, and it is updated approximately once per day.

The *navigation message of the high-accuracy signal (P)* structure is not officially published, but different research groups have decoded it. According to these investigations, each satellite transmits a superframe which is composed of 72 frames, each containing five strings of 100 bits. A frame needs 10 s to be transmitted, hence the total length of the message is 12 min. The first three frames contain the ephemeris for the transmitting satellite.

[31] See Keplerian elements in section 3.2.1 and ECEF coordinates in section A.2.2.

2.2.2.2 User Domain Glonass Performance

No standardised values are published on Glonass positioning and timing accuracy, and the performance can change according to author. Table 2.6 contains the figures for the Glonass SPS [United Nations, 2004].

Table 2.6: Glonass SPS performance according to document [United Nations, 2004].

SPS accuracy (95%)		
Horizontal	Vertical	Time
28 m	60 m	1 μs

Information on the PPS is not included in Table 2.6 due to the lack of data on this service's performance.

Previous performance figures changed dramatically with successive satellites. Indeed, the accuracy has been improving from about 25 m (1σ) in 2006 to about 5–7 m (1σ) in 2010 [Hein, G.W., 2011].

2.2.3 Galileo Signals

In FOC phase, each Galileo satellite will transmit 10 navigation signals in the frequency bands E1, E6, E5a and E5b, each right-hand circularly polarised. These signals are designed to support the different services that will be offered by Galileo and EGNOS,[32], based on various user needs as follows:

OS: The Open Service (OS) is free of charge to users worldwide. Up to three separate signal frequencies are offered within it. Single-frequency receivers will provide performances similar to GPS C/A. In general, OS applications will use a combination of Galileo and GPS signals, which will improve performance in severe environments such as urban areas.

PRS: The Public Regulated Service (PRS) is intended for the security authorities (police, military, etc.) who require a high continuity of service with controlled access. It is under governmental control. Enhanced signal modulation/encryption is introduced to provide robustness against jamming and spoofing. Two PRS navigation signals with encrypted ranging codes and data will be available.

CS: The Commercial Service (CS) provides access to two additional signals protected by commercial encryption (ranging data and messages). Higher data rates (up to 500 bps) for broadcasting data messages are introduced.

[32]The European Geostationary Navigation Overlay Service (EGNOS) is a Satellite-Based Augmentation System (SBAS) that enhances the US GPS satellite navigation system to make it suitable for safety-critical applications such as flying aircraft or navigating ships through narrow channels. More details can be found in [Ventura-Traveset, J. and Flament, 2006] and at *http://www.esa.int/esaNA/egnos.html*.

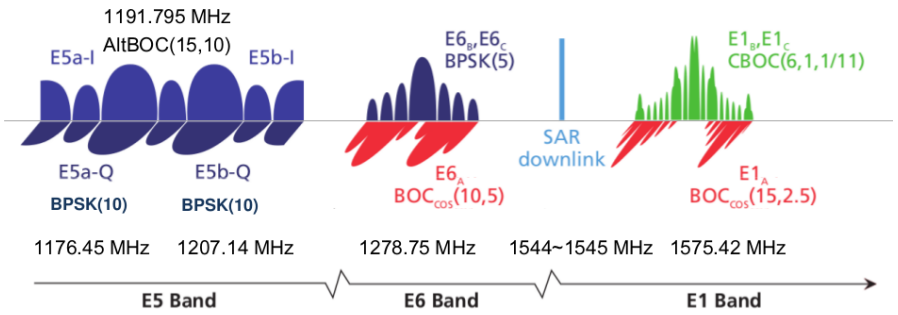


Figure 2.12: Spectra of Galileo signals. Courtesy of Stefan Wallner.

SAR: This service contributes to the international Cospas–Sarsat system for Search and Rescue (SAR). A distress signal will be relayed to the Rescue Coordination Centre and Galileo will inform users that their situation has been detected.

SoL: The Safety-of-Life (SoL) Service is already available for aviation to International Civil Aviation Organization (ICAO) standards thanks to EGNOS; Galileo will further improve the service performance.

As in GPS, all satellites share the same frequencies, and the signals are differentiated by the CDMA[33] technique [Galileo SIS ICD, EU, 2010]. As mentioned earlier, these signals can contain data and pilot channels. Both channels provide ranging codes, but the data channels also include navigation data. Pilot channels (or pilot tones) are data-less signals, so no bit transition occurs, thus helping the tracking of weak signals. The spectra of Galileo signals are given in Fig. 2.12, where the data and pilot channels are plotted in orthogonal planes.

A brief description of each signal follows:[34]

E1 supports the OS, CS, SoL and PRS services. It contains three navigation signal components in the L1 band. The first one, E1-A, is encrypted and only accessible to authorised PRS users; it contains PRS data. The other two components, E1-B and E1-C, are open access signals with unencrypted ranging codes accessible to all users. E1-B is a data channel and E1-C a pilot (or data-less) channel. The E1-B data stream, at 125 bps, also contains unencrypted integrity messages and encrypted commercial data. The MBOC modulation is used for the E1-B and E1-C signals, which is implemented by the Composite Binary Offset Carrier (CBOC), see Table 2.7 and Fig. 2.12. More details can be found in [Avila-Rodriguez et al., 2006] and [Avila-Rodriguez et al., 2007]. (Note that the E1 band is shared with GPS L1 and Beidou B1, see Fig. 2.15 below.)

E6 is a dedicated signal for supporting the CS and PRS services. It provides three navigation signal components transmitted in the E6 band. As with E1, the first one, E6-A, is encrypted and only accessible to authorised PRS users, carrying PRS data. The other two, E6-B and E6-C, are commercial access signals and include a data channel E6-B and a pilot (or data-less) channel E6-C. The E6 ranging codes and

[33] That is, where the spread spectrum codes enable the satellite to transmit at the same frequencies simultaneously.

[34] Mainly from [Galileo SIS ICD, EU, 2010].

data are encrypted. A data rate of 500 bps allows the transmission of added-value commercial data. (Note that the E6 band is shared with Beidou B3, see Fig. 2.15 below.)

E5a supports OS. It is an open access signal transmitted in the E5a band and includes two signal components, a data channel, E5a-I, and a pilot (or data-less) channel, E5a-Q. The E5a signal has unencrypted ranging codes and navigation data, which are accessible to all users. It transmits the basic data to support navigation and timing functions, using a relatively low 25 bps data rate that enables more robust data demodulation. (Note that the E5a band is shared with GPS L5, Beidou B2a and future Glonass L5 signals, see Fig. 2.15 below.)

E5b supports the OS, CS and SoL services. It is an open access signal transmitted in the E5b band and includes two other signal components: the data channel E5b-I and the pilot (or data-less) channel E5b-Q. It has unencrypted ranging codes and navigation data accessible to all users. The E5b data stream also contains unencrypted integrity messages and encrypted commercial data. The data rate is 125 bps. (Note that the E5b band is shared with Beidou B2b and Glonass G3 (slightly shifted), see Fig. 2.15 below.)

The E5a and E5b signal components are modulated onto a single E5 carrier frequency at 1 191.795 MHz using a technique known as Alternate Binary Offset Carrier (AltBOC). The composite of the E5a and E5b signals is denoted as the E5 signal and can be processed as a single large-bandwidth signal with an appropriate user receiver implementation, which results in a low-multipath and tracking noise signal, see [Avila-Rodriguez, 2008].

A summary of Galileo signals, frequencies and applied modulations is presented in Table 2.7. The ranging code rate and data rate are also given in the table.

Table 2.7: Galileo navigation signals. The two signals located in the E5a and E5b bands respectively are modulated onto a single E5 carrier frequency of 1191.795 MHz using the AltBOC technique: AltBOC(15,10).

Band	Carrier freq. (MHz)	Channel or sig. comp.	Modulation type	Code rate (Mcps)	Data rate (bps)	Services
E1	1575.420	E1-A data	$BOC_{cos}(15,2.5)$	2.5575	N/A	PRS
		E1-B data	MBOC(6,1,1/11)	1.023	125	OS, CS, SoL
		E1-C pilot			–	
E6	1278.750	E6-A data	$BOC_{cos}(10,5)$	5.115	N/A	PRS
		E6-B data	BPSK(5)		500	CS
		E6-C pilot			–	
E5a	1176.450	E5a-I data	BPSK(10)	10.23	25	OS
		E5a-Q pilot			–	
E5b	1207.140	E5b-I data	BPSK(10)	10.23	125	OS, CS, SoL
		E5b-Q pilot			–	

Table 2.8: Contents of the Galileo message types (from [Hofmann-Wellenhof et al., 2008]).

Message-type	F/NAV	I/NAV		C/NAV	G/NAV	
Galileo Services	OS	OS/CS/SoL		CS	PRS	
Channels	E5a-I	E1B	E5b-I	E6B	E1A	E6A
Navigation/Positioning	×	×	×		×	
Integrity		×	×		×	
Supplementary				×		
Public Regulated					×	
Search and Rescue		×				

2.2.3.1 Galileo Navigation Message

The Galileo satellites will broadcast five types of data in four navigation messages: the Freely accessible Navigation Message (F/NAV) and Integrity Navigation Message (I/NAV), the Commercial Navigation Message (C/NAV) and the Governmental Navigation Message (G/NAV). Table 2.8 summarises the contents of the Galileo messages, with an indication of the associated channels and services.

The Galileo ephemeris parameters are Keplerian-like orbital elements as in GPS (see section 3.3.1). The nominal period update is 3 h and is valid for a 4 h time interval. The 1 h overlap interval is intended to help with short outages or delays. The Galileo almanac is also similar to the GPS and Glonass ones.

The *Supplementary* data are expected to provide information to support the different envisaged commercial services as differential corrections to the High-Precision Positioning Service (HPPS), and different kinds of data for weather alerts, traffic information, etc. The data are encrypted to control access to authorised users by the service providers.

The *Public Regulated* data are under governmental control and are devoted to the PRS. The system will guarantee a high continuity of service with controlled access by encryption of data.

The *Search and Rescue* data will provide the capability to send acknowledgement SAR messages to a beacon equipped with a suitable Galileo receiver.

The complete navigation message is transmitted on each data channel as a sequence of frames (Fig. 2.13). A frame comprises a certain number of subframes, and a subframe comprises several pages (Table 2.9). This arrangement allows the three different main categories of data to be transmitted:

- Repeated at a fast rate (for urgent data): page.
- Medium rate (like data required for a warm start): subframe.
- Slow rates (like data required for a cold start): frame.

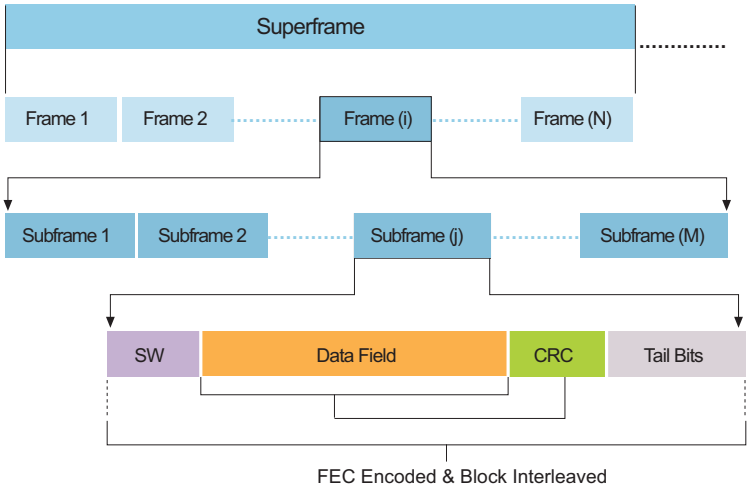


Figure 2.13: Galileo navigation message structure (from [Powe, M., 2006]).

The page starts with a Synchronisation Word (SW) followed by the data field. After the data, Cyclic Redundancy Check (CRC) parity bits are provided to detect the reception of corrupted data. The page ends with tail bits for the Forward Error Correction (FEC) encoding.

Three levels of error coding are applied to the Galileo message data stream:
(1) a CRC with error detection capabilities after recovery of the received data; (2) a one-half rate FEC, with tail bits (sequence of zeros) to allow Viterbi decoding; and (3) block interleaving of the resulting frames, to provide robustness to the FEC decoding algorithm by avoiding packets of errors. This scheme allows the bit error rate to be reduced at increased data rates.

Table 2.9: Frame structure.

Message	Channel	Data rate (bps)	Page duration (s)	Pages in a Subframe	Subframes in a frame
F/NAV	E5a-I	25	10	5	12
I/NAV	E1B, E5b-I	125	1	30	18
C/NAV	E6B	500	1	15	8
G/NAV	E1A, E6A	N/A			

2.2.3.2 User Domain Galileo Performance

The performance requirements for the Galileo services are summarised in Table 2.10.

Table 2.10: Performance requirements (from [Hofmann-Wellenhof et al., 2008], page 371) for the Galileo services: (1)single-frequency accuracy depends on which frequency is used; (2) offset Galileo to UT over 24 h. Integrity service performance is excluded from this table because this service is being redefined.

Satellite-only Service	Open Service	Commercial Service	Public Regulated Service
Coverage	Global	Global	Global
Accuracy (95%) Single-freq.(1) Dual-freq.	15 m/24 m H; 35 m V 4 m H; 8 m V		15 m/24 m H; 35 m V 6.5 m H; 12 m V
Timing accuracy (95%)(2)	30 ns	30 ns	30 ns
Service availability	99.5%	99.5%	99.5%
Access control	Free open access	Controlled access of ranging codes and navigation data message	Controlled access of ranging codes and navigation data message
Certification and service guarantees	–	Guarantee of service possible	Built for accreditation and guarantee of of service

2.2.4 Beidou Signals

Beidou Phase II/III satellites will transmit right-hand circularly polarised signals centred on three radio frequencies in the L band, referred to here as the B1, B2 and B3 bands, see Fig. 2.14.

Two services are foreseen for the Beidou system (in Phase II as a regional service and Phase III as a global service):

Open Service: The SPS (or Standard Accuracy Signal Service) is an open service, free of charge to all users.

Authorised Service: This service will ensure very reliable use, providing safer positioning, velocity and timing services, as well as system information, for authorised users [United Nations, 2010].

Like GPS, Galileo or the new Glonass signals, Beidou ranging signals are based on the CDMA technique. The different navigation signals, structure and supported services, according to the current signal plan for Phase II and Phase III, are summarised in Tables 2.11 and 2.12 and illustrated in Fig. 2.14. See also Fig. 2.15 which depicts all the GNSS signals.

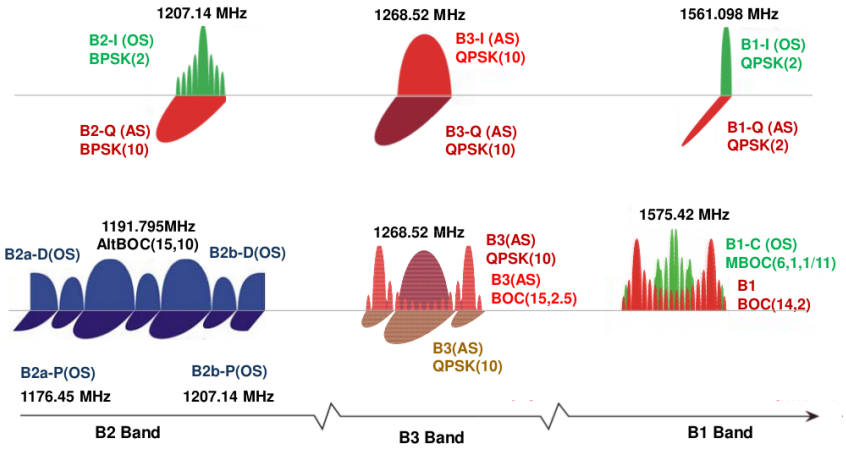


Figure 2.14: Spectra of Beidou signals: Phase II (*top*) and Phase III (*bottom*). Courtesy of Stefan Wallner.

In late December 2011, an English version of the 'test version' of the ICD for Beidou [BeiDou-SIS-ICD-Test, 2011] was published. This is an 11-page document that covers the open B1 civil signal centred at 165.098 MHz (see Table 2.11). The official ICD [BeiDou-SIS-ICD, 2012] was published one year after. This 77-page document provides details of the navigation message, including parameters of the satellite almanacs and ephemerides that were missing in previous 'test version' of the ICD.

Table 2.11: Beidou Phase II navigation signals [Hein, G.W., 2010]. Quadrature Phase-Shifted Keying (QPSK) and BPSK modulation squemes are applied.

Band	Carrier freq. (MHz)	PRN code	Modulation type	Code rate (Mcps)	Service
B1	1561.098	B1-I	QPSK(2)	2.046	Open
		B1-Q			Authorised
B2	1207.14	B2-I	BPSK(2)	2.046	Open
		B2-Q	BPSK(10)	10.23	Authorised
B3	1268.52	B3	QPSK(10)	10.23	Authorised

Table 2.12: Beidou Phase III navigation signals. [United Nations, 2010].

Band	Carrier freq. (MHz)	PRN code	Modulation type	Code rate (Mcps)	Data rate (bps)	Service
B1	1575.42	B1-C_D	MBOC(6,1,1/11)	1.023	50	Open
		B1-C_P			–	
		B1	BOC(14,2)	2.046	50	Authorised
B2	1191.795	B2-a_D	AltBOC(15,10)	10.23	25	Open
		B2-a_P			–	
		B2-b_D			50	
		B2-b_P			–	
B3	1268.52	B3	QPSK(10)	10.23	500	Authorised
		B3-A_D	BOC(15,2.5)	2.5575	50	
		B3-A_P			–	

2.2.4.1 User Domain Beidou Performances

Since 27 December 2011, Beidou has provided an initial Phase II operating service to China and the neighbouring area, using a reduced constellation of 10 satellites. According to the press statement at the conference of the State Council Information Office [Chengqi, R., 2011], 'Through previous testing upon the system, the main service performances are:

- Service coverage area: from 84°E to 160°E, from 55°S to 55°N.
- Positioning accuracy: Horizontally, 25 meters, Vertically, 30 meters.
- Velocity accuracy: 0.4 meters per second.
- Timing accuracy: 50 nanoseconds.'

Following the launching plan for the Beidou system, four more navigation satellites were launched in 2012 to further expand the service area, improve performances and provide services for most parts of the Asia-Pacific region. Completion of the 35-satellite constellation is projected for 2020 (Beidou Phase III).

The expected open free service performances in Beidou Phase III are summarised as follows: [United Nations, 2010]:

- Coverage area: global.
- Positioning accuracy: 10 m (95%).
- Velocity accuracy: 0.2 m/s.
- Timing accuracy: 20 ns.

2.2.5 Summary of GNSS Signals

To summarise the previous sections, Fig. 2.15 compares the different GNSS signals and their relative allocations on the associated frequency bands.

2.2.5.1 Authors' Disclaimer

The aim of this chapter was not to provide an extensive description of the current and future GNSSs but to give an overview of them. Note that the architectures, signal plans, performances, etc., presented above are subject to continuous evolution and upgrading, so some information may become outdated.

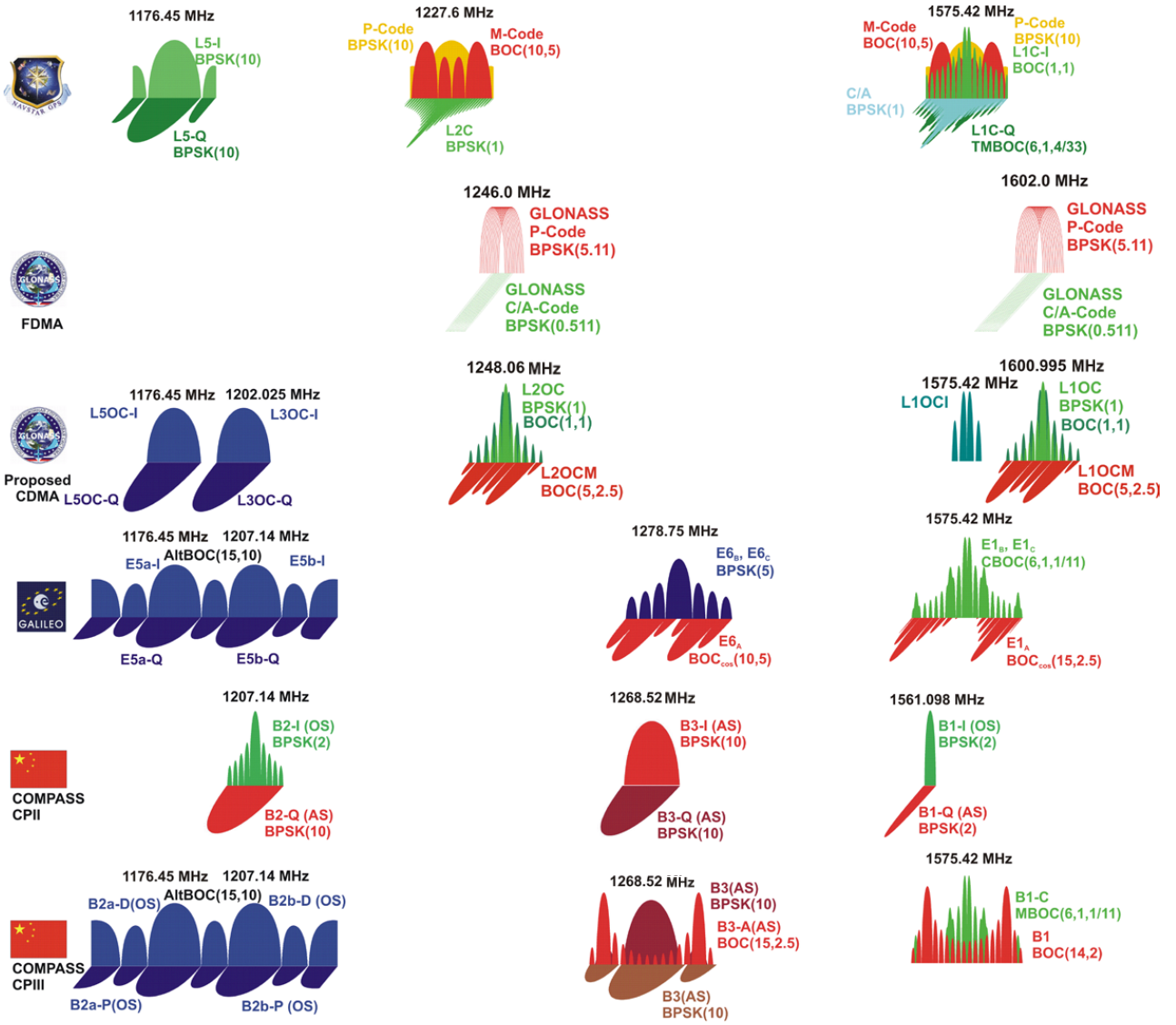


Figure 2.15: Global layout summarising the GPS, Glonass, Galileo and Beidou (Compass) signals. Courtesy of Stefan Wallner.

3. GNSS Time Reference, Coordinate Frames and Orbits

3.1 Time and Reference Frames

Accurate and well-defined time references and coordinate frames are essential in GNSSs, where positions are computed from signal travel time measurements and provided as a set of coordinates.

This section contains a brief review of the fundamentals of time and coordinate frames.The aim here is to provide the necessary background to follow the remaining chapters of this book. A deeper treatment of these topics is provided in Appendix A, where the concepts and equations for time and coordinate frame transformations are provided in detail.

3.1.1 Time

Everyday life follows the alternation of day and night, and the seasons of the year, so the initial conception of time was based on the motion of the Sun. However, as science and technology evolved, more precise, uniform and well-defined time scales were needed.

Several time references are currently in operation, based on different periodic processes associated with Earth's rotation, celestial mechanics or transitions between the energy levels in atomic oscillators. Table 3.1 summarises the different time systems and their associated periodic process.

Table 3.1: Different time systems (from [Hofmann-Wellenhof et al., 2003]).

Periodic process	**Time**
Earth's rotation	Universal Time (UT0, UT1, UT2) Greenwich Sidereal Time (Θ)
Earth revolution	Terrestrial Dynamic Time (TDT) Barycentric Dynamic Time (TDB)
Atomic oscillators	International Atomic Time (TAI) Coordinated Universal Time (UTC) GNSS Reference Time

Universal and sidereal times are associated with the daily rotation of the Earth. Universal Time (solar time) uses the Sun as a reference. Sidereal time uses the vernal equinox (the Aries point).[1] This leads to the fact that, in a year, both times differ by 24 h (one lap) and, hence, the sidereal day is shorter than the solar day by $3^{\mathrm{m}}\,56\overset{\mathrm{s}}{.}4$, see section A.1.1.3. That is,

$$1 \text{ mean sidereal day} = 1 \text{ mean solar day} - 3^{\mathrm{m}}\,56\overset{\mathrm{s}}{.}4$$

UT (UT0, UT1, UT2) is not completely uniform,[2] so atomic time TAI was introduced to achieve a more uniform time scale. TAI and UT accumulate drift over time and, consequently, UTC was defined, which is an atomic time that remains within 0.9 s of UT1. This is maintained by the systematic introduction of a certain number of seconds called *leap seconds*.[3] Over time, the difference between UTC and TAI varies in integer leaps of 1 s. UTC is generated by the Bureau International des Poids et Mesures (BIPM), located in Paris. Leap seconds are determined by the International Earth Rotation and Reference Systems Service (IERS).

UTC is obtained from about 250 caesium clocks and hydrogen masers located in about 65 different laboratories, distributed around the world, and applying a set of algorithms to ensure a uniform time. It is not determined in real time, but generated with a delay of about half a month. Real-time estimates of UTC are computed and provided by different centres, such as UTC(USNO), from the United States Naval Observatory (USNO); UTC(NIST), from the National Institute of Standards and Technology (NIST); and UTC(SU) from Russia. In general, UTC(k) is a realisation of UTC by a given laboratory k, see [Lewandowski et al., 2006].

The following relations are met (see Appendix A.1.5):

$$\mathrm{TAI}=\mathrm{UTC}+1^{\mathrm{s}}\times n, \qquad \mathrm{TAI}=\mathrm{TDT}-32\overset{\mathrm{s}}{.}184$$

$$\mathrm{UTC}=\mathrm{UT1}+dUT1, \qquad |dUT1| < 0\overset{\mathrm{s}}{.}9$$

where n is the number of leap seconds introduced for a given epoch (e.g. 1 Jan 1999 $n = 32$, 1 Jan 2006 $n = 33$, 1 Jan 2009 $n = 34$, 1 Jul 2012 $n = 35$); see Fig. 3.1.

GPS Time (GPST) is a continuous time scale (no leap seconds) defined by the GPS control segment on the basis of a set of atomic clocks at the MCS and onboard the satellites. It starts at 0^{h} UTC (midnight) of 5–6 January 1980 ($6\overset{\mathrm{d}}{.}0$). At that epoch, the difference TAI−UTC was 19 s, hence GPS−UTC $= n - 19^{\mathrm{s}}$. GPST is synchronised with UTC(USNO) at the 1 ms level (modulo 1 s), but actually kept within 25 ns.

[1] The Aries point is a fictitious direction pointing to the Aries Constellation about 2000 years ago. Due to precession of the equinox (see page 43), the Aries point moves continuously over the ecliptic (apparent orbit of the Sun relative to Earth) by about $50.26''$ each year, therefore in 2000 years it moves through an arc of 30° in a retrograde way. Currently, the Aries point is moving out of the Pisces Constellation and entering Aquarius.

[2] Earth's rotation is not uniform. Different phenomena such as tidal friction, mass transport due to seasonal changes, earthquakes, etc., affect the distribution of the Earth's mass (moment of inertia), introducing fluctuations in its rotation. UT0 is a time scale based on the instantaneous rotation of the Earth, UT1 is adjusted from observed periodic variations (polar motion effects, up to 0.06 s), and UT2 is obtained by correcting its other additional irregularities.

[3] From 1 January 1958 (where TAI = UTC = UT1) until 1 July 2012, 35 s were added.

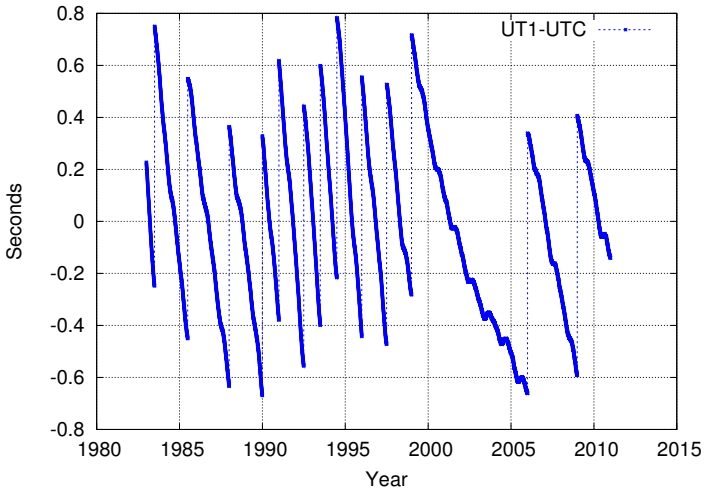


Figure 3.1: UT1−UTC and the leap second adjustments (data from USNO).

Glonass Time (GLNT) is generated by the Glonass Central Synchroniser. The difference between UTC(SU) and GLNT should not exceed 1 ms plus 3 h[4] (i.e., $t_{\text{Glonass}} = \text{UTC(SU)} + 3^{\text{h}} - \tau$, where $|\tau| < 1\,\text{ms}$), but τ is typically better than 1 μs. Note that, unlike GPS, Galileo or Beidou, the Glonass time scale implements leap seconds, like UTC.

Galileo System Time (GST) is a continuous time scale maintained by the Galileo central segment and synchronised to TAI with a nominal offset below 50 ns. The GST start epoch is 0^{h} UTC on Sunday, 22 August 1999.

BeiDou Time (BDT) is a continuous time scale starting at 0^{h} UTC on 1 January 2006, and is synchronised to UTC within 100 ns (modulo 1 s).

In order to facilitate calculations for long time intervals,[5] the Julian date is used (after Julio Scalier). It has as reference epoch the 1st of January of 4713 BC and, starting from there, days are counted in a correlative way. The Julian Day (JD) starts at 12^{h} of the corresponding civil day (e.g. $6^{\text{d}}.0$ January 1980 = JD 2 444 244.5). The current reference standard epoch for the scientific community is

$$\boxed{\text{J2000.0} = 1^{\text{d}}.5 \text{ January } 2000 = \text{JD } 2\,451\,545.0}$$

The Modified Julian Day (MJD) is also used, and is obtained by subtracting 2 400 000.5 days from the JD.

The following expression gives the JD for a civil date[6] (YYMMDDUT) (see equation (A.4) and, for instance, program `sub_cal2cal.f` in Volume II):

$$\text{JD} = \text{int}[365.25 \times y] + \text{int}[30.6001 \times (m+1)] + \text{DD} + \text{UT(h)}/24 + 1\,720\,981.5$$

where:

$$\begin{aligned} y &= \text{YY} - 1,\ m = \text{MM} + 12, && \text{MM} \leq 2 \\ y &= \text{YY},\ m = \text{MM}, && \text{MM} > 2 \end{aligned}$$

[4] The difference between Moscow time and Greenwich Mean Time (GMT).

[5] The calendar has undergone important adjustments throughout history, due to the fact that the length of a year is not exactly 365 days. For example, on 5 October 1582, Pope Gregory XIII introduced a leap period of 10 days – the Gregorian reform – so that the date became Friday the 15th. As an anecdote, Santa Teresa de Jesus died on Thursday the 4th and was buried on Friday the 15th, the following day (see *http://www.newadvent.org/cathen/14515b.htm*).

[6] This expression is valid between March 1900 and February 2100.

From the JD, and taking into account that GPS reference date ($6^{d}.0$ January 1980) corresponds to JD 2 444 244.5, one can immediately obtain the GPST. Then, using modulo 7, the GPS week[7] can be found.

The inverse transformation is given in Appendix A by expressions (A.6).

As another example, the 00:00 hours (Moscow time) on 1 January 1975 (21:00 hours GMT on 31 December 1974) used by Glonass as a reference time in some computations corresponds to JD 2 442 413.375 (GMT).

Greenwich Mean Sidereal Time (GMST) (θ_G) and UT are related by the following expression (see equation (A.34) and program `sub_sid.f` in Volume II):

$$\theta_G = r \cdot \mathrm{UT1} + 6^{h}41^{m}50^{s}.5481 + 8\,640\,184^{s}.812\,866\,T_u + 0^{s}.093\,104\,T_u^2 - 6^{s}.2 \cdot 10^{-6}\,T_u^3$$

where: $r = 1.002\,737\,909\,3$ and $T_u = (\mathrm{JD}_{[\mathrm{UT1\ date}]} - 2\,451\,545.0)/36\,525.0$

3.1.2 Reference Systems and Frames

Satellite coordinates and users' receivers must be expressed in a well-defined reference system. Therefore an accurate definition and determination of such systems is essential to ensure precise positioning in GNSS. Two of the main reference systems used in satellite navigation (see Fig. 3.2) are introduced below: the *Conventional Celestial Reference System* (also called the Conventional Inertial System, CIS) and the *Conventional Terrestrial Reference System* (also called the Coordinated Terrestrial System, CTS).

- *Conventional Celestial Reference System (CRS)*: This is a quasi-inertial reference system.[8] It has its origin at the Earth's centre of mass, or geocentre. The X-axis points in the direction of the mean vernal equinox at the J2000.0 epoch, the Z-axis is orthogonal to the plane defined by the mean equator at the J2000.0 epoch (fundamental plane) and the Y-axis is orthogonal to the former axes, so the system is directly (right-handed) oriented. The practical implementation is called the (conventional) Celestial Reference Frame (CRF)[9] and it is determined from a set of precise coordinates of extragalactic radio sources (i.e. it is fixed with respect to distant objects of the Universe). The mean equator and equinox J2000.0 were defined by International Astronomical Union (IAU) agreements in 1976, with 1980 nutation series as in [Seidelmann, 1982] and [Kaplan, 1981], which are valid analytic expressions for long time intervals (the former reference epoch was 1950.0).

- *Conventional Terrestrial Reference System (TRS)*: This is a reference system co-rotating with the Earth in its diurnal rotation, and

[7] The GPS week starts from Saturday night to Sunday.

[8] It is not strictly an inertial system, because it is affected by the acceleration due to Earth's motion around the Sun (annual revolution).

[9] Note that reference *system* and reference *frame* are different concepts. The first is understood as 'a theoretical definition', including models and standards for its implementation, while the second is its 'practical implementation' through observations and a set of reference coordinates (a set of fundamental stars for the CRF, or fiducial stations, for the Terrestrial Reference Frame).

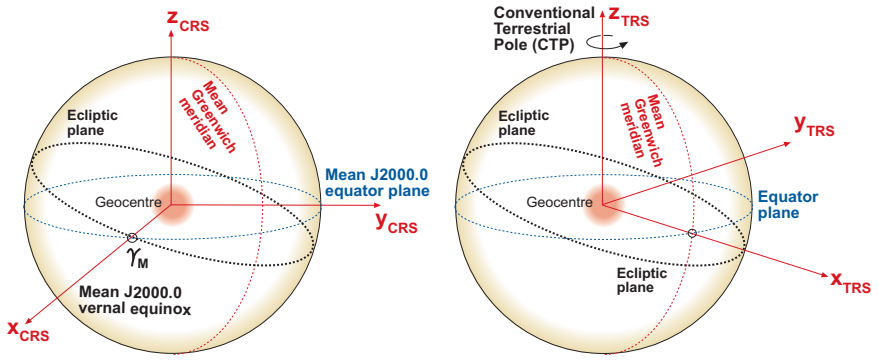


Figure 3.2: Reference systems: CRS (*left*) and TRS (*right*).

is also called the Earth-Centred, Earth-Fixed (ECEF) system. Its definition involves a mathematical model of the physical Earth in which point positions are expressed and have small temporal variations due to geophysical effects (plate motion, tides, etc.). The TRS has its origin in Earth's centre of mass. The Z-axis is identical to the direction of Earth's axis of rotation as defined by the Conventional Terrestrial Pole (CTP), the X-axis is defined as the intersection of the equatorial plane (orthogonal to the Z-axis) with the mean Greenwich meridian, and the Y-axis is orthogonal to both of them, directly orienting the system. The realisation of this system is called the (conventional) Terrestrial Reference Frame (TRF) and it is determined through the coordinates of a set of points on Earth serving as reference points.[10] An example of the TRF is the International Terrestrial Reference Frame (ITRF), introduced by the International Earth Rotation and Reference Systems Service (IERS) and updated every year (ITRF98, ITRF99, etc.). Other terrestrial reference frames are the World Geodetic System 84 (WGS-84), which is applied for GPS, the Parametry Zemli 1990 (Parameters of the Earth 1990) (PZ-90) for Glonass, the Galileo Terrestrial Reference Frame (GTRF) for Galileo, or the China Geodetic Coordinate System 2000 (CGCS2000) for Beidou.

3.1.2.1 Transformation between Celestial and Terrestrial Frames

Coordinate transformations between CRF and TRF frames are performed by mean of rotations corresponding to *precession*, *nutation* and *pole movement*, described briefly as follows.

- *Precession and nutation* (forced rotation): Earth's axis of rotation (and its equatorial plane) is not kept fixed in space (i.e. *in relation to so-called 'fixed stars'*), but rotates about the pole of the ecliptic, as shown in Fig. 3.3. This movement is due to the effect of the gravitational attraction of the Moon and the Sun and major planets over the terrestrial ellipsoid. The total movement can be split into a

[10] A conventional TRF is defined as a set of physical points with precisely determined coordinates in a specific coordinate system that is the realisation of an ideal TRS [Boucher and Altamimi, 2001].

secular component (*precession*, with a period of 26 000 years) and a periodic component (*nutation*, with a period of 18.6 years).

- *Pole movement* (free rotation): Due to the structure of Earth's distribution of mass and its variation, Earth's axis of rotation is not fixed *in relation to Earth's crust*. It moves about on the surface of Earth within a square of about 20 m side in relation to a point with fixed coordinates on Earth. This movement has a period of about 430 sidereal days (the Chandler period). On the other hand, Earth's velocity of rotation is not constant but changes in time (although in very small quantities[11]), as mentioned in the previous section.

Detailed expressions for the transformation between the CRF and TRF are given in Appendix A. The following equations briefly summarise such a transformation.

For a given epoch t, the coordinate transformation can be decomposed into a rotation matrix (i.e. orthogonal matrix)[12] product as

$$[\mathbf{TRF}] = \mathbf{R_M}(t)\,\mathbf{R_S}(t)\mathbf{N}(t)\mathbf{P}(t)[\mathbf{CRF}] \tag{3.1}$$

and the inverse transformation

$$[\mathbf{CRF}] = \mathbf{P}^T(t)\mathbf{N}^T(t)\mathbf{R_S}^T(t)\mathbf{R_M}^T(t)\,[\mathbf{TRF}] \tag{3.2}$$

where:

[CRF]	is the coordinate vector in the CRF,
[TRF]	is the coordinate vector in the TRF,
P	is the transformation matrix associated with the precession between the reference epoch and the epoch t,
N	is the transformation matrix associated with the nutation at epoch t,
$\mathbf{R_S}$	is the transformation matrix associated with Earth's rotation around the Celestial Ephemeris Pole (CEP) axis,
$\mathbf{R_M}$	is the transformation matrix associated with the polar motion.

The matrices **P** and **N** are associated with the rotations needed to transform the coordinates from the **[CRF]** to the **[CEP]**. They are provided by analytical expressions and do not require external parameters (see Appendix A.2.5.1).

The matrices $\mathbf{R_S}$ and $\mathbf{R_M}$ are associated with the rotations needed to transform the coordinates from the **[CEP]** to the **[TRF]**. Their computation requires the Earth Rotation Parameters (ERP) and Earth Orientation Parameters (EOP) files that are compiled periodically (see the website *http://acc.igs.org*). More details can be found in Appendix A.2.5.2.

The transformation matrix for the polar motion is (see equation (A.38) in Appendix A)

$$\mathbf{R_M}(t) = \mathbf{R}_2(-x_p)\,\mathbf{R}_1(-y_p) \tag{3.3}$$

[11] Due to the friction of water in shallow seas, atmospheric movements, abrupt displacements in Earth's interior (in 1955, the rotation suddenly slowed by $41 \cdot 10^{-6}$ s), etc. Note that TRS is tied to the Greenwich meridian and therefore rotates with Earth.

[12] Note that these matrices verify $\mathbf{R}^T(\theta) = \mathbf{R}^{-1}(\theta) = \mathbf{R}(-\theta)$.

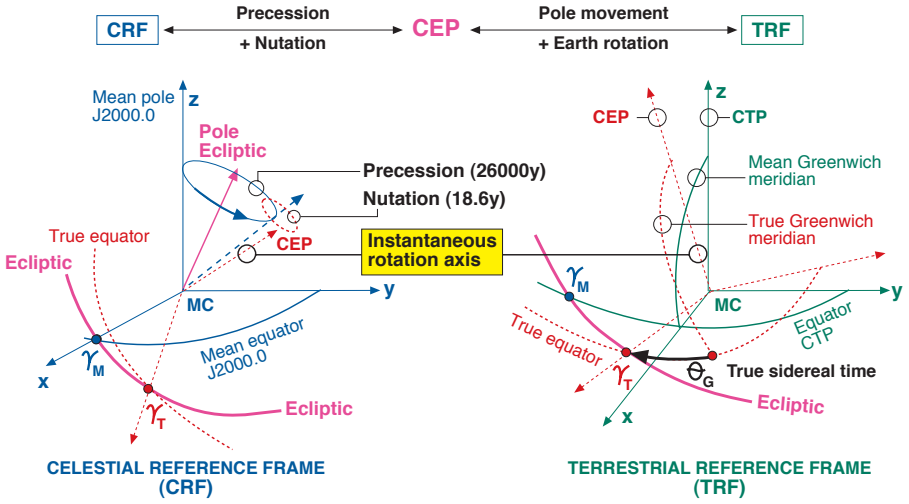


Figure 3.3: Transformations between the CRF and TRF.

where x_p and y_p are the coordinates of the CEP in the TRF, and $\mathbf{R}_1$ and $\mathbf{R}_2$ are the rotation matrices defined by equation (3.6) below.

The transformation matrix associated with Earth's rotation around the CEP-axis is given by (see also equation (A.32) in Appendix A)

$$\mathbf{R_S}(t) = \mathbf{R}_3(\Theta_G) \tag{3.4}$$

where Θ_G is the Greenwich true sidereal time at epoch t and $\mathbf{R}_3$ is from equation (3.6) below.

Figure 3.3 outlines the transformation between the celestial and terrestrial frames. Through precession and nutation corrections, the *mean* equator and mean equinox J2000.0 (γ_M) are transformed to the *true* equator and true equinox (γ_T) at the epoch of observation. They define a reference system with the Z-axis in the direction of the *instantaneous* axis of rotation of Earth (CEP), and with the X-axis pointing to the *true* Aries point. Finally, using ERP and EOP files, the CEP system is transformed to the TRF.[13]

3.1.2.2 Transformation between Terrestrial Frames

From elemental linear algebra, all transformations between two Cartesian coordinate systems can be decomposed into a shift vector ($\Delta\mathbf{X} = [\Delta x, \Delta y, \Delta z]$), three consecutive rotations around the coordinate axes (θ_1, θ_2, θ_3), and a scale factor (α). That is, they can be described by the following equation, which involves seven parameters:

$$\mathbf{X}_{\mathrm{TRF2}} = \Delta\mathbf{X} + \alpha\, \mathbf{R}_1[\theta_1]\, \mathbf{R}_2[\theta_2]\, \mathbf{R}_3[\theta_3]\, \mathbf{X}_{\mathrm{TRF1}} \tag{3.5}$$

[13] *Precession* and *nutation* series (defined for the mean equator and equinox J2000.0 in the CRF) have valid analytical expressions for long time intervals. But the rotation and orientation parameters of Earth cannot be modelled theoretically and must be periodically updated using observations.

where

$$\mathbf{R_1}[\theta] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}, \quad \mathbf{R}_2[\theta] = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$$
$$\mathbf{R}_3[\theta] = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

By adopting the convention used by IERS, equation (3.5) can be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\text{TRF2}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\text{TRF1}} + \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} + \begin{bmatrix} D & -R_3 & R_2 \\ R_3 & D & -R_1 \\ -R_2 & R_1 & D \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\text{TRF1}} \tag{3.7}$$

where T_1, T_2, T_3 are three translation parameters, D is a scale factor and R_1, R_2 and R_3 are three rotation angles.

Transformation parameters from ITRF2000 to past ITRFs are listed in table 4.1 of the IERS Conventions (2003) [McCarthy and Petit, 2004]. FORTRAN code `trnfsp3n.f` implementing this transformation and several parameter transformation files from the International GNSS Service (IGS) can be found in *ftp://igscb.jpl.nasa.gov/pub/resource/tutorial/APPENDIX_IGS_ITRF.txt*. See laboratory exercise 6 of session 3.3 in Volume II.

3.1.3 GNSS Reference Frames

Brief descriptions of the GPS WGS-84, Glonass PZ-90, Galileo GTRF and Beidou CGCS2000 reference frames are given below.

3.1.3.1 GPS Reference Frame WGS-84

From 1987, GPS has used the World Geodetic System WGS-84, developed by the US DoD, which is a unified terrestrial reference system for position and vector referencing.[14] Indeed, the GPS broadcast ephemerides are linked to the position of the satellite antenna phase centre in the WGS-84 reference frame. Thus, the user's receiver coordinates will be expressed in the same ECEF frame.

The initial implementation of WGS-84 was realised from a set of more than a thousand terrestrial sites, whose coordinates were derived from transit observations.[15] Successive refinements (which also led to some adjustments of the fundamental constants), using more accurate coordinates of the MS, approximate some ITRS realisations. For instance, realisations

[14] The document 'Modern Terrestrial Reference Systems PART 3: WGS 84 and ITRS' contains data and interesting references on WGS-84 and International Terrestrial Reference System (ITRS) (*http://www.ngs.noaa.gov/CORS/Articles/Reference-Systems-Part-3.pdf*).

[15] With an accuracy level of 1–2 m, while the accuracy of the ITRF reference stations is at the centimetre level.

WGS-84(G730)[16] and WGS-84(G873) correspond to ITRF92 and ITRF94, respectively. The refined frame WGS-84(G1150) was introduced in 2002, and agrees with ITRF2000 at the centimetre level.

The parameters of the WGS-84 ellipsoid are given in Table 3.2.

Table 3.2: Ellipsoidal parameters of WGS-84 (revised in 1997).

Ellipsoid		
Semi-major axis of the ellipse	a	$6\,378\,137.0\,\text{m}$
Flattening factor	f	$1/298.257\,223\,563$
Earth's angular velocity	ω_E	$7\,292\,115.0 \cdot 10^{-11}\,\text{rad/s}$
Gravitational constant	μ	$3\,986\,004.418 \cdot 10^{8}\,\text{m}^3/\text{s}^2$
Speed of light in vacuum	c	$2.997\,924\,58 \cdot 10^{8}\,\text{m/s}$

3.1.3.2 Glonass Reference Frame PZ-90

The Glonass broadcast ephemerides are given in the PZ-90 reference frame. As with WGS-84, this is an ECEF frame with a set of associated fundamental parameters (see table 3.3 from [GLONASS ICD, 2008]).

The determination of a set of parameters to transform the PZ-90 coordinates to ITRF97 was the target of the International Glonass Experiment (IGEX-98). [Boucher and Altamimi, 2001] present a review of the IGEX-98 experiment and, as a conclusion, they suggest the following transformation[17] from (x, y, z) in PZ-90 to (x', y', z') in WGS-84, to metre-level accuracy:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} -3\,\text{ppb} & -353\,\text{mas} & -4\,\text{mas} \\ 353\,\text{mas} & -3\,\text{ppb} & 19\,\text{mas} \\ 4\,\text{mas} & -19\,\text{mas} & -3\,\text{ppb} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0.07\,\text{m} \\ -0.0\,\text{m} \\ -0.77\,\text{m} \end{bmatrix} \quad (3.8)$$

Following the notation of equation (3.7), the previous transformation is defined by the following table of parameters:

T_1	T_2	T_3	D	R_1	R_2	R_3
(cm)			(ppb)	(mas)		
7	0	−77	−3	−19	−4	353

According to the Glonass modernisation plan, the ephemeris information implementing the PZ-90.02 reference system was updated on all operational Glonass satellites from 12:00 to 17:00 UTC, 20 September 2007. From this time on, the satellites have been broadcasting in PZ-90.02. This ECEF reference frame is an updated version of PZ-90, the closest one to ITRF2000.

[16] 'G' indicates that it has been exclusively obtained with GPS observations and '730' indicates the GPS week.

[17] Where mas stands for milliarcseconds ($1\,\text{mas} = 4.848\,136\,81 \cdot 10^{-9}\,\text{rad}$); and ppb stands for parts per billion ($1\,\text{ppb} = 10^{-9}$).

The transformation from PZ-90.02 to ITRF2000 contains only an origin shift vector, but no rotations or scale factor, as shown in the following equation [Revnivykh, 2007] (see laboratory exercise 6 of session 3.3 in Volume II):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\text{ITRF2000}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\text{PZ-90.02}} + \begin{bmatrix} -0.36\,\text{m} \\ 0.08\,\text{m} \\ 0.18\,\text{m} \end{bmatrix} \tag{3.9}$$

The parameters associated with PZ-90 and PZ-90.02 are given in Table 3.3 ([GLONASS ICD, 1998]; [GLONASS ICD, 2008]).

Table 3.3: Ellipsoidal parameters of PZ-90 and PZ-90.02.

Ellipsoid		
Semi-major axis of the ellipse	a	6 378 136.0 m
Flattening factor	f	1/298.257 839 303
Earth's angular velocity	ω_E	$7\,292\,115.0 \cdot 10^{-11}$ rad/s
Gravitational constant	μ	$3\,986\,004.4 \cdot 10^{8}\,\text{m}^3/\text{s}^2$
Speed of light in a vacuum	c	$2.997\,924\,58 \cdot 10^{8}$ m/s
Second zonal harmonic coefficient	J_2^0	$1\,082\,625.75 \cdot 10^{-9}$

3.1.3.3 Galileo Reference Frame (GTRF)

A consortium called the Galileo Geodetic Service Provider (GGSP) consisting of seven institutions under the leadership of GeoForschungsZentrum Potsdam, was in charge of building a prototype for the development of the GTRF and the establishment of a service with products and information for potential users.

The initial coordinates for the reference stations were provided using GPS observations, because the GTRF was already required by the time when the first Galileo signals were emitted during the In-Orbit Validation (IOV) phase. Subsequent GTRF versions will use both GPS and Galileo observations. Weekly solutions will be performed for the long-term maintenance of the GTRF.

The GTRF must be compatible with the latest ITRF to within a precision level of 3 cm (2σ), since its maintenance is one of the main functions of the GGSP, see [Gendt et al., 2011].

Connection to the ITRF is realised and validated by International GNSS Service (IGS) stations,[18] which are part of the ITRF, and especially by local ties to other geodetic techniques like satellite laser ranging and Very Long Baseline Interferometry (VLBI).

[18] The Galileo tracking stations (i.e. Galileo Sensor Stations (GSS)) form a sparse global network. Therefore, it is necessary to increase the density of the network with additional stations to get the highest possible precision and stability for the GTRF.

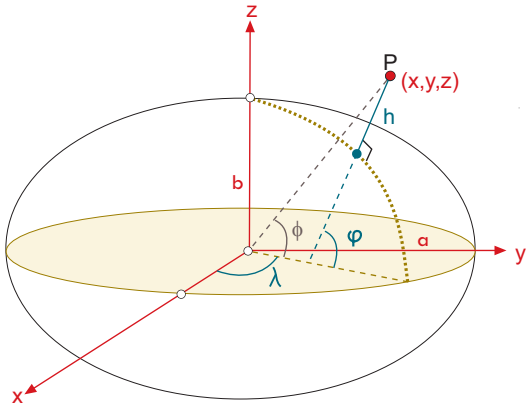


Figure 3.4: Cartesian (x, y, z) and ellipsoidal (φ, λ, h) coordinates. Geocentric latitude is denoted as ϕ.

As is the case for GPS and Glonass, the Galileo satellites broadcast their orbits in Galileo's own GTRF system.

3.1.3.4 Beidou Reference Frame (CGCS2000)

The Beidou system adopts CGCS2000, which is a geocentric coordinate system associated with Earth's ellipsoid given in Table 3.4. CGCS2000 is referred to ITRF97 at the epoch of 2000.0. The reference frame of CGCS2000 currently consists of the national GPS control network 2000 and the national astro-geodetic network after combined adjustment with the GPS network. The GNSS Continuously Operating Reference Stations (CORS) network will be the main part of China's geodetic infrastructure and a key technique for maintaining CGCS2000 [Cheng and Dang, 2011].

Table 3.4: Ellipsoidal parameters of CGCS2000.

Ellipsoid		
Semi-major axis of the ellipse	a	6 378 137.0 m
Flattening factor	f	1/298.257 222 101
Earth's angular velocity	ω_E	$7\,292\,115.0 \cdot 10^{-11}$ rad/s
Gravitational constant	μ	$3\,986\,004.418 \cdot 10^{8}$ m^3/s^2
Speed of light in a vacuum	c	$2.997\,924\,58 \cdot 10^{8}$ m/s

3.1.4 Cartesian and Ellipsoidal Coordinates

The (x, y, z) ECEF Cartesian coordinates of the above-mentioned terrestrial frames can be expressed in the associated ellipsoid as (φ, λ, h) ellipsoidal coordinates [19], where φ and λ are respectively the ellipsoidal latitude and longitude and h the height above the ellipsoid. Figure 3.4 illustrates the relation between Cartesian and ellipsoidal coordinates. The equations associated with this transformation are given in Appendix B.

This coordinate transformation is implemented in programs `car2geo.f` and `geo2car.f`. The input parameters are the ellipsoid semi-major axis a and ellipsoid flattening $f = 1 - b/a$.

[19] Also called *geodetic* or *geographic* coordinates.

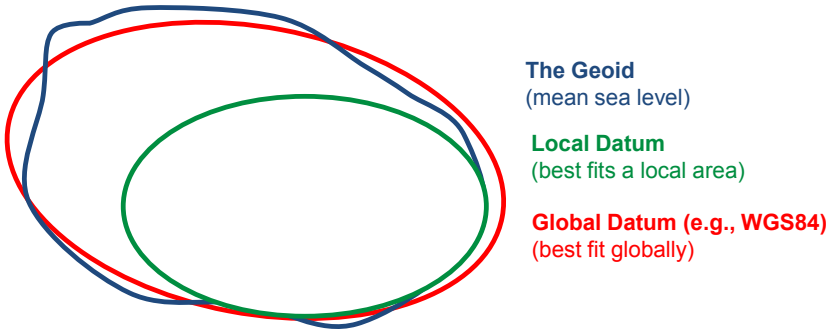


Figure 3.5: Global and local geodetic datums and the geoid.

3.1.5 Regional Datums and Map Projections

A datum consists of an ellipsoid relative to which the latitude and longitude of points are defined, and a geoid defining the surface at zero height.[20] Different organisations, countries or groups of countries have developed their own datums for civil and military applications.

These ellipsoids are conventionally defined as fitting the geoid in the region of interest, without being necessarily geocentric or having their axes constrained to a given orientation. A valuable contribution of GPS was to provide a Global Reference Frame, simplifying the mapping of Earth and unifying the diverse datums [Misra and Enge, 2001]. Figure 3.5 illustrates the concept of local and global datums, and their relation to the geoid.

The ellipsoidal coordinates of the datum must then be mapped to plane coordinates using a proper projection to build a map. Different projections have been defined (preserving areas, angles, distances, etc.), but all of them introduce distortions due to the mapping of a curved surface to a plane (see, for instance, [Hofmann-Wellenhof et al., 2008], chapter 8).

The local maps edited by the official organisations are referenced to a given datum. The coordinates expressed in one datum can be transformed into another with a seven-parameter transformation as explained in section 3.1.2.1 (i.e. a transformation in the space between two coordinate frames). In this way, most of the commercial receivers incorporate the parameters for several official datums to map the WGS-84 coordinates.

3.2 Satellite Orbits

Knowledge of the orbits and clocks of the satellites is fundamental for the right positioning. Any error in the GNSS satellite coordinates or satellite clock will affect the positioning accuracy. Information on orbital parameters and clocks is transmitted in the navigation message. Precise ephemerides and clocks are also provided by some organisations such as IGS.

In the following sections, the orbital elements are defined, the navigation message is introduced, and the algorithms for calculating the satellite coordinates from the GNSS ephemerides are detailed.

[20] More precisely, the equipotential surface of Earth's gravity field that on average coincides with mean sea level in the open oceans. Note that, due to variations in gravity, the geoid undulates significantly.

3.2.1 Keplerian Elements (Two-Body Problem)

According to Newtonian mechanics, the equation of motion of a mass m_2 relative to another mass m_1, considering only the central attractive force between them, is defined by the second-order homogeneous differential equation

$$\ddot{\mathbf{r}} + \frac{\mu}{r^3}\mathbf{r} = \mathbf{0} \tag{3.10}$$

where $\mathbf{r}$ is their relative position vector, $\mu = G(m_1 + m_2)$, and G is the universal gravitational constant. In the case of motion of an artificial Earth satellite, its mass can be neglected with respect to the mass of Earth (i.e. $\mu \simeq GM_E$).

Integration of this equation leads[21] to the Keplerian orbit of the satellite

$$\mathbf{r}(t) = \mathbf{r}(t; a, e, i, \Omega, \omega, \tau) \tag{3.11}$$

defined by the following six orbital parameters[22] (see Figs 3.6 and 3.7):

- Ω, the *right ascension of ascending node*, is the geocentric angle between the ascending node direction and the Aries point directions. Note that the intersection of the equatorial plane and orbital plane is called the nodal line. Its intersection with the unit sphere defines two points: the ascending node, through which the satellite crosses to the region of the positive Z-axis; and the descending node. Right ascension (Ω) is measured in a counterclockwise sense when viewed from the positive Z-axis.
- i, the *inclination of the orbital plane*, is the angle between the orbital plane and the equator.
- ω, the *argument of perigee*, is the angle between the ascending node and perigee directions, measured along the orbital plane. Perigee is the point of closest approach of the satellite to Earth's centre of mass. The most distant position is apogee. Both are in the direction of the orbital ellipse semi-major axis.
- a, the *semi-major axis of the orbital ellipse*, is the semi-major axis of the ellipse defining the orbit.
- e, the *numerical eccentricity of the orbit*, is the eccentricity of the orbital ellipse.
- T_0, the *perigee passing time*, is the time of satellite passage through the closest approach to Earth (perigee). The satellite's orbital position can be obtained at a given moment t using any one of the following 'anomalies', see Fig. 3.7:
  - $V(t)$, the *true anomaly*, is the geocentric angle between the perigee direction and the satellite direction. The sum of the true anomaly and the argument of perigee defines the *argument*

[21] The solution of this equation can be found in any textbook on celestial mechanics, or satellite orbital motion. See, for instance, [Seeber, 1993] or [Xu, 2007].

[22] We restrict ourselves to elliptic orbits.

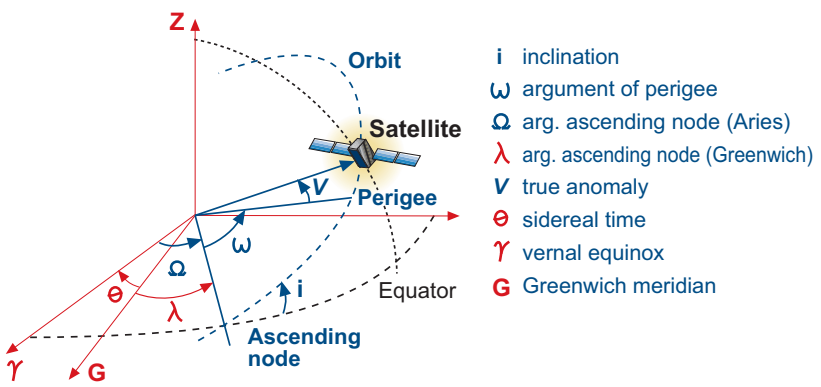


Figure 3.6: GNSS satellite orbital elements.

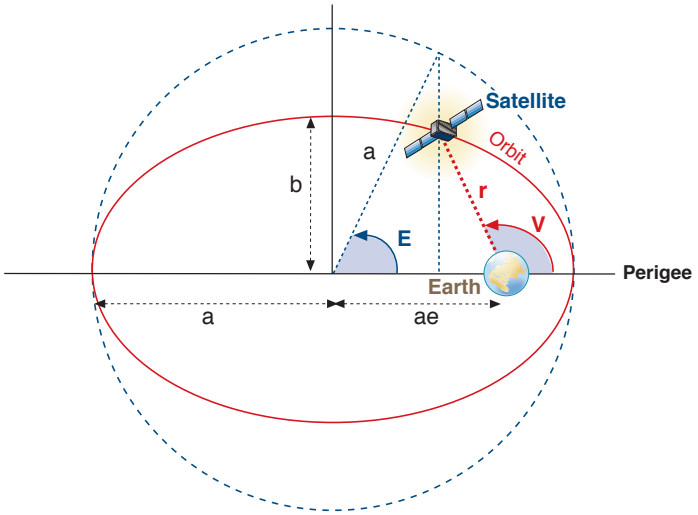


Figure 3.7: Elliptic orbit representation.

of latitude. Note that for a circular orbit ($e = 0$) the argument of perigee and the true anomaly are undefined. The satellite's position, however, can be specified by the argument of latitude.

– $E(t)$, the *eccentric anomaly*. A line that is normal to the major axis and passes through the satellite, it defines a point when it intersect a circle of radius a. Thus, the eccentric anomaly is the angle (measured from the centre of the orbit) between the perigee and such a point.

– $M(t)$, the *mean anomaly*, is a mathematical abstraction relating to mean angular motion.

The three anomalies are related by the formulae (see Fig. 3.7)

$$\begin{aligned} M(t) &= n(t - T_0) \\ E(t) &= M(t) + e \sin E(t) \\ V(t) &= 2 \arctan \left[\sqrt{(1+e)/(1-e)} \tan E(t)/2 \right] \\ n &= 2\pi/P = \sqrt{\mu/a^3} \end{aligned} \tag{3.12}$$

where n denotes the mean angular velocity of the satellite, or mean motion, with period of revolution P. Substituting the values for the semi-major axis $a = 26\,560$ km and gravitational constant $\mu = 3\,986\,004.418 \cdot 10^8\ \mathrm{m}^3/\mathrm{s}^2$ (see Table 3.2) into the last of the above equations (3.12), an orbital period of 12 sidereal hours is obtained for the GPS satellites.[23]

See Appendix C for the computation of the osculating orbital elements from the satellite's position and velocity, and vice versa. The equations are implemented in programs `rv2osc.f` and `osc2rv.f` in Volume II.

3.2.2 Perturbed Motion

The two-body problem considered in the previous section is only a first approximation to the real case. In practice, an additional set of accelerations or disturbing terms must be added to equation (3.10). They are mainly due to the following:

1. *Non-sphericity of Earth and non-homogeneous mass distribution*: As mentioned earlier, the shape of Earth can be approximated by an ellipsoid, with an equatorial radius about 20 km larger than the polar radius. On the other hand, the density of Earth is not uniform and the gravitational force depends on the latitude and longitude as well as the radial distance. Thus, Earth's potential[24] can be represented by a spherical harmonic expansion [Hofmann-Wellenhof et al., 2008] in the form

$$V = \frac{\mu}{r}\left[1 - \sum_{n=2}^{\infty}\left(\frac{a_E}{r}\right)^n J_n\, P_n(\sin\phi) + \sum_{n=2}^{\infty}\sum_{m=1}^{n}\left(\frac{a_E}{r}\right)^n (C_{nm}\cos m\lambda + S_{nm}\sin m\lambda)\, P_{nm}(\sin\phi)\right] \tag{3.13}$$

   where a_E is the semi-major axis of Earth, r is the geocentric distance of the satellite, and ϕ and λ are its geocentric latitude and longitude. The parameters $J_n = -C_{n0}$, C_{nm}, S_{nm} denote the zonal ($m = 0$) and tesseral ($m \neq 0$) coefficients of the harmonic development known from a model of Earth. P_n and P_{nm} are Legendre polynomials and the associated Legendre functions, respectively.

   The even-degree zonal coefficients produce primary secular variations of the orbital parameters, and the odd-degree zonal coefficients produce long-period perturbations. The tesseral coefficients produce short-periodic perturbations.

   The term with coefficient J_2 in equation (3.13), the harmonic coefficient of second order and degree 0, essentially models the ellipsoidal shape of Earth (with homogeneous density). It is smaller, by a factor of 10^4, than the acceleration due to a spherical Earth of uniform

[23] A sidereal day is $3^{\mathrm{m}}\,56^{\mathrm{s}}$ shorter than a solar day (see section 3.1.1).

[24] In equation (3.10), the term $-(\mu/r^3)\mathbf{r} = \nabla(\mu/r)$ corresponds to the acceleration from the potential produced by a spherical Earth with uniform density, see also equation (3.15).

density (i.e. $\nabla(\mu/r) = -(\mu/r^3)\mathbf{r}$). On the other hand, this term is approximately three orders of magnitude larger than the other coefficients (see Table 3.7). Including only this term in the potential, the previous expression becomes

$$V(r,\phi,\lambda) = \frac{\mu}{r}\left[1 + \frac{1}{2}\left(\frac{a_E}{r}\right)^2 J_2(1 - 3\sin^2\phi)\right] \tag{3.14}$$

The oblateness of Earth (J_2) has two effects. First, a torque which rotates the satellite's orbit in the equatorial plane produces a nodal regression $d\Omega/dt$. This is a secular effect (i.e. cumulative time) which depends upon the inclination of the orbit. It is zero for a polar orbit and maximum for an equatorial orbit. Table 3.5 shows the rate of change for the GPS, Beidou, Galileo and Glonass satellites. See [Hofmann-Wellenhof et al., 2008], equations (3.48).

Table 3.5: Secular precession of GNSS ascending node produced by J_2. The following values for orbital semi-major axis are assumed: $a_{GPS} = 26\,560$ km; $a_{Beidou_{MEO}} = 27\,906$ km; $a_{Galileo} = 29\,600$ km; $a_{Glonass} = 25\,480$ km. The excentricity is taken as $e = 0$.

	i	1 second	1 day	1 year
GPS	55.0°	$-4.5° \cdot 10^{-7}$	−0.039°	−14.16°
Beidou	55.0°	$-3.8° \cdot 10^{-7}$	−0.032°	−11.91°
Galileo	56.0°	$-3.0° \cdot 10^{-7}$	−0.026°	−9.44°
Glonass	64.8°	$-3.8° \cdot 10^{-7}$	−0.033°	−12.15°

The second effect of the non-central geopotential (J_2) produces a rotation of perigee $d\omega/dt$ (i.e. a rotation of the major axis in the orbital plane), see Table 3.6.

Table 3.6: Secular precession of GNSS perigee produced by J_2.

	i	1 second	1 day	1 year
GPS	55.0°	$2.5° \cdot 10^{-7}$	0.022°	7.96°
Beidou	55.0°	$2.1° \cdot 10^{-7}$	0.018°	6.69°
Galileo	56.0°	$1.5° \cdot 10^{-7}$	0.013°	4.76°
Glonass	64.8°	$-0.4° \cdot 10^{-7}$	−0.004°	−1.34°

This second harmonic of the geopotential (J_2) does not produce secular perturbations in the elements a, e and i. See the exercises of laboratory session 3.1 in Volume II.

2. *The presence of other celestial bodies – foremost, the Sun and Moon*: The gravitational field of the Sun and Moon act as perturbing forces on the satellites, the Moon being the body that produces the largest effects.[25] These gravitational forces also produce tides that deform the shape of Earth and affect its gravitational potential. Nevertheless, these tidal effects produce accelerations on GNSS satellites of the

[25] Note that, although the Sun is much more massive than the Moon, it is also much further away, and the gravitational effect is proportional to the inverse square of the distance.

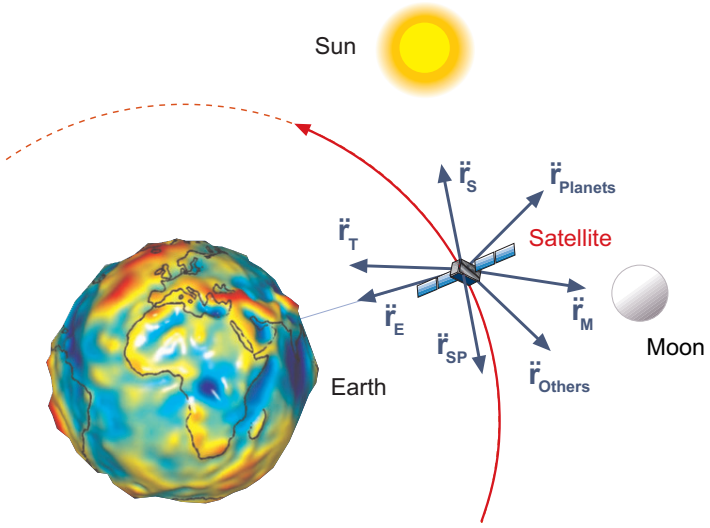


Figure 3.8: Different accelerations on the satellite orbit: Earth, non-spherical and non-homogeneous ($\ddot{r}_E$); celestial bodies ($\ddot{r}_S$, $\ddot{r}_M$, $\ddot{r}_{Planets}$); tidal effects ($\ddot{r}_T$); solar radiation pressure ($\ddot{r}_{SP}$); other sources ($\ddot{r}_{Others}$).

order of $10^{-9}\,\mathrm{m/s^2}$, which are three orders of magnitude lower than the lunar and solar gravitational accelerations (Table 3.7). See also the exercises of laboratory session 3.1.

3. *Solar radiation pressure*: This is produced by the photons impacting the surface of the satellite and mainly depends on the reflective properties and the area-to-mass ratio of the satellite. It also depends on an 'eclipse factor', which is zero when the satellite is in Earth's shadow. The effect of solar radiation pressure on a satellite's orbit is very difficult to model, and some contribution is estimated as stochastic acceleration parameters in the filter during estimation of the orbit.

Note that, although the gravitational effects of the Sun and Moon and the non-gravitational solar radiation pressure perturbing the accelerations appear very small, their accumulated effect can produce significant changes in GNSS satellite orbits over a long period of time (Fig. 3.8). Table 3.7 shows the different magnitudes of perturbation and their effect on GPS orbits.

Table 3.7: Different magnitudes of perturbation and their effects on GPS orbits (from [Seeber, 1993]).

Perturbation	Acceleration ($\mathrm{m/s^2}$)	Orbital effect	
		in 3 hours	in 3 days
Central force (as a reference)	0.56		
J_2	$5 \cdot 10^{-5}$	2 km	14 km
Rest of the harmonics	$3 \cdot 10^{-7}$	50–80 m	100–1500 m
Solar + Moon grav.	$5 \cdot 10^{-6}$	5–150 m	1000–3000 m
Tidal effects	$1 \cdot 10^{-9}$	–	0.5–1.0 m
Solar rad. pressure	$1 \cdot 10^{-7}$	5–10 m	100–800 m

3.2.3 GNSS Broadcast Orbits

As mentioned previously, the user's receiver computes the satellite coordinates from the information broadcast in the navigation messages by the GNSS satellites.

Two different approaches are followed by the GPS/Galileo/Beidou and Glonass satellites to account for satellite orbit perturbations. These approaches define what their messages contain.

In the case of the GPS, Galileo or Beidou satellites, the orbits are seen as Keplerian in a first approximation, and the perturbations are treated as temporal variations in the orbital elements (see comment below).

Indeed, an extended set of 16 quasi-Keplerian parameters (see Table 3.8) is broadcast to the user in the navigation message and regularly updated. This expanded set consists of the six orbital elements $(a(t), e(t), i(t), \Omega(t), \omega(t), M(t))$ and three rate parameters to account for the linear changes with time $(\dot{\Omega}, \dot{i}, \Delta n)$, three pairs of sinusoidal corrections (C_c, C_s) (i.e. $C_c \cos(2\phi)$, $C_s \sin(2\phi)$), and the reference ephemeris epoch t_{oe} (see section 3.3.1).

For Glonass satellites, the navigation message broadcasts the initial conditions of position and velocity $(\mathbf{r}_0, \mathbf{v}_0)$ and the vector components of the lunar and solar gravitational acceleration perturbations (see Table 3.9) to perform numerical integration of the orbit. The integration is based on applying a fourth-order Runge–Kutta method to the equation

$$\ddot{\mathbf{r}} = \nabla V + \mathbf{k}_{sun\_moon} \tag{3.15}$$

where V is the potential defined by equation (3.14) and $\mathbf{k}_{sun\_moon}$ are the lunar–solar accelerations expressed in an inertial coordinate system (see section 3.3.2). Note that, in the differential equation system (3.20) below, the term expressions are given as in [GLONASS ICD, 2008].

Comment: At any epoch the state of motion of the satellite is given by six parameters, namely the position and velocity vector components $(\mathbf{r}, \mathbf{v})$, or the six Keplerian elements $(a, e, i, \Omega, \omega, V)$; therefore, a 'point-to-point' transformation can be done between them. The orbital elements are the natural representations of the orbit, because (in the absence of perturbations) the motion along the orbit is described by a single parameter $(V(t))$, see Fig. 3.7. In the presence of perturbing forces, time-varying Keplerian elements defining an ellipse tangent to the orbit at any epoch can be considered, that is an osculating orbit.[26]

The transformations between position and velocity, and the osculating orbital elements, are given in Appendix C.

[26] From the Latin *osculor* (to kiss).

3.3 Computation of GNSS Satellite Coordinates

The algorithms for computing the satellite coordinates from the broadcast navigation messages of GPS, Glonass, Galileo and Beidou are provided as follows.

3.3.1 Computation of GPS, Galileo and Beidou Coordinates

Table 3.8 lists the GPS, Galileo or Beidou broadcast ephemeris parameters for computing the satellite coordinates at any observational epoch. These parameters are renewed periodically (typically every two hours for GPS) and must not be used after the prescribed time (about four hours), because the extrapolation error grows exponentially beyond this validity period.

The algorithm given below is from [GPS/SPS-SS, 1995], table 2-15. The Galileo and Beidou satellites follow the same scheme.

Table 3.8: GPS/Galileo/Beidou broadcast ephemeris and clock message parameters.

Parameter	Explanation
t_{oe}	Ephemerides reference epoch in seconds within the week
$\sqrt{a}$	Square root of semi-major axis
e	Eccentricity
M_o	Mean anomaly at reference epoch
ω	Argument of perigee
i_o	Inclination at reference epoch
Ω_0	Longitude of ascending node at the beginning of the week
Δn	Mean motion difference
$\dot{i}$	Rate of inclination angle
$\dot{\Omega}$	Rate of node's right ascension
c_{uc}, c_{us}	Latitude argument correction
c_{rc}, c_{rs}	Orbital radius correction
c_{ic}, c_{is}	Inclination correction
a_0	Satellite clock offset
a_1	Satellite clock drift
a_2	Satellite clock drift rate

In order to compute satellite coordinates from the navigation message, the algorithm here must be used (see, for instance, program `sub_orbit.f` in Volume II). An accuracy of about 5 m (RMS) is achieved for GPS satellites with S/A off and several tens of metres with S/A on[27] (see the laboratory exercises of session 3.2 in Volume II).

- Compute the time t_k from the ephemerides reference epoch t_{oe} (t and t_{oe} are expressed in seconds in the GPS week):

$$t_k = t - t_{oe}$$

  If $t_k > 302\,400$ s, subtract $604\,800$ s from t_k. If $t_k < -302\,400$ s, add $604\,800$ s.

[27] Actually, S/A was mainly applied to the satellite clocks and, apparently, not so often to the ephemerides.

- Compute the mean anomaly for t_k:

$$M_k = M_o + \left(\frac{\sqrt{\mu}}{\sqrt{a^3}} + \Delta n \right) t_k$$

- Solve (iteratively) the Kepler equation for the eccentric anomaly E_k:

$$M_k = E_k - e \sin E_k$$

- Compute the true anomaly v_k:

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

- Compute the argument of latitude u_k from the argument of perigee ω, true anomaly v_k and corrections c_{uc} and c_{us}:

$$u_k = \omega + v_k + c_{uc} \cos 2 (\omega + v_k) + c_{us} \sin 2 (\omega + v_k)$$

- Compute the radial distance r_k, considering corrections c_{rc} and c_{rs}:

$$r_k = a (1 - e \cos E_k) + c_{rc} \cos 2 (\omega + v_k) + c_{rs} \sin 2 (\omega + v_k)$$

- Compute the inclination i_k of the orbital plane from the inclination i_o at reference time t_{oe}, and corrections c_{ic} and c_{is}:

$$i_k = i_o + \dot{i}\, t_k + c_{ic} \cos 2 (\omega + v_k) + c_{is} \sin 2 (\omega + v_k)$$

- Compute the longitude of the ascending node λ_k (with respect to Greenwich). This calculation uses the right ascension at the beginning of the current week (Ω_o), the correction from the apparent sidereal time variation in Greenwich between the beginning of the week and reference time $t_k = t - t_{oe}$, and the change in longitude of the ascending node from the reference time t_{oe}:

$$\lambda_k = \Omega_o + \left(\dot{\Omega} - \omega_E \right) t_k - \omega_E t_{oe}$$

- Compute the coordinates in the TRS frame, applying three rotations (around u_k, i_k and λ_k):

$$\begin{bmatrix} X_k \\ Y_k \\ Z_k \end{bmatrix} = \mathbf{R}_3 (-\lambda_k)\, \mathbf{R}_1 (-i_k)\, \mathbf{R}_3 (-u_k) \begin{bmatrix} r_k \\ 0 \\ 0 \end{bmatrix}$$

  where $\mathbf{R}_1$ and $\mathbf{R_3}$ are the rotation matrices defined in equation (3.6).

3.3.2 Computation of Glonass Satellite Coordinates

The Glonass satellite coordinates are computed according to the specifications in the Glonass-ICD document. An accuracy level of about 3 m can be achieved using the algorithm provided by this ICD.

Table 3.9 lists the broadcast ephemeris parameters which are used to compute the Glonass satellite coordinates. Essentially, the ephemerides contain the initial conditions of position and velocity to perform numerical integration of the Glonass orbit within the interval of measurement $|t-t_e| < 15$ min. The accelerations due to solar and lunar gravitational perturbations are also given.

Table 3.9: Glonass broadcast ephemeris and clock message parameters.

Parameter	Explanation
t_e	Ephemerides reference epoch
$x(t_e)$	Coordinate at t_e in PZ-90
$y(t_e)$	Coordinate at t_e in PZ-90
$z(t_e)$	Coordinate at t_e in PZ-90
$v_x(t_e)$	Velocity component at t_e in PZ-90
$v_y(t_e)$	Velocity component at t_e in PZ-90
$v_z(t_e)$	Velocity component at t_e in PZ-90
$X''(t_e)$	Moon and Sun acceleration at t_e
$Y''(t_e)$	Moon and Sun acceleration at t_e
$Z''(t_e)$	Moon and Sun acceleration at t_e
$\tau_n(t_e)$	Satellite clock offset
$\gamma_n(t_e)$	Satellite relative frequency offset

In order to compute PZ-90 Glonass satellite coordinates from the navigation message, the following algorithm must be used [GLONASS ICD, 1998] (see program `GLOeph2sp3.f` and the laboratory exercises of session 3.3 in Volume II).

3.3.2.1 Computation Equations and Algorithm

1. *Coordinate transformation to an inertial reference frame*: The initial conditions $(x(t_e), y(t_e), z(t_e), v_x(t_e), v_y(t_e), v_z(t_e))$ broadcast in the Glonass navigation message, are in the ECEF Greenwich coordinate system PZ-90. Therefore, before orbit integration they must be transformed to an absolute (inertial) coordinate system using the following expressions.[28]

   Position:

$$\begin{aligned} x_a(t_e) &= x(t_e)\cos(\theta_{G_e}) - y(t_e)\sin(\theta_{G_e}) \\ y_a(t_e) &= x(t_e)\sin(\theta_{G_e}) + y(t_e)\cos(\theta_{G_e}) \\ z_a(t_e) &= z(t_e) \end{aligned} \tag{3.16}$$

[28] Note that, over small integration intervals, a simple rotation of the angle θ_{G_e} around the Z-axis is enough to perform this transformation. Nutation and precession of Earth and polar motion are very slow processes and will not introduce significant deviations over such short integration time intervals. See equation (3.2).

Velocity:

$$\begin{aligned} v_{x_a}(t_e) &= v_x(t_e)\cos(\theta_{G_e}) - v_y(t_e)\sin(\theta_{G_e}) - \omega_E\, y_a(t_e) \\ v_{y_a}(t_e) &= v_x(t_e)\sin(\theta_{G_e}) + v_y(t_e)\cos(\theta_{G_e}) + \omega_E\, x_a(t_e) \\ v_{z_a}(t_e) &= v_z(t_e) \end{aligned} \tag{3.17}$$

The $(X''(t_e), Y''(t_e), Z''(t_e))$ acceleration components broadcast in the navigation message are the projections of lunisolar accelerations to axes of the ECEF Greenwich coordinate system.[29] These accelerations must then be transformed to the inertial system by

$$\begin{aligned} (Jx_a m + Jx_a s) &= X''(t_e)\cos(\theta_{G_e}) - Y''(t_e)\sin(\theta_{G_e}) \\ (Jx_a m + Jx_a s) &= X''(t_e)\sin(\theta_{G_e}) + Y''(t_e)\cos(\theta_{G_e}) \\ (Jx_a m + Jx_a s) &= Z''(t_e) \end{aligned} \tag{3.18}$$

where θ_{G_e} is the sidereal time in the Greenwich meridian at epoch t_e, to which the initial conditions are referred (see program `sub_sid.f` in Volume II), that is

$$\theta_{G_e} = \theta_{G_0} + \omega_E(t_e - 3\,\mathrm{h}) \tag{3.19}$$

and where:

ω_E is Earth's rotation rate ($0.729\,211\,5 \cdot 10^{-4}$ rad/s)

θ_{G_0} is the sidereal time (in radians) in Greenwich at midnight GMT of a date for which the epoch t_e is specified, see equation (A.35). Note that t_e is in Glonass time (i.e. UTC(SU)+3 h).

2. *Numerical integration of the differential equations that describe the motion of the satellites*: According to the Glonass-ICD, recalculation of ephemerides from epoch t_e to epoch t_i within the interval of measurement ($|t_i - t_e| < 15$ min) is performed by numerical integration of the differential equations (3.20) describing the motion of the satellites. These equations are integrated in a direct absolute geocentric coordinate system OXa, OYa, OZa, connected with the current equator and vernal equinox, using the fourth-order Runge–Kutta technique given next:

$$\begin{cases} dx_a/dt = v_{x_a}(t) \\ dy_a/dt = v_{y_a}(t) \\ dz_a/dt = v_{z_a}(t) \\ dv_{x_a}/dt = -\bar{\mu}\,\bar{x}_a + \frac{3}{2}C_{20}\,\bar{\mu}\,\bar{x}_a\,\rho^2(1 - 5\bar{z}_a^2) + Jx_a m + Jx_a s \\ dv_{y_a}/dt = -\bar{\mu}\,\bar{y}_a + \frac{3}{2}C_{20}\,\bar{\mu}\,\bar{y}_a\,\rho^2(1 - 5\bar{z}_a^2) + Jy_a m + Jy_a s \\ dv_{z_a}/dt = -\bar{\mu}\,\bar{z}_a + \frac{3}{2}C_{20}\,\bar{\mu}\,\bar{z}_a\,\rho^2(3 - 5\bar{z}_a^2) + Jz_a m + Jz_a s \end{cases} \tag{3.20}$$

where

$\bar{\mu} = \mu/r^2$, $\bar{x_a} = x_a/r$, $\bar{y_a} = y_a/r$, $\bar{z_a} = x_a/r$

$\bar{\rho} = a_E/r$, $r = \sqrt{x_a^2 + y_a^2 + z_a^2}$

[29] These accelerations can also be computed using the analytical expressions given in [GLONASS ICD, 1998], after fixing some typographical errors; see subroutines `sub_Moon_pos_GLO.f` and `sub_Sun_pos_GLO.f` of program `GLOeph2sp3.f` in Volume II.

$a_E = 6\,378.136\,\text{km}$	Equatorial radius of Earth (PZ-90)
$\mu = 398\,600.44\,\text{km}^3/\text{s}^2$	Gravitational constant (PZ-90)
$C_{20} = -1082.63 \cdot 10^{-6}$	Second zonal coefficient of spherical harmonic expression

Note that, in the differential equations system (3.20), the term $C_{20} = -J_2 = +\sqrt{5}\bar{C}_{20}$ is used instead of J_2 in equations (3.14) and (3.15) to keep the same expressions as in the Glonass-ICD.

The right-hand side of the previous equation system (3.20) takes into account the accelerations determined by the central body gravitational constant μ, the second zonal coefficient C_{20} (which characterises the polar flattening of Earth) and the accelerations due to lunisolar gravitational perturbation. System (3.20) is implemented in program `sub_orb_deriv.f` in Volume II.

Runge–Kutta integration algorithm:

- Given the following initial value problem

$$\begin{cases} dy_1/dt = f_1(t, y_1, \ldots, y_n) \\ \qquad\vdots \\ dy_n/dt = f_1(t, y_1, \ldots, y_n) \end{cases} \iff \mathbf{Y}'(t) = \mathbf{F}(t, \mathbf{Y}(t)) \quad (3.21)$$

  $\mathbf{Y}(t_0) = [y_1(t_0), \ldots, y_n(t_0)]^T$, $\mathbf{Y}'(t_0) = [y_1'(t_0), \ldots, y_n'(t_0)]^T$
  it is desired to find $\mathbf{Y}(t_f)$ at some final time t_f, or $\mathbf{Y}(t_k)$ at some discrete list of points t_k (e.g. at tabulated intervals).

- The classical fourth-order Runge–Kutta method (RK4) is based on the following algorithm:

$$\begin{aligned} \mathbf{K}_1 &= \mathbf{F}(t_n, \mathbf{Y}_n) \\ \mathbf{K}_2 &= \mathbf{F}(t_n + h/2, \mathbf{Y}_n + h\,\mathbf{K}_1/2) \\ \mathbf{K}_3 &= \mathbf{F}(t_n + h/2, \mathbf{Y}_n + h\,\mathbf{K}_2/2) \\ \mathbf{K}_4 &= \mathbf{F}(t_n + h, \mathbf{Y}_n + h\,\mathbf{K}_3) \\ \mathbf{Y}_{n+1} &= \mathbf{Y}_n + \frac{h}{6}\,(\mathbf{K}_1 + 2\,\mathbf{K}_2 + 2\,\mathbf{K}_3 + \mathbf{K}_4) + O(h^5) \end{aligned} \quad (3.22)$$

  The method is initialised with the initial conditions $\mathbf{Y}(t_0)$ and $\mathbf{Y}'(t_0)$. For the numerical integration of Glonass satellite orbits, the function $\mathbf{F}(t, \mathbf{Y})$ is given by equations (3.20). The fourth-order Runge–Kutta algorithm is implemented in subroutine `sub_GlonassephRK.f` of program `GLOeph2sp3.f` in Volume II.

3. *Coordinate transformation back to the PZ-90 reference system*: The coordinates $(x(t), y(t), z(t))$, obtained from the numerical integration of the equations of motion, are transformed back to the Earth-fixed reference frame PZ-90 by the following equations:

$$\begin{aligned} x(t) &= x_a(t)\cos(\theta_G) + y_a(t)\sin(\theta_G) \\ y(t) &= -x_a(t)\sin(\theta_G) + y_a(t)\cos(\theta_G) \\ z(t) &= z_a(t) \end{aligned} \quad (3.23)$$

where θ_G is the sidereal time at the Greenwich meridian at time t, that is

$$\theta_G = \theta_{G_0} + \omega_E(t - 3\,h) \tag{3.24}$$

Note that t is in Glonass time, see equation (3.19).

It must be pointed out that the Glonass satellite coordinates are computed in the PZ-90 or PZ-90.02 reference system[30] instead of WGS-84 where the GPS coordinates have been calculated. To bring the PZ-90 coordinate system into coincidence with WGS-84 the transformation given by equation (3.8) must be applied (see section 3.1.3.2). The transformation from PZ-90.02 to WGS-84 (actually ITRF2000) is given by $\Delta x = -0.36\,\text{m}$, $\Delta y = +0.08\,\text{m}$, $\Delta z = +0.18\,\text{m}$, with no rotation, that is equation (3.9).

An example of the software code implementing the full algorithm for Glonass orbit computation, based on RK4, is provided in program `GLOeph2sp3.f`.

3.3.3 Computation of Precise GNSS Satellite Coordinates

Precise orbits and clocks for GPS and Glonass satellites can be found on the International GNSS Service (IGS) web server *http://igscb.jpl.nasa.gov*.[31] These products are in the public domain and free of charge to all users. They are available as American Standard Code for Information Interchange (ASCII) files, providing *precise orbits and clock files* with a sampling rate of 15 min, as well as *precise clock files* with a sample rate of 5 min and 30 s in SP3 format.[32]

Table 3.10 summarises the different products available and their accuracy, latency and sampling rate. It must be pointed out that IGS has adopted the ITRS as its reference system. Thus, IGS products are referred to this system.

Some centres also provide GPS satellite clocks with a 5 s sampling rate, like the files obtained from the Crustal Dynamics Data Information System (CDDIS) site.[33] Bear in mind that orbits and clock files must be consistent (i.e. they must come from the same centre). Therefore, Centre for Orbit Determination in Europe (CODE) clocks must be used with 'CODE' orbits, and not with 'JPL' or 'EMR' products, for instance.

The ANTenna EXchange format (ANTEX) files, together with other useful additional information on IGS antennas (see section 5.6), can be found at *http://www.epncb.oma.be/ftp/station/general*.

The satellite coordinates between epochs can be computed by polynomial interpolation. A 10th-order polynomial is enough for a centimetre level of accuracy with 15 min data. Stable clocks with a sampling rate of 30 s or higher can be interpolated with a first-order polynomial to a few

[30] PZ-90.02 was implemented on 20 September 2007.

[31] Other useful sites with data products and information are the CDDIS at *ftp://cddis.gsfc.nasa.gov/gps* and *ftp://cddis.gsfc.nasa.gov/glonass*, and *http://igscb.jpl.nasa.gov/components/compindex.html*.

[32] See the format files description (SP3, ANTEX, RINEX, etc.) in the HTML files associated with laboratory session 2.1 in Volume II.

[33] *ftp://cddis.gsfc.nasa.gov/pub/gps/products/*.

Table 3.10: IGS orbits and clock products: RMS accuracy, latency and sampling (see http://igscb.jpl.nasa.gov/components/prods.html).

Products	Broadcast	Ultra-rapid		Rapid	Final
		Predicted	Observed		
(delay)	(real time)	(real time)	(3–9 h)	(17–41 h)	(12–18 d)
Orbit GPS	∼100 cm	∼5 cm	∼3 cm	∼2.5cm	∼ 2.5 cm
(sampling)	(∼2 h)	(15 min)	(15 min)	(15 min)	(15 min)
Glonass					∼5 cm
(sampling)					(15 min)
Clock GPS	∼5 ns	∼3 ns	∼150 ps	∼75 ps	∼75 ps
(sampling)	(daily)	(15 min)	(15 min)	(5 min)	(30 s)
Glonass					∼ TBD
(sampling)					(15 min)

centimetres of accuracy (see exercise 7b of laboratory session 3.2 in Volume II). Clocks with a lower sampling rate should not be interpolated, because clocks evolve as random walk processes.

As an example of polynomial interpolation, the Lagrange method is presented as follows. Given a table of values (x_i, y_i), $i = 1, \ldots, n$, the interpolated value $y \simeq P_n(x)$ at a given x can be computed as

$$\begin{aligned} P_n(x) &= \sum_{i=1}^{n} y_i \frac{\prod_{j\neq i}(x - x_j)}{\prod_{j\neq i}(x_i - x_j)} \\ &= y_1 \frac{x - x_2}{x_1 - x_2} \cdots \frac{x - x_n}{x_1 - x_n} + \cdots \\ &\quad + y_i \frac{x - x_1}{x_i - x_1} \cdots \frac{x - x_{i-1}}{x_i - x_{i-1}} \frac{x - x_{i+1}}{x_i - x_{i+1}} \cdots \frac{x - x_n}{x_i - x_n} + \cdots \\ &\quad + y_n \frac{x - x_1}{x_n - x_1} \cdots \frac{x - x_{n-1}}{x_n - x_{n-1}} \end{aligned} \tag{3.25}$$

3.3.4 Computation of Coordinates from Almanac Data

Almanac data allow computation of a raw estimate of the satellite coordinates (1 – 2 km, 1σ error), which is needed for signal acquisition by the receiver.

The almanac data consist of a reduced set of parameters (with respect to broadcast ephemerides), as given in Table 3.11 for GPS satellites. Similar data are broadcast for Galileo and Beidou.

Thus, the same equations as in section 3.3.1 can be used to compute the coordinates from the almanac just by taking

$$\Delta n = \dot{i} = C_{uc} = C_{us} = C_{rc} = C_{rs} = C_{ic} = C_{is} = 0\,. \tag{3.26}$$

Table 3.11: GPS/Galileo/Beidou broadcast almanac parameters.

Parameter	Explanation
ID	Satellite identification number
WEEK	Current week
t_a	Reference epoch in seconds within the week
$\sqrt{a}$	Square root of semi-major axis
e	Eccentricity
M_o	Mean anomaly at reference epoch
ω	Argument of perigee
δ_i	Inclination offset relative to $i = 54°$ (GPS) or $i = 56°$ (Galileo) or $i = 54°$ (Beidou MEO) or $i = 0°$ (Beidou GEO)
Ω_0	Longitude of ascending node at beginning of week
$\dot{\Omega}$	Rate of node's right ascension

4. GNSS Measurements and Data Preprocessing

The basic GNSS observable is the travel time ΔT of the signal to propagate from the phase centre of the satellite antenna (the emission time) to the phase centre of the receiver antenna (the reception time). This value multiplied by the speed of light gives the *apparent*[1] range $R = c\,\Delta T$ between them.

As mentioned in section 2.2, the GNSS signals contain ranging codes to allow users to compute the travel time ΔT. Indeed, the receiver determines ΔT by correlating the received code (P) from the satellite with a replica of this code generated in the receiver, so this replica moves in time (ΔT) until the maximum correlation is obtained (see Fig. 4.1).

The measurement $R = c\Delta T$ is what is known as the *pseudorange*. It is called pseudorange, because it is an 'apparent range' between the satellite and the receiver which does not match its geometric distance because of, among other factors, synchronisation errors between receiver and satellite clocks. Taking explicitly into account possible synchronisation errors between such clocks, the travel time between transmission and reception is obtained as the difference in time measured on two different clocks or time scales: the satellite (t^{sat}) and the receiver (t_{rcv}). Thus, considering

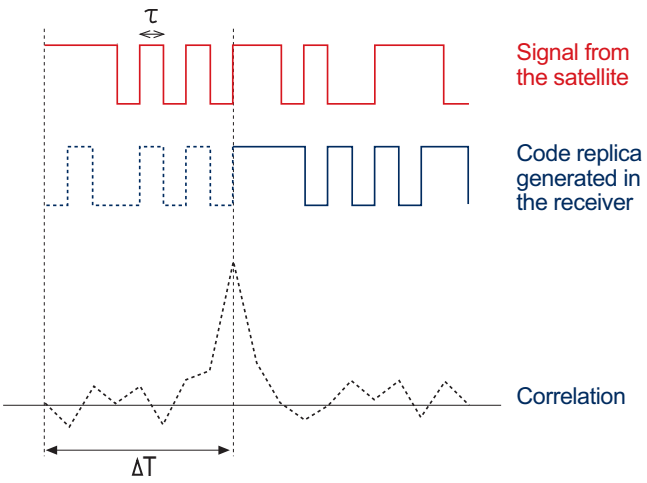


Figure 4.1: Determination of the signal travel time.

[1] It is called apparent, or pseudorange, to distinguish it from the true range, since different effects cause them to differ.

a reference time scale T (i.e. GNSS), the measured pseudorange (using the code P for the frequency signal f) for the satellite and receiver may be expressed as

$$R_{P_f} = c\,[t_{rcv}(T_2) - t^{sat}(T_1)] \tag{4.1}$$

where: c is the speed of light in a vacuum; $t_{rcv}(T_2)$ is the time of signal reception, measured on the time scale given by the receiver clock; and $t^{sat}(T_1)$ is the time of signal transmission, measured on the time scale given by the satellite clock.

The pseudorange R_{P_f} measurement obtained by the receiver using this procedure includes, besides the geometric range ρ between the receiver and the satellite and clock synchronisation errors, other terms due to signal propagation through the atmosphere (ionosphere and troposphere), relativistic effects, instrumental delays (of satellite and receiver), multipath and receiver noise (see Fig. 5.1 on page 96). Taking explicitly into account all these terms, the previous equation can be rewritten as follows, where R_{P_f} represents any GNSS code measurement at frequency f (from GPS, Glonass, Galileo or Beidou):

$$R_{P_f} = \rho + c(dt_{rcv} - dt^{sat}) + Tr + \alpha_f\, STEC + K_{P_f,rcv} - K^{sat}_{P_f} + \mathcal{M}_{P_f} + \varepsilon_{P_f} \tag{4.2}$$

Here:

- ρ is the geometric range between the satellite and receiver Antenna Phase Centres (APCs) at emission and reception time, respectively (section 5.1). Note: The APC is frequency dependent, but we neglect this effect here for simplicity (see section 5.6 and Volume II, Chapter 6).
- dt_{rcv} and dt^{sat} are the receiver and satellite clock offsets from the GNSS time scale, including the relativistic satellite clock correction (section 5.2).
- Tr is the tropospheric delay, which is non-dispersive (section 5.4.2).
- $\alpha_f\, STEC$ is a frequency-dependent ionospheric delay term, where α_f is the conversion factor between the integrated electron density along the ray path ($STEC$), and the signal delay at frequency f. That is, $\alpha_f = \frac{40.3}{f^2} 10^{16}$ $\mathrm{m}_{(\text{signal delay at frequency } f)}/\mathrm{TECU}$, where the frequency f is in Hz and 1 TECU$= 10^{16} \mathrm{e}^-/\mathrm{m}^2$. See equation (5.40), section 5.4.1.
- $K_{P_f,rcv}$ and $K_{P_f}{}^{sat}$ are the receiver and satellite instrumental delays, which are dependent on the code and frequency (sections 4.1 and 5.3).
- $\mathcal{M}_{P_f}$ represents the effect of *multipath*, also depending on the code type and frequency, and ε_{P_f} is the receiver noise (section 4.2).

Besides the code, the carrier phase itself is also used to obtain a measure of the apparent distance between satellite and receiver. These carrier phase measurements are much more precise than the code measurements (typically two orders of magnitude more precise), but they are ambiguous by an unknown integer number of wavelengths (λN). Indeed, this ambiguity changes arbitrarily every time the receiver loses the lock on the signal, producing jumps or range discontinuities.

The carrier phase measurements ($\Phi_{L_f} = \lambda_{L_f}\,\phi_{L_f}$) can be modelled as

$$\begin{aligned}\Phi_{L_f} = \;& \rho + c(dt_{rcv} - dt^{sat}) + Tr - \alpha_f\, STEC + k_{L_f,rcv} - k^{sat}_{L_f} + \\ +\; & \lambda_{L_f}\, N_{L_f} + \lambda_{L_f}\, w + m_{L_f} + \epsilon_{L_f}\end{aligned} \tag{4.3}$$

where this equation, besides the terms in equation (4.2), includes the *wind-up* ($\lambda_{L_f}\, w$) due to the circular polarisation of the electromagnetic signal[2] and the integer ambiguity N_{L_f} (see section 5.5 and Fig. 4.2 below). The terms $k_{L_f,rcv}$ and $k^{sat}_{L_f}$ are frequency dependent and correspond to carrier phase instrumental delays (see Fig. 6.4) associated with the receiver and satellite, respectively. The m_{L_f} and ϵ_{L_f} terms are the carrier phase multipath and noise, respectively.

Note that the ionospheric term has opposite signs for code and phase. This means that the ionosphere produces an advance in the carrier phase measurement equal to the delay on the code measurement (see section 5.4.1).

4.1 Combinations of GNSS Measurements

Starting from the basic observables as described previously, the following combinations can be defined (where R_i and Φ_i, $i = 1, 2$, indicate measurements in the frequencies f_1 and f_2 and P and L are omitted for simplicity):

- *Ionosphere-free combination*: This removes the first-order (up to 99.9%) ionospheric effect, which depends on the inverse square of the frequency ($\alpha_i \propto 1/f_i^2$, see section 5.4.1)

$$\Phi_{_C} = \frac{f_1^2\,\Phi_{_1} - f_2^2\,\Phi_{_2}}{f_1^2 - f_2^2}, \qquad R_{_C} = \frac{f_1^2\,R_{_1} - f_2^2\,R_{_2}}{f_1^2 - f_2^2} \tag{4.4}$$

  Satellite clocks are defined relative to the $R_{_C}$ combination, see section 4.1.1.

- *Geometry-free (or ionospheric) combination*: This cancels the geometric part of the measurement, leaving all the frequency-dependent effects (i.e. ionospheric refraction, instrumental delays, wind-up). It can be used to estimate the ionospheric electron content or to detect cycle slips in the carrier phase, as well. Note the change in the order of terms in $\Phi_{_I}$ and $R_{_I}$:

$$\Phi_{_I} = \Phi_{_1} - \Phi_{_2}, \qquad R_{_I} = R_{_2} - R_{_1} \tag{4.5}$$

- *Wide-laning combinations*: These combinations are used to create a measurement with a significantly wide wavelength. This longer wavelength is useful for carrier phase cycle-slip detection and fixing ambiguities:

$$\Phi_{_W} = \frac{f_1\,\Phi_{_1} - f_2\,\Phi_{_2}}{f_1 - f_2}, \qquad R_{_W} = \frac{f_1\,R_{_1} - f_2\,R_{_2}}{f_1 - f_2} \tag{4.6}$$

[2] A rotation of 360°of the receiver antenna, keeping its position fixed, would mean a variation of one wavelength in the phase-obtained measurement of the apparent distance between receiver and satellite (see section 5.5).

- *Narrow-laning combinations*: These combinations create measurements with a narrow wavelength. The measurement in this combination has a lower noise than each separate component:

$$\Phi_{N} = \frac{f_1\,\Phi_1 + f_2\,\Phi_2}{f_1 + f_2}, \qquad R_{N} = \frac{f_1\,R_1 + f_2\,R_2}{f_1 + f_2} \tag{4.7}$$

  Φ_{W} and R_{N} have the same ionospheric dependence, which is exploited by the Melbourne–Wübbena (MW) combination (see equations (4.19) and Fig. 4.10 below) to remove the ionospheric refraction.

4.1.1 Combining Pairs of Signals

Before applying the previous equations to the code and phase measurements, some rearrangement of expressions (4.2) and (4.3) will be done in order to refer the clocks to the code ionosphere-free combination of f_1 and f_2 frequencies, $R_{C_{12}}$ (other pairs of frequencies could be used as well). With this clock redefinition, the code instrumental delay will cancel in the ionosphere-free combination. Moreover, such delays will appear in terms of a differential code bias between both frequencies, always joining the ionospheric term, with a frequency-dependent coefficient (i.e. $\alpha_{(\cdot)}\,(I + K_{21})$), see equations (4.19).

4.1.1.1 Clock Redefinition and Differential Code Biases

By defining a new clock δt as

$$c\,\delta t = c\,dt + K_{C_{12}}, \quad \text{where} \quad K_{C_{12}} = \frac{f_1^2 K_1 - f_2^2 K_2}{f_1^2 - f_2^2} \tag{4.8}$$

it is not difficult to find that

$$\begin{aligned} c\,dt + K_1 &= c\,\delta t + \tilde{\alpha}_1(K_2 - K_1) \\ c\,dt + K_2 &= c\,\delta t + \tilde{\alpha}_2(K_2 - K_1) \end{aligned} \tag{4.9}$$

with

$$\tilde{\alpha}_i \equiv \frac{\alpha_i}{\alpha_2 - \alpha_1} \quad (i = 1, 2) \;\text{ and }\; \alpha_i = \frac{40.3}{f_i^2} 10^{16}\,\mathrm{m}_{\mathrm{delay(signal}\;\Phi_{f_i})}/\mathrm{TECU} \tag{4.10}$$

where dt, δt, K_1, K_2 stand for either the satellite or the receiver clocks, and their instrumental delays.

The term $K_2 - K_1$ is the *interfrequency bias* between the code instrumental delays at frequencies f_1 and f_2, also called the Differential Code Bias (DCB), which can be defined for both the receiver and satellite instrumental delays:

$$K_{21,rcv} \equiv K_{2,rcv} - K_{1,rcv}, \qquad K_{21}^{sat} \equiv K_2^{sat} - K_1^{sat} \tag{4.11}$$

In the same way, and following previous notation, the combined instrumental delay (receiver and satellite) can be written as

$$\begin{aligned} K_{21,rcv}^{sat} &= (K_{2,rcv} - K_{1,rcv}) - (K_2^{sat} - K_1^{sat}) = K_{21,rcv} - K_{21}^{sat} \\ &= (K_{2,rcv} - K_2^{sat}) - (K_{1,rcv} - K_1^{sat}) = K_{2,rcv}^{sat} - K_{1,rcv}^{sat} \end{aligned} \tag{4.12}$$

To simplify the notation, hereafter we will use

$$K_{21} \equiv K^{sat}_{21,rcv} = K_{21,rcv} - K^{sat}_{21} = K^{sat}_{2,rcv} - K^{sat}_{1,rcv} \tag{4.13}$$

and

$$\begin{aligned} K_i &\equiv K_{i,rcv} - K^{sat}_i \\ k_i &\equiv k_{i,rcv} - k^{sat}_i \end{aligned} \tag{4.14}$$

for the code and carrier instrumental delays corresponding to frequency f_i.

From expressions (4.11) to (4.14), it follows that whatever is applied to the individual receiver or satellite's DCB, or the combined one, it is always

$$K_{21} = K_2 - K_1 \tag{4.15}$$

The DCBs of the satellites are broadcast in the GNSS navigation messages as the Total Group Delay (TGD), see equation (5.20). For instance, TGD $= -\tilde{\alpha}_1 K^{sat}_{21}$ for the GPS P1 code pseudorange measurements (i.e. $R_{P_{f_1}}$) is broadcast in the legacy GPS message for the P2–P1 code inter-frequency bias (see section 2.2.1.1), and the DCBs between the codes on frequencies $E1, E5a$ and frequencies $E1, E5b$ are broadcast in the Galileo F/NAV and I/NAV messages, respectively (see section 2.2.3.1). IGS also provides DCBs for the GPS satellites and station receivers. These products match the Global Ionospheric Maps (GIM)[3] and the precise satellite orbits and clocks (see section 3.3.3), which always refer to the code ionosphere-free combination $R_{C_{12}}$ of the GPS P1 and P2 codes.

4.1.1.2 Rewriting the Equations

Using the previous results, the code and carrier measurement equations can be referred to the new clock δt defined in equation (4.8), in such a way that the DCBs cancel in the ionosphere-free combination of code measurements.

Thus, taking I as the ionospheric delay (units of $\mathrm{m}_{\mathrm{delay}_{(\Phi_1-\Phi_2)}}$) in the geometry-free combination,

$$I \equiv (\alpha_2 - \alpha_1)\,\mathrm{STEC} = \frac{40.3(f_1^2 - f_2^2)}{f_1^2 f_2^2} 10^{16}\,\mathrm{STEC} \tag{4.16}$$

and substituting expressions (4.9) into (4.2) and (4.3), it follows that

$$R_i = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr + \tilde{\alpha}_i(I + K_{21}) + \mathcal{M}_i + \varepsilon_i \tag{4.17}$$

$$\Phi_i = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr - \tilde{\alpha}_i(I + K_{21}) + b_i + \lambda_i N_i + \lambda_i w + m_i + \epsilon_i \tag{4.18}$$

where $b_i = k_i - K_i + 2\tilde{\alpha}_i K_{21}$ is a frequency-dependent bias[4] $(i = 1, 2)$.

Comment: It may be surprising to see the 'code DCBs' (i.e. K_{21}) included in the equation of carrier phase measurements joining the ionospheric term, but see the first remark on page 72.

[3] See [Hernández-Pajares, 2004] and *ftp://cddis.gsfc.nasa.gov/gps/products/ionex*. Note that the IONEX DCBs are related to the TGD broadcast in the GPS message (associated with the P1 code) by TGD $\rightarrow -\tilde{\alpha}_1\,\mathrm{DCB}_{\mathrm{IONEX}}$. On the other hand, such DCBs can be aligned to a different reference, and a global bias can appear. See exercise 3 in laboratory session 5.3 in Volume II.

[4] Note that, as with the code instrumental delay, the carrier phase bias b can be split into two different terms associated with the receiver and the satellite. That is, $b_i = b_{i,rcv} - b^{sat}_i$.

4.1.1.2.1 Combinations of Measurements Written in Closed Form

By replacing the expressions for R_i and Φ_i, $i = 1, 2$, in definitions (4.4) to (4.7), the following expressions can be found (the derivations are left as an exercise). Remark: the APC effect is neglected here for simplicity (see section 5.6 and Volume II, exercise 2 in session 6.1).

Input measurements R_i and Φ_i ($i = 1, 2$):

$$R_i = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr + \tilde{\alpha}_i(I + K_{21}) + \mathcal{M}_i + \varepsilon_i$$
$$\Phi_i = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr - \tilde{\alpha}_i(I + K_{21}) + B_i + \lambda_i w + m_i + \epsilon_i$$

where the ambiguity B_i is given by

$$B_i = b_i + \lambda_i N_i,\ \lambda_i = c/f_i,\ \tilde{\alpha}_1 = 1/(\gamma_{12} - 1),\ \tilde{\alpha}_2 = \gamma_{12}\tilde{\alpha}_1 = 1 + \tilde{\alpha}_1,$$
$$\gamma_{12} = (f_1/f_2)^2$$

with the bias b_i a real number and N_i an integer ambiguity.
Note that $K_{21} = K_{21,rcv} - K_{21}^{sat},\ b_i = b_{i,rcv} - b_i^{sat}$.

Ionosphere-free combination:

$$R_C = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr + \mathcal{M}_C + \varepsilon_C$$
$$\Phi_C = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr + B_C + \lambda_N w + m_C + \epsilon_C$$

where the bias B_C is given by

$$B_C = b_C + \lambda_N\,(N_1 + (\lambda_W/\lambda_2)N_W)$$

Geometry-free combination:

$$R_I = I + K_{21} + \mathcal{M}_I + \varepsilon_I$$
$$\Phi_I = I + K_{21} + B_I + (\lambda_1 - \lambda_2)w + m_I + \epsilon_I$$

where the bias B_I is given by

$$B_I = b_I + \lambda_1\,N_1 - \lambda_2\,N_2$$

Wide-lane (phase) and narrow-lane (code) combinations:

$$R_N = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr + \tilde{\alpha}_W(I + K_{21}) + \mathcal{M}_N + \varepsilon_N$$
$$\Phi_W = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr + \tilde{\alpha}_W(I + K_{21}) + B_W + m_W + \epsilon_W$$

where the bias B_W is given by

$$B_W = b_W + \lambda_W N_W$$

Other combinations involving code and phase measurements:

The Melbourne–Wübbena combination

$$\Phi_W - R_N = b_W + \lambda_W N_W + \mathcal{M}_{MW} + \varepsilon_{MW}$$

The GRAPHIC (Group and Phase Ionospheric Calibration) combination

$$\tfrac{1}{2}(R_i + \Phi_i) = \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr + \tfrac{1}{2}B_i + \tfrac{1}{2}\lambda_i w + \mathcal{M}_G + \varepsilon_G$$

Definitions and relationships (where $(\cdot)_X \equiv (\cdot)_{X_{12}}$) :

$$N_W \equiv N_1 - N_2,$$
$$\lambda_W \equiv c/(f_1 - f_2),\quad \lambda_N \equiv c/(f_1 + f_2),$$
$$\tilde{\alpha}_W \equiv \sqrt{\tilde{\alpha}_1\,\tilde{\alpha}_2} = f_1 f_2/(f_1^2 - f_2^2) = \sqrt{\gamma_{12}}/(\gamma_{12} - 1),\quad \gamma_{12} = (f_1/f_2)^2,$$
$$b_W \equiv (f_1 b_1 - f_2 b_2)/(f_1 - f_2),\quad b_C \equiv (f_1^2 b_1 - f_2^2 b_2)/(f_1^2 - f_2^2),$$
$$b_I \equiv b_1 - b_2,\quad b_W - b_C = \tilde{\alpha}_W\,b_I,$$

the same expressions for B_X as for b_X. (4.19)

The effect of a jump in the integer ambiguities in terms of ΔN_1, ΔN_2 and N_W is given next:

$\Delta\Phi_W$, $\Delta\Phi_I$, $\Delta\Phi_C$ variations

$$\begin{aligned}
\Delta\Phi_W &= \lambda_W \Delta N_W = \lambda_W \left(\Delta N_1 - \Delta N_2\right) \\
\Delta\Phi_I &= \lambda_1 \Delta N_1 - \lambda_2 \Delta N_2 = \left(\lambda_2 - \lambda_1\right) \Delta N_1 + \lambda_2 \, \Delta N_W \\
\Delta\Phi_C &= \lambda_N \left(\tfrac{\lambda_W}{\lambda_1} \Delta N_1 - \tfrac{\lambda_W}{\lambda_2} \Delta N_2\right) = \lambda_N \left(\Delta N_1 + \tfrac{\lambda_W}{\lambda_2} \Delta N_W\right)
\end{aligned} \tag{4.20}$$

The different wavelengths for the wide- and narrow-lane combinations of frequencies in GPS, Glonass and Galileo, as well as the values for the associated parameters, are given in Tables 4.1 and 4.2.

Table 4.1: Wide- and narrow-lane combinations of signals for different frequencies of GPS, Glonass (only the channel $k = 0$ is given for G1 and G2 signals) and Galileo. The Galileo E5 and E6 signals have not been included to simplify the table.

System	Signal	Frequency (MHz)	Wavelength λ_i (m)	Signals combined	Wide lane λ_W (m)	Narrow lane λ_N (m)
	i	f_i	$\lambda_i = c/f_i$	ij	$c/(f_i - f_j)$	$c/(f_i + f_j)$
GPS	L1	1575.420	$\lambda_{L1} = 0.190$	L1,L2	0.862	0.107
	L2	1227.600	$\lambda_{L2} = 0.244$	L1,L5	0.751	0.109
	L5	1176.450	$\lambda_{L5} = 0.255$	L2,L5	5.861	0.125
Glonass	G1	1602.000	$\lambda_{G1} = 0.187$	G1,G2	0.842	0.105
	G2	1246.000	$\lambda_{G2} = 0.241$	G1,G3	0.750	0.107
	G3	1202.025	$\lambda_{G3} = 0.249$	G2,G3	6.817	0.122
Galileo	E1	1575.420	$\lambda_{E1} = 0.190$	E1,E5b	0.814	0.108
	E5b	1207.140	$\lambda_{E5b} = 0.248$	E1,E5a	0.751	0.109
	E5a	1176.450	$\lambda_{E5a} = 0.255$	E5b,E5a	9.768	0.126

Table 4.2: Values for the different parameters related to the combinations of GPS, Glonass and Galileo measurements shown in Table 4.1.

System	Frequency f_i (MHz)	$\alpha_i = \frac{40.3}{f_i^2} 10^{16}$ ($m_{delay}/TECU$)	i,j	$\gamma_{ij} = (f_i/f_j)^2$
GPS	$f_{L1} = 154 \times 10.23$	$\alpha_{L1} = 0.1624$	L1,L2	$(77/60)^2$
	$f_{L2} = 120 \times 10.23$	$\alpha_{L2} = 0.2674$	L1,L5	$(154/115)^2$
	$f_{L5} = 115 \times 10.23$	$\alpha_{L5} = 0.2912$	L2,L5	$(24/23)^2$
Glonass	$f_{G1} = 2848 \times \frac{9}{16}$	$\alpha_{G1} = 0.1570$	G1,G2	$(9/7)^2$
	$f_{G2} = 2848 \times \frac{7}{16}$	$\alpha_{G2} = 0.2596$	–	–
	$f_{G3} = 117.5 \times 10.23$	$\alpha_{G3} = 0.2789$	–	–
Galileo	$f_{E1} = 154 \times 10.23$	$\alpha_{E1} = 0.1624$	E1,E5b	$(77/59)^2$
	$f_{E5b} = 118 \times 10.23$	$\alpha_{E5b} = 0.2766$	E1,E5a	$(154/115)^2$
	$f_{E5a} = 115 \times 10.23$	$\alpha_{E5a} = 0.2912$	E5b,E5a	$(118/115)^2$

Remarks on the previous equations:

- The code DCBs have been included in the equations of carrier phase measurements,[5] joining the ionospheric term to provide closed expressions for the different combinations of measurements, namely equations (4.4) to (4.7), where the ionospheric term always joins the DCB as '$I + K_{21}$', with a frequency-dependent scale factor, see equations (4.19). In this way, combinations removing the ionosphere, such as MW, will also be free from DCBs. Note that, as the carrier measurement contains an unknown bias $B_{(\cdot)}$, it can be redefined without loss of generality to include such DCBs. That is, $\tilde{\alpha}_{(\cdot)}(I + K_{21}) + B_{(\cdot)} \equiv \tilde{\alpha}_{(\cdot)} I + \mathcal{B}_{(\cdot)}$.
- Although, for simplicity, the previous equations (4.19) are written in terms of two signals at frequencies f_1 and f_2, they are valid for any pair of frequencies f_k and f_m (see Table 4.1).
- Note that the wind-up term does not appear in the wide-lane carrier phase combination. This is because the signals are subtracted in cycles, see equation (4.3). That is, $\Phi_W = (f_1\Phi_1 - f_2\Phi_2)/(f_1 - f_2) = c(\phi_1 - \phi_2)/(f_1 - f_2)$, and both signals are affected by the same fraction of cycle by the wind-up.
- The GRAPHIC combination [Yunck, 1993] provides an ionosphere-free single-frequency measurement with reduced noise (half the code noise, see Fig. 4.3 below), but contains the unknown ambiguity of the carrier phase. This combination is used, for example, for GPS single-frequency orbit determination for Low Earth Orbit (LEO) satellites; see, for instance, [Montenbruck and Ramos-Bosch, 2007].

4.1.1.3 Additional Comments on the Equations

A brief discussion on some issues related to the previous equations is provided to clarify their content and meaning. Several of these issues will be discussed in more detail in subsequent chapters.

4.1.1.3.1 Code-Based Positioning

The GNSS satellites broadcast in their navigation messages ephemeris and satellite clock information linked to the code ionosphere-free combination (R_{PC}). In this way, dual-frequency users can navigate the ionosphere-free combination without needing either ionospheric corrections or DCBs. In the case of the legacy GPS navigation message, such DCBs correspond to the P1–P2 codes (see Table 2.2).

For single-frequency users, the navigation message provides the parameters of an ionospheric model together with the associated TGDs, see section 5.3.

[5] Bear in mind that the clock has been redefined to include this 'code interfrequency bias', and the clock is common for both code and carrier measurements.

4.1.1.3.2 Carrier-Phase-Based Positioning

Due to the (satellite and receiver) clock redefinition (from expression (4.8)), the ionosphere-free combination of code measurements is free of DCBs (either for the satellite or receivers), but this is not the case for the carrier phase measurements, which still contain the real bias $b_C = b_{C,rec} + b_C^{sat}$ because the clocks have been referred to the ionosphere-free combination of 'code measurements'.

Nevertheless, this carrier phase bias b_C is not a problem for carrier-phase-based positioning techniques like Precise Point Positioning (PPP) presented in section 6.2, because it is assimilated into the unknown ambiguity when it is estimated by the Kalman filter as a real number (i.e. 'floating' the ambiguities). This does not affect the ambiguity fixing in differential positioning techniques like Real-Time Kinematics (RTK), because any remaining bias cancels out when forming the double differences between satellites and receivers $\nabla\Delta$, keeping only the integer part of the ambiguities ($\nabla\Delta B_i = \nabla\Delta N_i$, $i = 1, 2$), see section 6.3.1. Nevertheless, this bias (the fractional part of the ambiguities) must be considered if the ambiguities are tried to be fixed in undifferenced mode, i.e. *undifferenced ambiguity fixing*, see section 6.3.2.

4.1.1.3.3 Ionosphere and DCB Estimation

The redefinition of the satellite and receiver clocks allowed expressions (4.19) to be arranged in such a way that the DCBs join the ionospheric refraction in all the equations, with the same frequency-dependent factor.

The carrier phase bias B_I in the geometry-free combination can be estimated directly from the geometry-free combination of code measurements. Afterwards, the sum $I + K_{21}$ can be computed. Because this estimate is based on the code measurements, its accuracy will depend on the level of code noise and multipath. Nevertheless, geodesy can dramatically improve the accuracy of such estimates, due to the relationship between the ambiguities B_I, B_W and B_C. Indeed, B_C can be accurately estimated from geodetic positioning and B_W can be easily fixed using the MW combination. Hence, B_I can be computed from

$$B_W - B_C = \tilde{\alpha}_W B_I \tag{4.21}$$

Once the sum $I + K_{21}$ is computed, the DCBs (K_{21}) can be decorrelated from the ionospheric refraction (I) using a geometric description of the ionosphere (i.e. describing the satellite–receiver ray ionospheric sounding), due to the motion of the satellite along its path. In this way, the accuracy of the ionosphere and DCB determinations mostly relies on the geometric model used to describe the ionosphere (i.e. the one-layer model or two-layer voxel model [HJS and Colombo, 1999]).

4.1.2 Combining Trios of Signals

The previous expressions (4.19) have been referred to the new clock δt defined by equation (4.8), which is linked to the ionosphere-free combination of code measurements in the frequencies f_1 and f_2. As a result of this clock redefinition, the instrumental delay appears as the DCB (K_{21}), which cancels out when forming the ionosphere-free combination of codes in both frequencies.

Next, the code and carrier measurements of a third frequency will be rearranged and expressed in terms of this clock δt (relative to $R_{C_{12}}$).

4.1.2.1 Referring Measurements at a Third Frequency to the New Clock

When considering a third frequency f_3, a new instrumental delay K_3 appears, which leads to the following expressions when referring equations (4.2) and (4.3) to the clock δt:

$$\begin{aligned} R_3 &= \rho + c\,(\delta t_{rcv} - \delta t^{sat}) + T + \tilde{\alpha}_3(I + \tilde{K}_{31}) + \tilde{\alpha}_1 K_{21} + \mathcal{M}_3 + \varepsilon_3 \\ \Phi_3 &= \rho + c(\delta t_{rcv} - \delta t^{sat}) + T - \tilde{\alpha}_3(I + \tilde{K}_{31}) - \tilde{\alpha}_1 K_{21} + \\ &\quad + B_3 + \lambda_3 w + m_3 + \epsilon_3 \end{aligned} \tag{4.22}$$

where the ambiguity B_3 is given by

$$B_3 = b_3 + \lambda_3 N_3\,,\ \lambda_3 = c/f_3\,,\ \alpha_3 = 40.3/f_3^2\,,\ \tilde{\alpha}_3 = \alpha_3/(\alpha_2 - \alpha_1)\,,$$

$$\tilde{K}_{31} = (1/\tilde{\alpha}_3)(K_3 - K_1)\,,\ \ b_3 = k_3 - K_1 + \tilde{\alpha}_3\,\tilde{K}_{31} + 2\tilde{\alpha}_1 K_{21}\,,\quad N \in \mathbb{Z}$$

Note that, on comparing expression (4.22) with those of Φ_i and R_i, at the top of equations (4.19), the K_{21} DCB appears among the $\tilde{K}_{31}$ ones. This is because the clock is referred to the ionosphere-free combination of codes at frequencies f_1 and f_2.

4.1.2.2 New Combinations Involving Trios of Signals

New combinations of measurements can be built from three-frequency signals, as follows. Some examples are analysed in the exercises of session 4.2 in Volume-II.

- *The geometry-free and first-order ionosphere-free combination*: This combination would allow measurement of the second-order ionospheric effect (see section 5.4.1) with well-calibrated antenna phase centres. Nevertheless, the measurement noise is of the same order as or higher than this effect.

- *The first-order and second-order ionosphere-free combination*: This combination would allow removal of the ionosphere up to the millimetre level, but, again, the measurement noise is greatly amplified.

4.2 Measurement Features and Noise

In the previous sections we described the GNSS receiver measurements and their combinations. Now we are going to analyse the main features of these data, their measurement noise, and how it affects the combinations.

Two main features of code and carrier phase measurements are depicted in Fig. 4.2. While the code measurements (in green) show a noisy but continuous pattern, the carrier phase track (in blue) is thinner and broken by sudden jumps. The magnitude of such jumps is always unknown; they happen every time the receiver recovers carrier phase lock after a loss of signal. This behaviour can be summarised in the following two concepts:

- Code measurements are noisy but unambiguous.
- Carrier measurements are precise but ambiguous.

The previous example allows us to introduce the distinction between two different concepts that are often synonymous: namely, precision and accuracy.

Precision refers to the dispersion of a measurement around a mean value (not necessarily the true one). Accuracy refers to the closeness of the measurement to the true value. In this way, note that a measurement like the carrier phase can be very precise (i.e. at a very low noise level), but completely wrong due to the unknown bias, which can reach thousands of kilometres (in the figure the carrier phase has been shifted to fit into the plot frame).

The code and carrier phase measurement noise depends on different factors like the signal power, the method used for the analog-to-digital conversion, the correlation process, the design of the antenna, etc. These errors are independent of each receiver and each signal, and cannot be removed by differential techniques (differences between receivers or satellites) or by combinations of measurements.

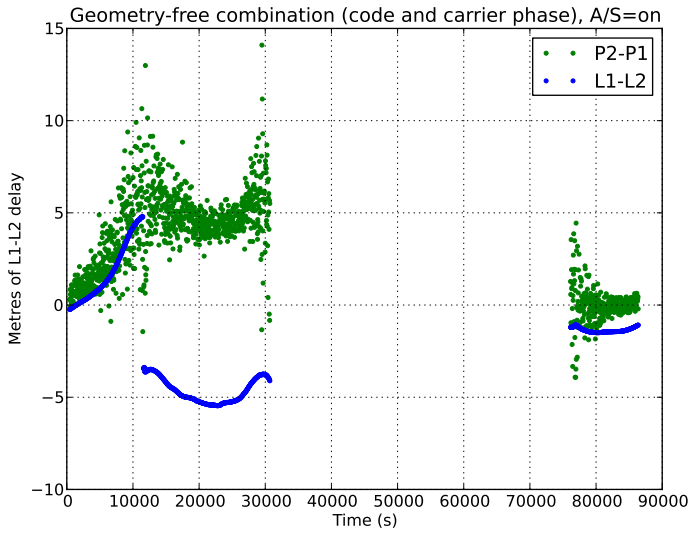


Figure 4.2: GPS code and carrier phase measurement features. The geometry-free combinations of code ($R_{P_2} - R_{P_1}$), in green, and carrier ($\Phi_{L_1} - \Phi_{L_2}$), in blue, are plotted as function of time for a given satellite.

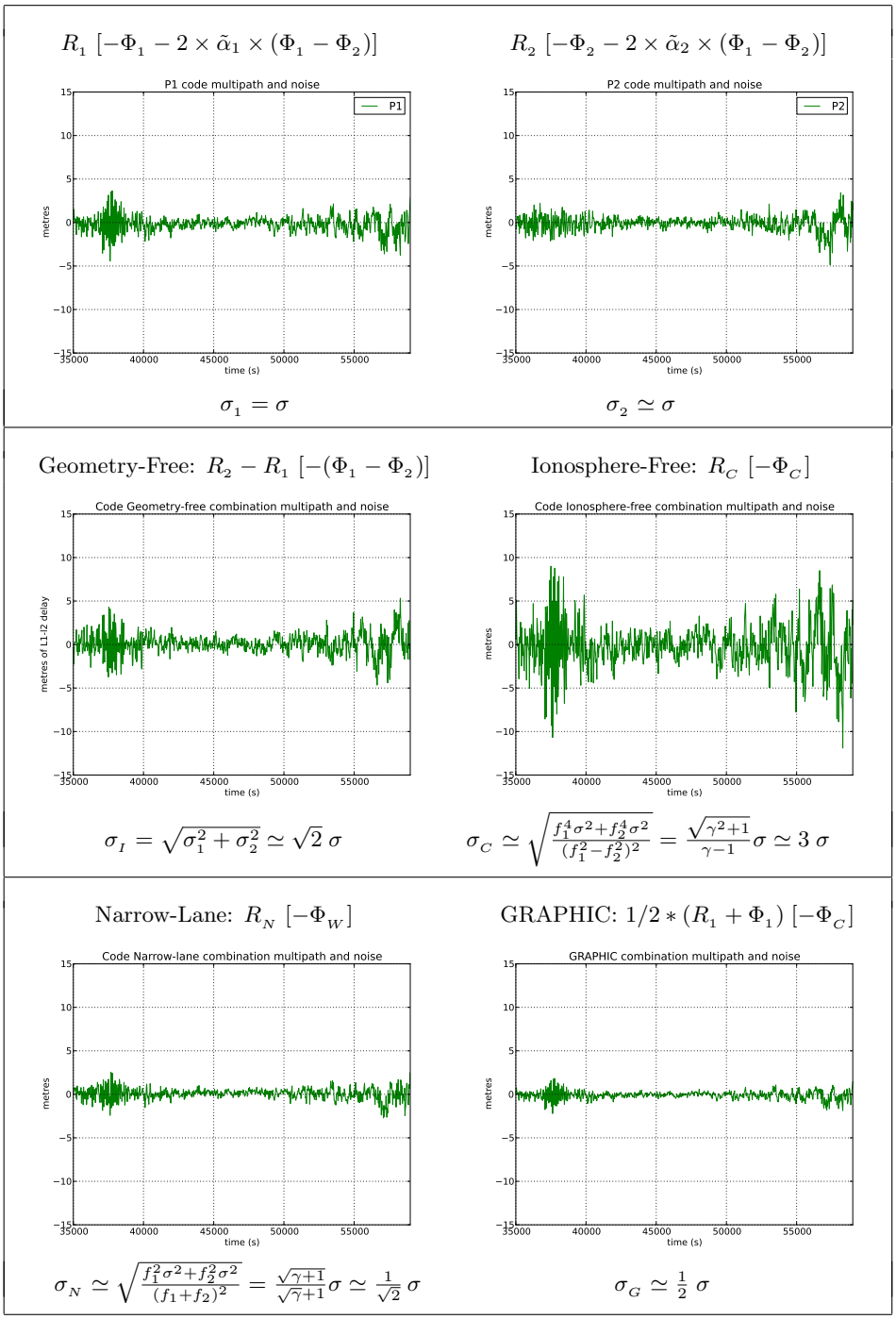


Figure 4.3: Comparison of code pseudorange noise for the original GPS R_1 and R_2 raw code measurements and their main combinations (see captions). The geometry and/or the dispersive delays have been removed by a combination of carrier phases (indicated within the brackets). The plots are shifted to remove carrier ambiguity. The approximate expression of σ noise is given above each plot. Note: $\tilde{\alpha}_1 = 1/(\gamma-1)$, $\tilde{\alpha}_2 = 1+\tilde{\alpha}_1$ and $\gamma = (77/60)^2 \simeq 1.65$.

Figure 4.3 shows a comparison of code pseudorange noise for the original GPS P1 and P2 code raw measurements and their combinations. The data were collected on 16 December 2008, in Barcelona, Spain, with an Ashtec Z-XII receiver and an AT2775-42W antenna.

In the plots of Fig. 4.3, the geometry and dispersive delays have been removed by a combination of carriers (indicated within the brackets at the top of each plot) in order to allow better visualisation of the code measurement noise. The approximate expression of σ noise for each combination (see section 4.1), assuming uncorrelated signal noise[6], is also given under each

[6]The expressions are based on the following results: given two independent random variables X, Y, and two constants a, b, then $\sigma_{aX+bY} = \sqrt{a^2\sigma_X^2 + b^2\sigma_Y^2}$.

plot. Notice the agreement between the noise shown in the different plots and the theoretical values derived from this simple approach, assuming the same noise in both frequencies. Many other examples with Galileo, GPS and Glonass signals can be found in the laboratory exercises of sessions 4.2 and 4.3 in Volume II.

The two main sources of receiver measurement error are presented below: namely, receiver noise and multipath.

4.2.1 Receiver Noise

The receiver code noise is a white-noise-like error and can be smoothed using a low-pass filter.

This error affects both the code and carrier measurements, but at different magnitudes. The accuracy of pseudorange measurements is about 1% of the wavelength ('chip'), or better. This means, for instance, noise with a maximum value of 3 m for the GPS civil C1 code (i.e. C/A code) and about 30 cm for the protected P codes (see Table 2.1). However, when smoothing the code with the carrier phase, the C1 code noise can be reduced to about 50 cm.

The carrier phase noise is at the level of few millimetres (about 1% of the carrier phase wavelength).

4.2.2 Multipath

The interference by multipath is generated when a signal arrives, by different ways, at the antenna (see Fig. 4.4). Its principal cause is the closeness of the antenna to the reflecting structures, and it is important when the signal comes from a satellite with low elevation. This error is different for different frequencies. It affects the phase measurements, as well as the code measurements. In the case of the code, it can reach a theoretical value of 1.5 times the wavelength ('chip'). This means, for instance, that multipath in the GPS C1 code can reach up to 450 m (see Table 2.1), although higher values than 15 m are difficult to observe. Typically, it is less than 2 or 3 m. In the case of the carrier, its theoretical maximum value is a quarter of the wavelength. This means about 5 cm for the GPS L1 or L2 signals, but it is typically less than 1 cm.

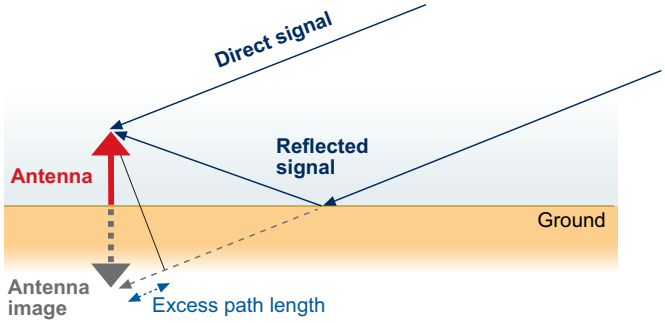


Figure 4.4: Difference in the optical path between the direct signal and the reflected signal.

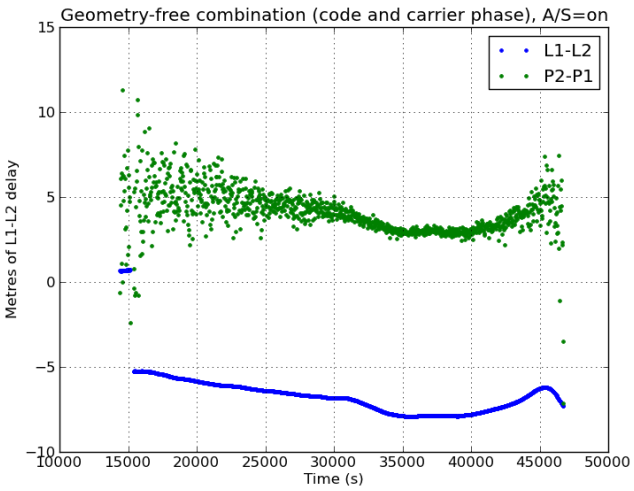


Figure 4.5: Effect of code multipath on GPS signals. The plot at the top shows the 'butterfly effect' due to the larger multipath at the satellite's rising and setting (i.e. due to the low-elevation rays). The plot at the bottom shows the repeatability of multipath signatures with the geometry: the code minus phase signal combination $R_1 - \Phi_1$ is plotted for a given satellite over three consecutive sidereal days. The drift is due to the ionospheric refraction. See exercises 3 to 7 of session 4.1 in Volume II.

The multipath error can be minimised by improving antenna directivity: that is, by attenuating the signal coming from certain low-elevation directions, and moving the antenna away from reflecting objects.

The effect of code multipath for low-elevation rays is depicted in the top plot of Fig. 4.5, where the geometry-free combinations of code $(R_2 - R_1)$ and carrier $(\Phi_1 - \Phi_2)$ measurements are plotted for a satellite–receiver track. The larger noise at the edges is due to the low-elevation rays at the rising and setting of the satellite. Notice that the carrier phase multipath (less than 1 cm) cannot be seen due to the scale of the figure and its 2 m divisions.

The geometric nature of multipath is illustrated in the bottom plot of Fig. 4.5, where the code minus phase combination $R_1 - \Phi_1$ is plotted for a given satellite over three consecutive days. The measurements were collected with a static GPS receiver and, therefore, the geometry of the satellite–receiver rays repeats every sidereal day.[7] Notice in this plot how the multipath signatures overlap after aligning the plots to the first day (i.e. by shifting 236 and 2×236 s, the second- and third-day plots, respectively) to account for the difference between the sidereal and solar times. The global drift is due to the ionospheric refraction $2I_1$, see equation (4.26) (data set collected on 28–30 September 1998, in Barcelona, Spain, with the single-frequency Lassen-SK8, Trimble receiver).

[7] That is, every 24 h minus 236 s of GPS time (see sections 2.1.1.1 and 3.1.1).

4.2.3 Carrier Smoothing of Code Pseudoranges

The noisy (but unambiguous) code pseudorange measurements can be smoothed with the precise (but ambiguous) carrier phase measurements. A simple algorithm (the Hatch filter) is given as follows.

Let $R(s;k)$ and $\Phi(s;k)$ be the code and carrier measurements of a given satellite s at k-th epoch. Then, the smoothed code $\hat{R}(s;k)$ can be computed as

$$\widehat{R}(s;k) = \frac{1}{n}R(s;k) + \frac{n-1}{n}\left[\widehat{R}(s;k-1) + (\Phi(s;k) - \Phi(s;k-1))\right] \quad (4.23)$$

The algorithm is initialised with $\widehat{R}(s;1) = R(s;1)$, where $n = k$ when $k < N$ and $n = N$ when $k \geq N$.

This algorithm must be initialised every time that a carrier phase cycle slip occurs.

The algorithm can be interpreted as a real-time alignment of the carrier phase with the code measurement. That is,

$$\begin{aligned}
\widehat{R}(k) &= \frac{1}{n}R(k) + \frac{n-1}{n}\left[\widehat{R}(k-1) + (\Phi(k) - \Phi(k-1))\right] \\
&= \Phi(k) + \frac{n-1}{n}\left(\widehat{R}(k-1) - \Phi(k-1)\right) + \frac{1}{n}\left(R(k) - \Phi(k)\right) \\
&= \Phi(k) + \frac{n-1}{n}\left\langle R - \Phi\right\rangle_{(k-1)} + \frac{1}{n}\left(R(k) - \Phi(k)\right) \\
&= \Phi(k) + \left\langle R - \Phi\right\rangle_{(k)}
\end{aligned} \quad (4.24)$$

where the mean bias[8] $\langle R - \Phi\rangle$ between the code and carrier phase is estimated in real time and used to align the carrier phase with the code.

4.2.3.1 Code–Carrier Divergence Effect: Single-Frequency Smoothing

The time-varying ionosphere induces a bias in the single-frequency smoothed code when it is averaged in the smoothing filter. This effect is analysed as follows.

The single-frequency code (R_1) and carrier (Φ_1) measurements given by the first two equations of (4.19) can be written in a simplified form as

$$\begin{aligned}
R_1 &= r + I_1 + \varepsilon_1 \\
\Phi_1 &= r - I_1 + B_1 + \epsilon_1
\end{aligned} \quad (4.25)$$

where r includes all non-dispersive terms such as geometric range, satellite and receiver clock offset and tropospheric delay. I_1 represents the frequency-dependent terms as the ionospheric and instrumental delays. B_1 is the carrier phase ambiguity term, which is constant along continuous

[8] The mean value of a set of measurements $\{x_1, \ldots, x_n\}$ can be computed recursively as $\langle x\rangle_k = (1/k)x_k + [(k-1)/k]\langle x\rangle_{k-1}$. Equation (4.24) is a variant of the previous expression and provides an estimate of the moving average over a window of N samples. Note that, when $k \geq N$, the weighting factors $1/N$ and $(N-1)/N$ are used instead of $1/k$ and $(k-1)/k$.

carrier phase arcs. ε_1 and ϵ_1 account for the code and carrier thermal noise and multipath.

Since the ionospheric term has opposite sign in code and carrier measurements, it does not cancel in the R–Φ combination, but, on the contrary, its effect is twofold. That is,[9]

$$R_1 - \Phi_1 = 2I_1 - B_1 + \varepsilon_1 \tag{4.26}$$

The term $2I_1$ is often called *code–carrier divergence*, because it results from the fact that the ionosphere affects code and carrier in different ways, that is, the ionosphere delays the code and advances the carrier by the same amount.

Substituting equations (4.25) and (4.26) into (4.24) results in

$$\widehat{R}_1(k) = \Phi_1(k) + \langle R_1 - \Phi_1 \rangle_{(k)} = r(k) - I_1(k) + B_1 + \langle 2I_1 - B_1 \rangle_{(k)} \tag{4.27}$$

Since the carrier ambiguity term B_1 is a constant bias and the average $\langle \cdot \rangle$ is a linear operator, we can assume that $B_1 \approx \langle B_1 \rangle_{(k)}$ and then they cancel in the previous equation (4.27), which can be rewritten as

$$\widehat{R}_1(k) = \Phi_1(k) + \langle R_1 - \Phi_1 \rangle_{(k)} = r(k) + I_1(k) + \underbrace{2\left(\langle I_1 \rangle_{(k)} - I_1(k)\right)}_{bias_I} \tag{4.28}$$

(where the ionosphere is a time-varying term). If the ionosphere were constant, the averaged value $\langle I_1 \rangle_{(k)}$ would coincide with the instantaneous value $I_1(k)$ and, hence, the bias $bias_I$ would cancel. However, the time-varying ionosphere will result in a bias that depends on the magnitude of the temporal gradient.

That is, the time-varying ionosphere produces a bias in the single-frequency carrier-smoothed code due to the code–carrier divergence effect, in such a way that the first equation of (4.25) becomes, for the smoothed code $\widehat{R}_1$,

$$\widehat{R}_1 = r + I_1 + bias_I + \nu_1 \tag{4.29}$$

where ν_1 is the noise term after the filter smoothing. The magnitude of this bias is a function of the smoothing window N.

In order to assess such an effect, let us assume a simple model where STEC varies linearly over time:

$$I_1(t) = I_{0_1} + \dot{I}_1\, t \tag{4.30}$$

In exercise 17 of session 4.1 in Volume II, it is shown that, assuming equation (4.30), the bias in the smoothed code, in the steady state, is given by

$$bias_I = 2\left(\langle I_1 \rangle_{(k)} - I_1(k)\right) = -2\tau\, \dot{I}_1 \tag{4.31}$$

where τ is the filter smoothing time constant ($\tau = N\Delta t$ in equation (4.23)).

Figures 4.6 and 4.7 show two examples of the error induced by the ionosphere in the single-frequency smoothed code (see exercises 15 and 16 of session 4.1 in Volume II). Figure 4.6 corresponds to the Halloween storm (on 30 October 2003) with high ionospheric temporal gradients producing up to 8 m of error. Figure 4.7 is for a quiet scenario and compares the performance using different smoothing time constants.

[9] Where the carrier term ϵ_1 is negligible compared to the code noise and multipath ε_1.

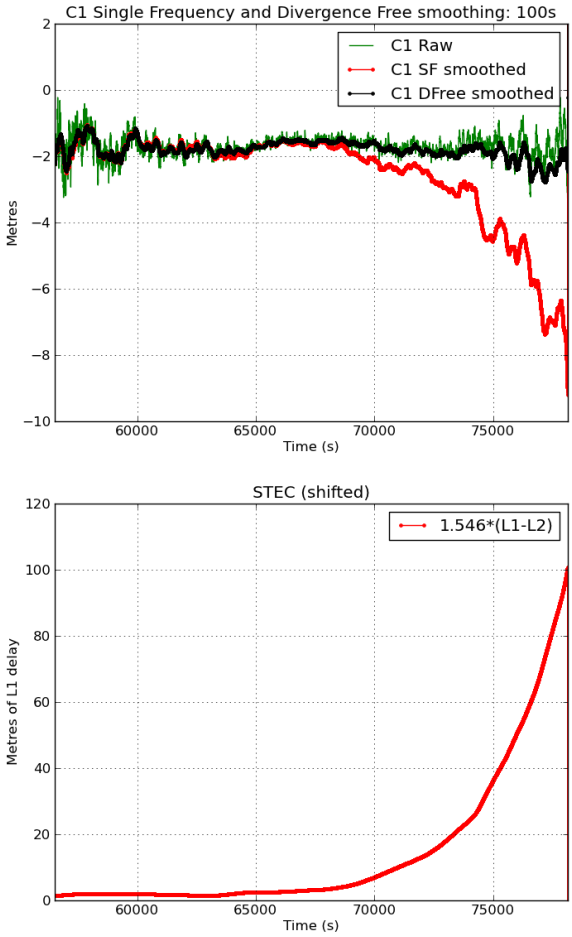


Figure 4.6: Effect of 100 s smoothing during the Halloween storm. The top plot shows the C1–carrier smoothing using equation (4.26), in red (single-frequency smoother), and using equation (4.32), in black (divergence-free smoother). The raw measurements are shown in green. STEC is depicted in the bottom plot. As it is shown, the larger temporal ionospheric gradients lead to larger code–carrier divergence-induced error in the single-frequency smoothed solution, which reaches up to about 8 m in this example. See exercise 16 of session 4.1 in Volume II.

4.2.3.1.1 Divergence-Free Smoother

With dual frequency measurements, a combination of carriers with the same ionospheric delay as the code (the same sign) can be generated ($\Phi_{1_{DF}}$). Thus, neglecting the carrier noise and multipath ϵ_1 and ϵ_2 against the code ε_1,

$$\begin{aligned} R_1 - \Phi_{1_{DF}} &= B_{12} + \varepsilon_1 \\ \Phi_{1_{DF}} &= \Phi_1 + 2\tilde{\alpha}_1(\Phi_1 - \Phi_2) \end{aligned} \tag{4.32}$$

Using this new carrier measurement $\Phi_{1_{DF}}$ in the Hatch filter (either equation 4.23 or 4.24) instead of Φ_1, a smoothed code is obtained, free from the code–carrier ionospheric divergence effect, and having the same ionospheric delay as the original unsmoothed code R_1:

$$\widehat{R}_1 = r + I_1 + \nu_{12} \tag{4.33}$$

where ν_{12} is the noise after the filter smoothing. This smoothed code is called *divergence-free* carrier-smoothed code.

Comment: This smoothed code is not biased by the temporal gradients of the ionosphere (unlike the single-frequency smoother), being the same ionospheric delay as in the original raw code (i.e. I_1). Nevertheless, as it

is still affected by the ionosphere, its spatial decorrelation[10] must be taken into account in differential positioning.

Figures 4.6 and 4.7 compare the results of single-frequency and divergence-free smoothers. The sinusoidal-like oscillations are due to the code multipath. These oscillations are smoothed by increasing the time constant. This works well with the divergence-free smoother, but a larger ionospheric bias is introduced in the case of the single-frequency smoothing.

4.2.3.1.2 Ionosphere-Free Smoother

The previous combination of equation (4.32) uses two-frequency carriers (Φ_1,Φ_2) but only a single-frequency code (R_1).

Using both code and carrier dual-frequency measurements, it is possible to remove the frequency-dependent effects[11] using the code and carrier ionosphere-free combinations (R_C, Φ_C), see equations (4.4). Thus, equations (4.25) become

$$\begin{aligned} R_C &= r + \varepsilon_C \\ \Phi_C &= r + B_C + \epsilon_C \end{aligned} \tag{4.34}$$

Now, using the previous ionosphere-free combinations of code and carrier, equation (4.26) becomes

$$R_C - \Phi_C = B_C + \varepsilon_C \tag{4.35}$$

Applying this combination as the input in the Hatch filter (either equation 4.23 or 4.24), a smoothed solution completely free of the (first-order) ionosphere is computed:

$$\widehat{R}_C = r + \nu_C \tag{4.36}$$

where ν_C is the noise after filter smoothing. This smoothed code is called the *ionosphere-free* carrier-smoothed code.

The right side of Fig. 4.7 shows the ionosphere-free carrier-smoothed solution computed using the same smoothing time constants as with the C1 code. Notice the larger noise in this combination regarding the C1 code, as well as the unbiased smoothed solution (see also Fig. 4.3).

Comment: This dual-frequency smoothing is based on the ionosphere-free combination of measurements, and therefore it is unaffected by either the spatial or the temporal ionospheric gradients, but has the disadvantage that the noise is amplified by a factor of 3 (using the legacy GPS signals).

Note that the left-hand plots of Fig. 4.3 are generated in exercise 15 of session 4.1 in Volume II.

[10] The spatial decorrelation of the ionosphere is typically at a level of 1 to 4 mm/km, but in extreme conditions can reach up to 400 mm/km. Such spatial gradients moving at 0.75 km/s are being considered in the Ionospheric Threat Model, which results in potential temporal gradients at a given location of about 300 mm/s [Konno, 2007], which can lead to a large bias in the single-frequency smoother.

[11] That is, the first-order ionosphere, which accounts for up to 99.9% of the ionospheric effect, and the interfrequency bias, see equations (4.19).

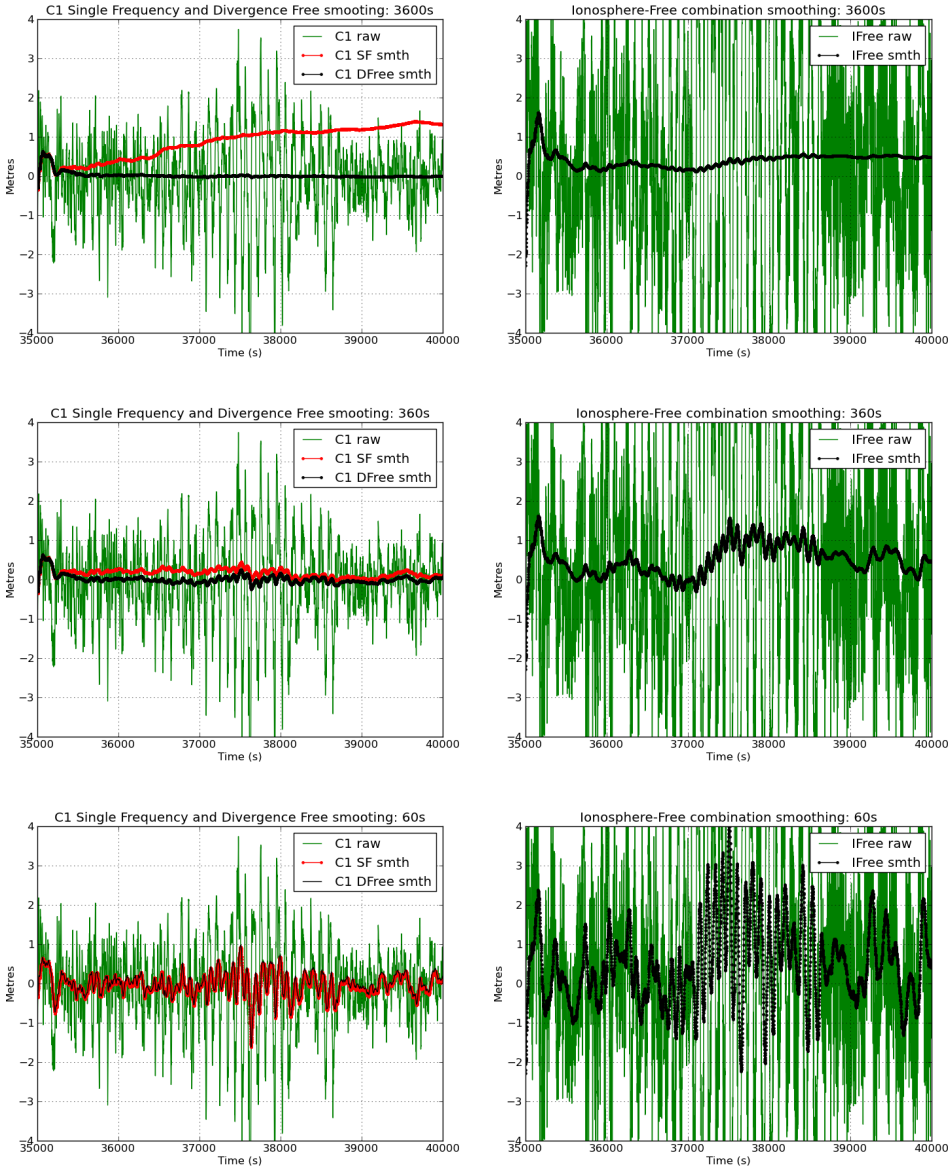


Figure 4.7: Carrier-smoothed codes with different time filter lengths: 1 h (first row), 6 min (second row), 1 min (third row). Left-hand plots show the Single Frequency (SF) smoothing using equation (4.26) in red and the Divergence Free (DFree) smoothing using equation (4.32) in black; right-hand plots show the Ionosphere-Free (IFree) smoothing using equation (4.35). The red curves (left-hand plots) accumulate the bias due to the divergence of the ionosphere. The larger noise in the right-hand plots is due to the ionosphere-free combination of codes (which amplifies the noise by a factor of 3). The unsmoothed measurements in green correspond to R_1 (*left*) and R_C (*right*) of Fig. 4.3 (zoom to [35000:40000] s). Note that in this example R_1 is the GPS C1 code measurement and R_C is the IFree combination of GPS codes. See exercise 15 of session 4.1 in Volume II.

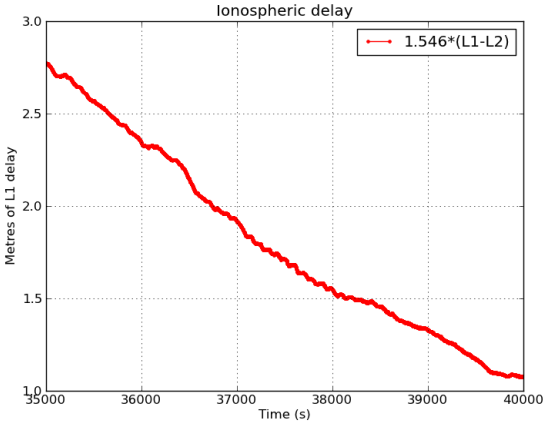


Figure 4.8: The ionospheric delay variation associated with Figure 4.7 computed from $\Phi_1 - \Phi_2$ (shifted) in metres of L1 signal delay.

4.3 Carrier Phase Cycle-Slip Detection

As already mentioned, receiver losses of lock cause discontinuities in the phase measurements (*cycle slips*) that are seen as jumps of integer numbers of wavelengths λ (i.e. the integer ambiguity N changes by an arbitrary integer value).

Different heuristic methods are used for cycle-slip detection, operating over undifferenced, single-differenced or double-differenced measurements between pairs of satellites and receivers.

The methods presented in this section[12] are oriented towards single-receiver positioning, and thus do not require any differencing of data between receivers, being suitable for implementation in real time. Moreover, they are based on using only combinations of measurements at different frequencies, or just one frequency measurement. That is, they do not need any geometric delay modelling.

4.3.1 Examples of Multifrequency Cycle-Slip Detectors

With two-frequency signals (or multifrequency signals in general) it is possible to build combinations of measurements to enhance the reliability of cycle-slip detection. The target is to remove the geometry, which is the largest varying effect,[13] the clocks and the other non-dispersive delays, as well as ionospheric delays.

Two types of examples of detectors will be presented in this section: detectors based on carrier phase measurements only; and detectors based on code and carrier phase data. In the first type, carrier phase measurements of signals at two different frequencies are subtracted in order to remove the geometry and all non-dispersive effects. This provides a very precise test signal (multipath and noise less than 1 cm), although it is affected by the ionospheric refraction. However, this effect varies as a smooth function and can be modelled by a low-degree polynomial fit. Nevertheless, high ionospheric activity conditions can degrade the performance of this detector, mainly with low sampling rate data.

As the cycle slips can occur in each of the signals independently, two independent combinations must be use to ensure that all possible jumps are taken into account. In this way, the simultaneous use of two independent detectors protects against those situations where the combination of ΔN_1 and ΔN_2 cycle slips would produce inappreciable jumps in the geometry-free combination.[14]

The second type of detector is based on the MW combination of code and carrier phase measurements [Blewitt, 1990]. This combination cancels not only the non-dispersive effects, but also the ionospheric refraction. Nevertheless, the resulting test signal (i.e. MW combination) is affected by

[12] Most are examples based on our own algorithms presently in use.

[13] The range ρ varies up to hundreds of metres in 1 s.

[14] For instance, with GPS signals, $(\Delta N_1, \Delta N_2) = (9, 7)$ or $(68, 53)\ldots$, produces jumps of few millimetres in the geometry-free combination. In particular, no jump happens when $(\Delta N_1, \Delta N_2) = (77, 60)$, but this event produces a jump of $17\lambda_W \simeq 15$ m in the wide-lane combination.

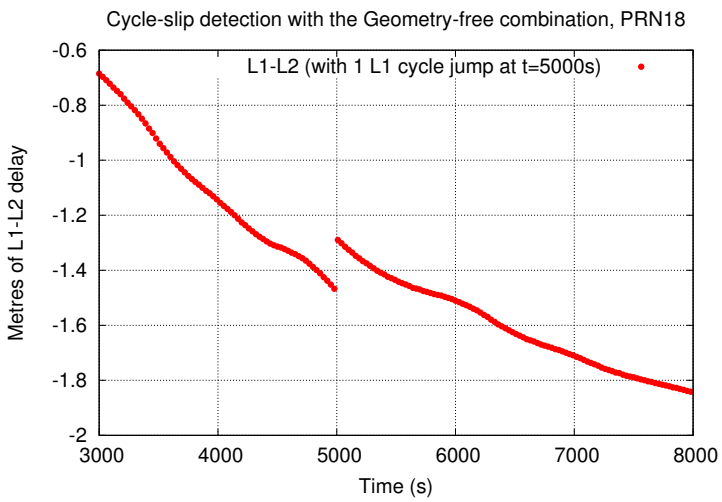


Figure 4.9: Effect of one-cycle jump in the GPS Φ_1 carrier phase signal on the ionosphere-free combination. The horizontal axis is seconds of day; the vertical axis is in metres.

the code multipath, which can reach up to several metres. The impact of this noise is partially reduced by the increased ambiguity spacing of the wide-lane combination of carrier phases, on the one hand, and the noise reduction due to the narrow-lane combination of code measurements, on the other hand (both of which are involved in the MW combination). Nevertheless, and in spite of these benefits, the performance is worse than in the carrier-phase-based detector and it is used as a secondary test.

4.3.1.1 Detector Based on Carrier Phase Data: The Geometry-Free Combination

With two-frequency signals it is possible to obtain the carrier phase geometry-free combination, in order to remove the geometry, including clocks, and all non-dispersive effects in the signal. As commented previously, in non-disturbed conditions, this very precise (i.e. with very low-noise) test signal performs as a smooth function, driven by the ionospheric refraction, with very few changes between close epochs. Indeed, although, for instance, the jump produced by a simultaneous one-cycle slip in both signal components is smaller in this combination than in the original signals,[15] it can provide reliable detection, even for small jumps.

The easiest way to build a cycle-slip detector is to consider the differences in time of consecutive samples (see Fig. 4.9). A refinement of this procedure is the use of nth-order differences to take advantage of the jump amplitude enlargement produced by the differencing process (see Table 4.3).

This approach allows us to make a reasonable enough detector for many applications. Nevertheless, it must be taken into account that, as the jumps are enlarged, also the signal noise (i.e. signal instabilities) is amplified, which can lead to false detections in some scenarios (for instance, with low signal-to-noise ratios, scintillation, etc.).

One way to mitigate the impact of these effects is to use a low-order polynomial fit, reducing the test signal noise. This concept is the basis of the detector example presented next.

[15] For GPS signals, this jump is $\lambda_2 - \lambda_1 = 5.4\,\text{cm}$, which is about 3–4 times shorter than $\lambda_1 = 19.0\,\text{cm}$ or $\lambda_2 = 24.4\,\text{cm}$ (see Table 4.1).

Table 4.3: Computational scheme of differences: a jump in amplitude ϵ happens at time t_4 and its effect is propagated and amplified by the n-th order differences (from [Hofmann-Wellenhof et al., 2008]).

t_i	$y(t_i)$	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
t_1	0				
		0			
t_2	0		0		
		0		ϵ	
t_3	0		ϵ		-3ϵ
		ϵ		-2ϵ	
t_4	ϵ		$-\epsilon$		3ϵ
		0		ϵ	
t_5	ϵ		0		$-\epsilon$
		0		0	
t_6	ϵ		0		
		0			
t_7	ϵ				

Description of Algorithm: The detection is based on fitting a second-degree polynomial over a sliding window of N_I samples (e.g. $N_I = 10$). The predicted value from this polynomial is compared with the observed value to detect the cycle slip. As the geometry-free combination is affected by the ionospheric refraction, a sampling-rate-dependent threshold is considered.

Input data: Geometry-free combination of carrier phase measurements

$$\Phi_I(s;k) = \Phi_1(s;k) - \Phi_2(s;k) \tag{4.37}$$

Output: [satellite, time, cycle-slip flag].

For each epoch (k)

For each tracked satellite (s)

- Declare cycle slip when data gap greater than $tol_{\Delta t}$.[16]
- Fit a second-degree polynomial $P(s;x)$ to the previous values (after the last cycle slip)[17] $[\Phi_I(s;k-N_I),\ldots,\Phi_I(s;k-1)]$.
- Compare the measured $\Phi_I(s;k)$ and the predicted value $P(s;k)$ at epoch k. If the discrepancy exceeds a given *threshold*, then declare cycle slip. That is,
  if $|\Phi_I(s;k) - P(s;k)| > threshold$, then cycle slip.
- Reset algorithm after cycle slip.

End

End

An example of this algorithm implementation can be found in program `PreProcess.c` in Volume II.

[16] For instance, 60 s.

[17] A simpler approach is to construct a Lagrange interpolating polynomial (see section 3.3.3) with the last three values. For $N_I = 3$, this would be equivalent to considering the third-order differences in time of consecutive samples (i.e. $\Delta^3 P_2(t) \equiv 0$).

Comments: As mentioned previously, the relative variation between two consecutive epochs t_1 and t_2 depends on the elapsed time interval between them (i.e. $\Delta t = t_2 - t_1$), due to the variation in ionospheric refraction. Thus, a Δt-dependent threshold is recommended to account for the measurement sampling rate. For instance, $threshold = a_0 - a_1 \exp(-\Delta t/T_0)$ can be used.

Note that a_0 is the maximum threshold. On the other hand, taking for example $a_1 = a_0/2$ and $T_0 = 60\,\text{s}$, then the minimum threshold is $a_0[1 - \exp((-\Delta t/T_0)/2)]$, which corresponds to about $a_0/2$ for 1 s sampling rate and about $2a_0/3$ for 30 s sampling rate data. In this way, for instance, if a_0 is set as $a_0 = \frac{3}{2}(\lambda_2 - \lambda_1)$, the minimum detectable jump between two contiguous measurements will be $\frac{3}{4}(\lambda_2 - \lambda_1)$ or $\lambda_2 - \lambda_1$ for a 1 s or 30 s sampling rate, respectively.[18]

4.3.1.2 Detector Based on Code and Carrier Phase Data: The MW Combination

The MW combination provides a noisy estimate of the wide-lane ambiguity B_W, according to the equation

$$B_W = \Phi_W - R_N = \lambda_W N_W + b_W + \varepsilon \tag{4.38}$$

where $N_W = N_1 - N_2$ is the integer wide-lane ambiguity, b_W accounts for the satellite and receiver instrumental delays and ε is the measurement noise, including carrier phase and code multipath (see equations (4.19)).

This combination has a double benefit. On the one hand, the wide-lane combination has a larger wavelength $\lambda_W = c/(f_1 - f_2)$ than each signal individually (see Table 4.1), which leads to an enlargement of the ambiguity spacing.[19] On the other hand, the measurement noise is reduced by the narrow-lane combination of code measurements,[20] reducing the dispersion of values around the true bias.

A simple algorithm, suitable for running in real time, is presented as follows (see [Blewitt, 1990]).[21]

Description of Algorithm: The detection is based on the real-time computation of mean and sigma values of the measurement test data B_W. A cycle slip is declared when a measurement differs from the mean bias value over a predefined number of standard deviations (S_{B_W}), that is the threshold.

Input data: MW combination of B_W of code–carrier phase.

Output: [satellite, time, cycle-slip flag].

[18] Note that, from equation (4.20), $\lambda_2 - \lambda_1$ is the jump produced on the geometry-free combination Φ_I when a jump of one cycle occurs simultaneously in both carriers. This jump, for instance for the GPS L1 and L2 signals, is $\lambda_2 - \lambda_1 = 5.4\,\text{cm}$ (see Table 4.1).

[19] The noisy measurements are concentrated around discrete levels of separated multiples of λ_W units (see Fig. 4.10, right). That is, the jumps are integer numbers of λ_W.

[20] Namely, $\sigma_W^2 = (f_1^2\sigma_1^2 + f_2^2\sigma_2^2)/(f_1 + f_2)^2 \simeq 1/2\sigma_1^2$ (see Fig. 4.3).

[21] This algorithm was designed to work with the wide-lane combination ($\lambda_W \simeq 75$–$85\,\text{cm}$), see Table 4.1. Of course, its performance will dramatically improve if the extra-wide-lane combination ($\lambda_{EW} \simeq 6$–$10\,\text{m}$) is used, see Table 4.1.

For each epoch (k)

For each tracked satellite (s)

- Declare cycle slip when data gap greater than $tol_{\Delta t}$ (e.g. 60 s).
- Compare the measurement $B_W(s;k)$ at epoch k with the mean bias $m_{B_W}(s;k-1)$ computed from the previous values. If the discrepancy is over a *threshold* of $K_{factor} \times S_{B_W}$ (e.g. $K_{factor} = 6$), declare cycle slip. That is,
  if $|B_W(s;k) - m_{B_W}(s;k-1)| > K_{factor} \times S_{B_W}(s;k-1)$,
  then cycle slip.
- Update the mean and sigma values according to the equations.

$$m_{B_W}(s;k) = \frac{k-1}{k} m_{B_W}(s;k-1) + \frac{1}{k} B_W(s;k)$$
$$S^2_{B_W}(s;k) = \frac{k-1}{k} S^2_{B_W}(s;k-1) + \frac{1}{k}(B_W(s;k) - m_{B_W}(s;k-1))^2 \quad (4.39)$$

  Note: S_{B_W} can be initialised with an a priori $S_0 = \lambda_W/2$.
- Reset algorithm after cycle slip.

End

End

An example of this algorithm implementation can be found in program `PreProcess.c` in Volume II.

Comments:

- The calculation of the *mean* is exact (over the $k = 1, \ldots, n, \ldots$ samples), while the computation of *sigma* is a good approximation ($O(1/k^2)$). The equations (4.42) can be used as well.
- A lower, or lower and upper, limit for the threshold can be set in order to protect the test from unrealistic noise estimates of S_{B_W}.
- The mean bias estimate m_{B_W} can be greatly affected by strong code multipath at the beginning of the data arch (due to low-elevation rising satellites), but as the number of averaged samples increases, this estimate becomes more stable and robust. This improvement of m_{B_W} with the number of processed samples does not necessarily benefit S_{B_W} as an estimate of the data noise to define the detection threshold. Indeed, as the number of samples increases, the value of $S^2_{B_W}$ is frozen, becoming more insensitive to the measurement noise variations. A solution to this problem is to compute a *moving sigma* (only for S_{B_W}, not m_{B_W}) over the last N-sample window as in section 4.3.2.

The effect of the ambiguity space widening produced by the MW combination is shown in Fig. 4.10 and compared with the single-frequency *phase minus code combination* (see section 4.3.2)[22] as a reference. As in Fig. 4.9,

[22] This combination (Φ–R) cancels all non-dispersive effects (geometry, clocks, etc.) so only the ionospheric refraction remains (among the instrumental delays), producing the drift seen in the figure.

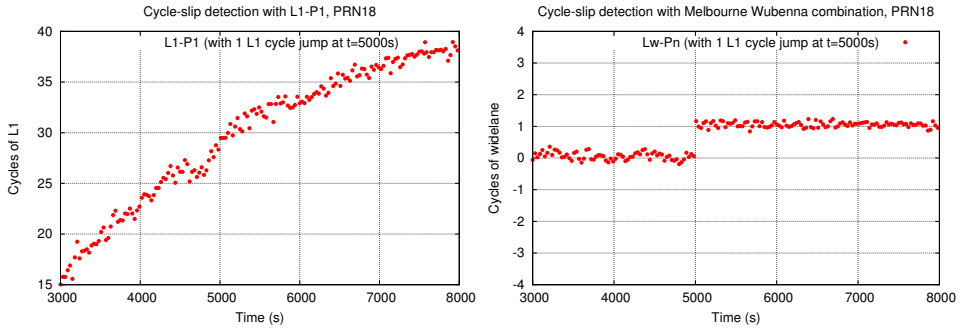


Figure 4.10: Effect of one-cycle jump in the GPS L1 signal in the Φ–R (*left*) and MW (*right*) combination (raw measurements without smoothing). Vertical axes are in cycles of $\lambda_1 \simeq 19\,\text{cm}$ (*left*) and $\lambda_W \simeq 86\,\text{cm}$ (*right*).

a jump of one cycle is introduced in the Φ_1 carrier phase measurement at time 5000 s. This jump cannot be identified from the Φ–R combination shown in the left plot of Fig. 4.10, due to the receiver code noise and multipath, and to the ionospheric drift. On the contrary, it is clearly seen in the MW combination plot shown on the right of the figure, which has a lower relative noise and is not affected by the ionospheric refraction.

Figure 4.10 (right) shows a nice example of cycle-slip detection with the MW combination, where the jump is well defined.[23] Unfortunately, this is not the case on many occasions, because the detection threshold is 'fussier' due to the code receiver noise and multipath. This noise can be smoothed by filter averaging (i.e. by computing the mean bias B_W), but small jumps can still escape from the detector in the first epochs following a filter reset.

Large code multipath or, even worse, undetected cycle slips at the beginning of the filter convergence can propagate forward large errors that can lead to large S_{B_W} values, increasing the detection threshold ($K_{factor} \times S_{B_W}$) and hence causing misdetections. On the other hand, use of small K_{factor} values to increase the detector's sensibility can lead to a higher fault-detection probability.

4.3.1.3 Cycle-Slip Detection with Three-Frequency Signals

The reliability of cycle-slip detection can be significantly improved with three-frequency signals. Indeed, on the one hand the extra-wide-lane combination provides a wavelength of several metres that can dramatically improve the sensibility of this code-based detector, overcoming most of the problems mentioned previously in section 4.3.1.2. On the other hand, two independent geometry-free combinations are possible to generate, which allow us to relay the detection in carrier phase data detectors as in those described in section 4.3.1.1.

Thus, a simple scheme can be based on the following elements: (1) two carrier-phase-based detectors for two independent geometry-free combinations; and (2) an MW detector using the extra-wide-lane combination.

[23] The measurements shown in this figure are unsmoothed. They were collected under *A/S off* conditions (IGS station CASA, California, USA, 18 October 1995).

4.3.2 Examples of Single-Frequency Cycle-Slip Detectors

The examples of single-frequency detectors presented next are based only on data measurements of a single receiver and do not use any geometric delay model. They are simple algorithms, suitable for running in real time, but with a worse performance than the two-frequency detectors of previous sections. Other algorithms using geometric modelling or single or double differences between satellites, or satellites and receivers, can improve the detection threshold and the reliability, but their study is beyond the scope of this book.

The non-dispersive delays (geometry, clocks, troposphere, etc.) are cancelled when forming the code pseudorange and carrier phase combination for a given satellite and receiver measurement, that is $\Phi - R = \lambda N - 2I + K + \varepsilon$, where the ionospheric refraction I is affected by a factor of 2. The terms N, K and ε indicate the ambiguity, instrumental delays and measurement noise, respectively. Two example algorithms will be presented in this section based on the following considerations:

- The ionospheric term I varies slowly with time, with small changes between consecutive epochs (typically less than 1–2 cm in 30 s).
- The measurement noise ε can reach up to several cycles, but it can be smoothed by a polynomial fit of the data measurements over a moving window (Example 1); or a smoothed prediction can be computed by averaging the samples as in the detector of the previous section (Example 2).

Example 1: Description of Algorithm.

The detection is based on fitting an nth-degree polynomial over a sliding window of N samples (e.g. $N = 200$ at 1 Hz). The residuals (predicted − observed) are compared to detect the carrier phase cycle slips.[24] As the (geometry-free) combination used, equation (4.40), is two times the ionospheric refraction, a sampling-rate-dependent threshold can also be considered.

Input data: Code minus phase combination of single-frequency data

$$d(s;k) = \Phi(s;k) - R(s;k) \tag{4.40}$$

Output: [satellite, time, cycle-slip flag].

[24] This approach is inspired by the algorithms developed by [Blewitt, 1990] for the geometry-free combination of two-frequency signals. Note that a similar scheme is applied here for the code and phase combination of a single-frequency signal.

For each epoch (k)

For each tracked satellite (s):

- Declare cycle slip when data gap greater than $tol_{\Delta t}$.[25]
- If no data gap larger than $tol_{\Delta t}$, then:
- Update an array with the previous N values (after the last cycle slip) $[d(s;k-N),\ldots,d(s;k-1)]$.
- Fit an nth-degree polynomial $P(s;x)$ to the previous epochs.[26]
- Declare cycle slip when $|d(s;k) - P(s;k)| > Threshold.$
- Reset algorithm after cycle slip.

End

End

Example 2: Description of Algorithm.

The detection is based on computing the mean and sigma values of the code pseudorange and carrier phase (Φ–R) differences over a sliding window of N samples (e.g. $N = 100$ with 1 Hz data). A cycle slip is declared when a measurement differs from the mean bias value over a predefined threshold.

Input data: Code pseudorange (R) and carrier phase (Φ) measurements.

Output: [satellite, time, cycle-slip flag].

For each epoch (k)

For each tracked satellite (s):

- Declare cycle slip when data gap greater than $tol_{\Delta t}$.[27]
- If no data gap larger than $tol_{\Delta t}$, then:
- Update an array with the last N differences of

$$d(s;k) = \Phi(s;k) - R(s;k) \tag{4.41}$$

  That is, $[d(s;k-N),\ldots,d(s;k-1)]$.
- Compute the mean and sigma discrepancy over the previous N epochs (after the last cycle slip) $[k-N,\ldots,k-1]$:

$$\begin{aligned} m_d(s;k-1) &= \frac{1}{N}\sum_{i=1}^{N} d(s;k-i) \\ m_{d^2}(s;k-1) &= \frac{1}{N}\sum_{i=1}^{N} d^2(s;k-i) \\ S_d(s;k-1) &= \sqrt{m_{d^2}(s;k-1) - m_d^2(s;k-1)} \end{aligned} \tag{4.42}$$

[25] For instance, 60 s.
[26] Typically, $n = 2$ or 3 can be used.
[27] For instance, 15 s with 1 Hz data.

- Compare the difference at epoch k with the mean value of differences computed over the previous N-sample window. If the value is over a $threshold = n_T \times S_d$ (e.g. $n_T = 6$), declare cycle slip.[28]
  That is,
  if $|d(s;k) - m_d(s;k-1)| > n_T \times S_d(s;k-1)$, then cycle slip.
- Reset algorithm after cycle slip.

End

End

An example of this algorithm implementation can be found in program `PreProcess.c` in Volume II.

Comments:

- This algorithm can be seen as a particular case of Example 1, using a zero-degree polynomial fit.
- This detector is affected by the pseudorange noise and multipath, as well as the divergence of the ionosphere. Thus, higher sampling rates improve detection performance, but the shortest jumps can still escape from this detector. On the other hand, a minimum number of samples are needed for filter initialisation, in order to ensure a reliable value of S_d for the detection threshold.
- To avoid unrealistic estimates of sigma during the first iterations of the filter, the following weighted average with an initial value of S_0^2 can be used (see Fig. 4.11 for an example of implementation performance):

$$\widetilde{S}_d^2(s;n) = \frac{n-1}{n} S_d^2(s;n) + \frac{1}{n} S_0^2 \tag{4.43}$$

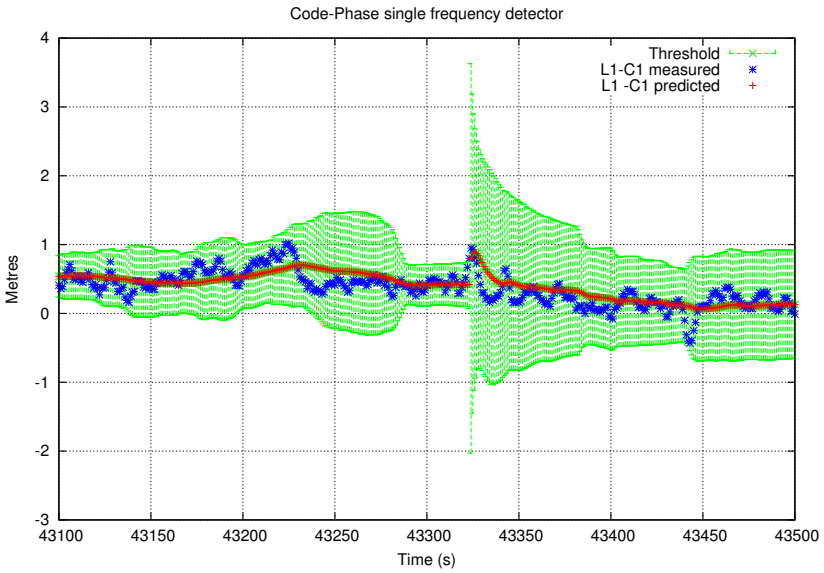


Figure 4.11: Fault cycle-slip detection on the GPS L1 signal. A multipath drift, after a quite stable period (i.e. with small sigma), produces a fault detection and the detector is reinitialised. The measured and predicted differences are compared with the confidence threshold (computed with the sliding window algorithm, using equation (4.42) combined with (4.43)). See also the plots of exercises 9 to 11 of session A.1 in Volume II.

[28] The mean bias between code and phase computed from the 'previous N samples' is taken as the prediction of the actual difference at epoch k (i.e. $\hat{d}(s;k) \equiv m_d(s;k-1)$) with confidence $n_T \times S_d(s;k-1)$.

where S_0 is a predefined initial value for sigma (e.g. $S_0 = 1\,\mathrm{m}$). Another easier approach is to fix a lower, or a lower and an upper, bound for the sigma threshold. That is, to take $threshold = n_T \times S_d$, with $th_{min} \leq threshold \leq th_{max}$.

- The Hatch filter can be used instead of the finite window of equations (4.42) to compute the mean values m_d, m_{d^2} in order to simplify the code:

$$\begin{aligned} m_d(s;n) &= \frac{a-1}{a} m_d(s;n-1) + \frac{1}{a} d(s;n) \\ m_{d^2}(s;n) &= \frac{a-1}{a} m_{d^2}(s;n-1) + \frac{1}{a} d^2(s;n) \end{aligned} \tag{4.44}$$

  where $a = n$ when $n < N$ and $a = N$ when $n \geq N$. These equations allow computation of a sequential estimate of the mean and sigma values, but this filter has infinite memory, propagating forward the divergence of the ionospheric refraction.[29] Nevertheless, such an accumulated effect, although biasing the ambiguity estimate, should not affect the cycle-slip detection, because it varies smoothly and the detector looks for large jumps.

Additional Comments. Differences in the Time Detector: As in the previous section, a detector based on the nth-order time differences of carrier phase Φ measurements between consecutive epochs could be considered (see section 4.3.1.1). Nevertheless, it must be taken into account that such differences are affected by changes in geometric range and in clocks.[30] In spite of this, most of the geometric range variation is cancelled from Δ^3, because it varies as a smooth function. Thence, this detector is mainly affected by receiver clock instabilities and high ionospheric fluctuations, as well.

In the case of the receiver clock effects, these can be removed by considering single differences of measurements between pairs of satellites in view, although this will enlarge the noise. The ionospheric effects can not be completely removed with such differences, but they should not be a problem, except in highly disturbed scenarios.

[29] On the contrary, equations (4.42) provide an estimate based on a (sliding) window, with a finite number of points.

[30] In the previous case in section 4.3.1.1, the geometry-free combination was used, cancelling all the non-dispersive effects (geometric range, clocks, etc.).

5. Measurement Modelling

As explained in the previous chapter, the code and carrier phase measurements contain several additional time delays associated with the signal propagation or with the clocks, among the geometric range between the satellite and receiver antenna phase centres. These delays are common in code and carrier measurements, except for the wind-up (see section 5.5) and the ambiguities that affect only the carrier measurements, and for a sign in the ionospheric refraction, delaying the code and advancing the carrier measurements, see section 5.4.1.

Figure 5.1 shows the layout of the different time delay terms in the pseudorange to illustrate the contents of the measurements.

Recalling equations (4.19), the measurements can be modelled by[1]

$$\begin{aligned} R_i &= \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr + \tilde{\alpha}_i(I + K_{21}) + \mathcal{M}_i + \varepsilon_i \quad (i = 1, 2) \\ \Phi_i &= \rho + c(\delta t_{rcv} - \delta t^{sat}) + Tr - \tilde{\alpha}_i\, I + b_i + \lambda_i\, N_i + \lambda_i w + m_i + \epsilon_i \end{aligned} \tag{5.1}$$

where the satellite clock offset δt^{sat} includes the relativistic clock correction and range ρ the relativistic path correction.

The aim of this chapter is to provide the background for the measurement modelling of the Standard and Precise Point Positioning. In applying this modelling, the user's receiver would correct each measurement from the known delay.

The residuals between the measured and predicted code or carrier pseudoranges contain the receiver position error and clock offset, together with mismodelling and measurement noise errors. These residuals, hereafter called *prefit residuals*, are the input data for the navigation equations studied in the next chapter. Such equations consist of a linear system whose solution by least squares or Kalman filtering allows us to decorrelate (i.e. separate) the different error components from the prefit residuals and then to determine the receiver coordinates and clock. The more accurate the measurement modelling, the better the decorrelation of errors and the more accurate the coordinates obtained.

[1] Where the term '$-\tilde{\alpha}_i K_{21}$' added in equations (4.19) to the carrier measurements (to have closed expressions for the combinations of measurements) is assimilated here into the unknown bias '$b_i + \lambda_i\, N_i$', for simplicity. Nevertheless, such closed expressions are especially useful when working with combinations of pairs of frequencies, as in section 6.3.2, to simplify the relationships between equations and parameters (see also sections 4.1.1.2 and 4.1.1.3).

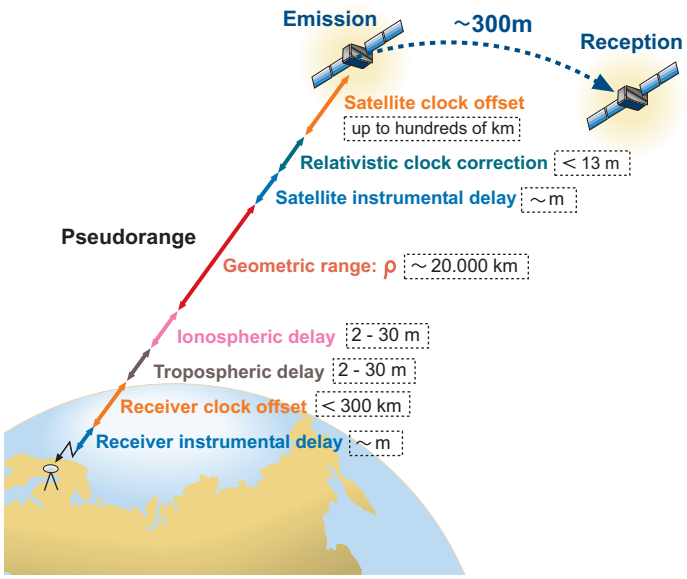


Figure 5.1: Pseudorange measurement contents.

5.1 Geometric Range Modelling

The geometric range ρ_{rcv}^{sat} is the Euclidean distance[2] between the satellite and receiver antenna phase centre coordinates at transmission and reception time, respectively:

$$\rho_{rcv}^{sat} = \left\| \mathbf{r}^{sat} - \mathbf{r}_{rcv} \right\| = \sqrt{(x^{sat} - x_{rcv})^2 + (y^{sat} - y_{rcv})^2 + (z^{sat} - z_{rcv})^2} \tag{5.2}$$

The algorithms for computing the transmission time from the measurement time, the satellite coordinates as well as the geometric range pre-fitting are as follows.

5.1.1 Satellite Coordinates

As the measurements are linked to the signal reception time, which is given by the receiver time tags (i.e. in the receiver clock), an algorithm is needed to obtain the signal emission time in the GNSS (GPS, Glonass, Galileo, etc.) time scale. Two algorithms will be presented next in section 5.1.1.1 for computing the transmission time, one of which is based on the pseudorange measurement and the other on a pure geometric approach.

Once the transmission time is obtained, the satellite coordinates can be computed.[3] This computation can be done in either an inertial or a rotating ECEF coordinate frame. Obviously, an ECEF frame is more suitable for providing the user's position on Earth and, therefore, this system is used as a reference frame for the satellite and receiver coordinates.

[2] The relativistic path range correction adds an extra amount (less than 2 cm) to this Euclidean range, see section 5.1.2.

[3] The satellite's speed is about 5 km/s and the travel time about 0.1 s; thus satellites move hundreds of metres during signal flight time, see Fig. 5.2.

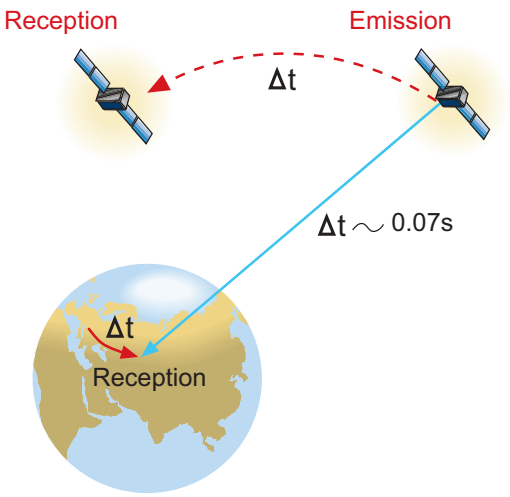


Figure 5.2: Coordinates at emission and reception time.

The ECEF frame is taken at reception time because it is a common reference for all measurements.[4] Then, Earth's rotation during the signal travel time must be considered to account for the ECEF rotation during this elapsed time. An algorithm for this computation is given in section 5.1.1.2.

Finally, besides the satellite coordinates (at transmission time), the receiver coordinates (at reception time) are also needed to compute the geometric range (from equation (5.2)). In fact, the receiver coordinates are the unknowns to be determined in the navigation problem and can seem quite contradictory in needing some 'knowledge of them to estimate themselves' (i.e. to compute the receiver's position). As explained in section 6.1, a simple approach to solve the navigation problem is to initialise the positioning algorithm with an approximation of the user's position.[5] In this approach, the unknowns to be estimated are the deviations of such approximate user locations from the true coordinates, or, essentially the same, the correction to apply in order to obtain more accurate receiver coordinates.

Different options can be considered to obtain an initial value for the receiver coordinates, that is a user location a priori, to initialise the navigation algorithm. The simplest way would be just to use Earth's centre coordinates (i.e. $(0, 0, 0)$) or an arbitrary point on Earth's surface. After a few iterations, the solution will converge to the actual position of the receiver. A different approach that does not require any 'a priori' knowledge of the receiver's location is the Bancroft method presented in Appendix D.

[4] Note that reception time is common for all satellites, but not transmission time, which depends on the satellite–receiver range and, therefore, is different for each satellite.

[5] These approximate coordinates (x_0, y_0, z_0) or *'a priories'* are used to linearise the geometric range in the neighbourhood of this point (i.e. $\rho = \rho_0 + \nabla\rho|_{\mathbf{r}=\mathbf{r}_0} \cdot \Delta\mathbf{r}$, where $\Delta\mathbf{r} = [\Delta x, \Delta y, \Delta z]^T$). This is done in order to build a linear model for estimating the position, since the deviations from this nominal value of $\Delta x = x - x_0$, $\Delta y = y - y_0$, $\Delta z = z - z_0$ are the unknowns to be estimated, together with the receiver clock offset δt and other parameters as well (see Chapter 6).

5.1.1.1 Computation of the Emission Time

Two different algorithms for computing the satellite transmission time from the receiver measurement time are presented as follows. The first of them is based on using the pseudorange measurement, which is a link between the receiver time tags (i.e. the reception time in the receiver clock) and the satellite transmission time (in the satellite clock). The second one is a pure geometric algorithm, which does not require any receiver measurement; it only needs the satellite coordinates and an approximate position of the receiver.

A Pseudorange-Based Algorithm

The emission time can be directly obtained from the reception time by taking into account that the pseudorange R is a direct measurement of the time difference between both epochs, each one being measured in the corresponding clock:

$$R = c\left(t_{rcv}[reception] - t^{sat}[emission]\right) \tag{5.3}$$

So the signal emission time, measured with the satellite clock (t^{sat}), is given by

$$t^{sat}[emission] = t_{rcv}[reception] - \Delta t \tag{5.4}$$

where

$$\Delta t = R/c \tag{5.5}$$

Thus, if δt^{sat} is the satellite clock offset regarding the GNSS (GPS, Glonass, Galileo, etc.) time scale (see section 5.2), the transmission time $T[emission]$ in this system time scale can be computed from the receiver measurement time tags (t_{rcv}) as

$$T[emission] = t^{sat}[emission] - \delta t^{sat} = t_{rcv}[reception] - R/c - \delta t^{sat} \tag{5.6}$$

This equation has the advantage of providing the signal emission time directly, without iterative calculations, although it does need pseudorange measurements in order to connect both epochs.

The accuracy in determining $T[emission]$ is very high, and essentially depends on the δt^{sat} error. For instance, in the case of the GPS system it is less than 10 or 100 ns with S/A off and S/A on, respectively. This allows us to calculate the satellite coordinates with an error of less than one- tenth of a millimetre in both cases.[6]

This algorithm is implemented in the `gLAB` software provided with this book.

[6] The speed of the GPS, Glonass or Galileo satellites is a few kilometres per second.

A Purely Geometric Algorithm

The previous algorithm (equation (5.4)) provides the signal emission time tied to the satellite clock (t^{sat}). The next algorithm ties this epoch to the receiver clock (t_{rcv}):

$$t_{rcv}[emission] = t_{rcv}[reception] - \Delta t \tag{5.7}$$

where Δt is now calculated by iteration assuming that an approximate receptor position $\mathbf{r}_{0_{rcv}}$ is known (it converges very quickly).

The algorithm is based on the following steps:

1. Calculate the position $\mathbf{r}^{sat}$ of the satellite at signal reception time t_{rcv}.
2. Calculate the geometric distance between the satellite coordinates obtained previously and the receiver's position[7] and, from it, calculate the signal travel time between both points:
$$\Delta t = \frac{\left\| \mathbf{r}^{sat} - \mathbf{r}_{0_{rcv}} \right\|}{c} \tag{5.8}$$
3. Calculate the satellite's position at the time $t = t_{rcv} - \Delta t \Longrightarrow \mathbf{r}^{sat}$.
4. Compare the new position $\mathbf{r}^{sat}$ with the former position. If they differ by more than a certain threshold value, iterate the procedure starting from step 2.

Finally, the emission time at the system time scale is given by[8]

$$T[emission] = t_{rcv}[emission] - \delta t_{rcv} \tag{5.9}$$

where δt_{rcv} is the receiver clock offset referred to the system time and that may be obtained from a navigation solution (although 'a posteriori').

Comments: This algorithm for calculating satellite coordinates at the reception epoch possesses an efficient modularity because pseudorange measurements are not needed to compute the transmission time.

If the receiver clock offset is small,[9] then δt_{rcv} may be neglected. On the other hand, the receiver clock estimates from the navigation solution can be used (extrapolated from the previous epoch). In any case, it must be taken into account that neglecting this term when δt_{rcv} reaches large values (e.g. 1 ms) may introduce errors up to the meter level in the satellite coordinates, and this must be considered when building the navigation model;[10] or, more

[7] Again, note that the satellite and receiver coordinates must be given in the same reference system, because the satellite–receiver ray must be generated in a common reference frame, see next section 5.1.1.2.

[8] Rigorously, equation (5.9) is $T[emission] = f(T[reception]) = f(t_{rcv}[reception] - \delta t_{rcv}) \simeq t_{rcv}[emission] - \delta t_{rcv}$ where function $f(\cdot)$ represents the geometric algorithm.

[9] Some receivers apply clock steering to adjust their clocks epoch by epoch and provide offsets of a few nanoseconds. However, in many cases the receiver waits until it has gathered an offset of 1 ms before adjusting the clock.

[10] That is, in the 'design matrix' or Jacobian matrix, obtained when linearising the model with respect to coordinate and receiver clock errors, see section 6.1 and Appendix E.

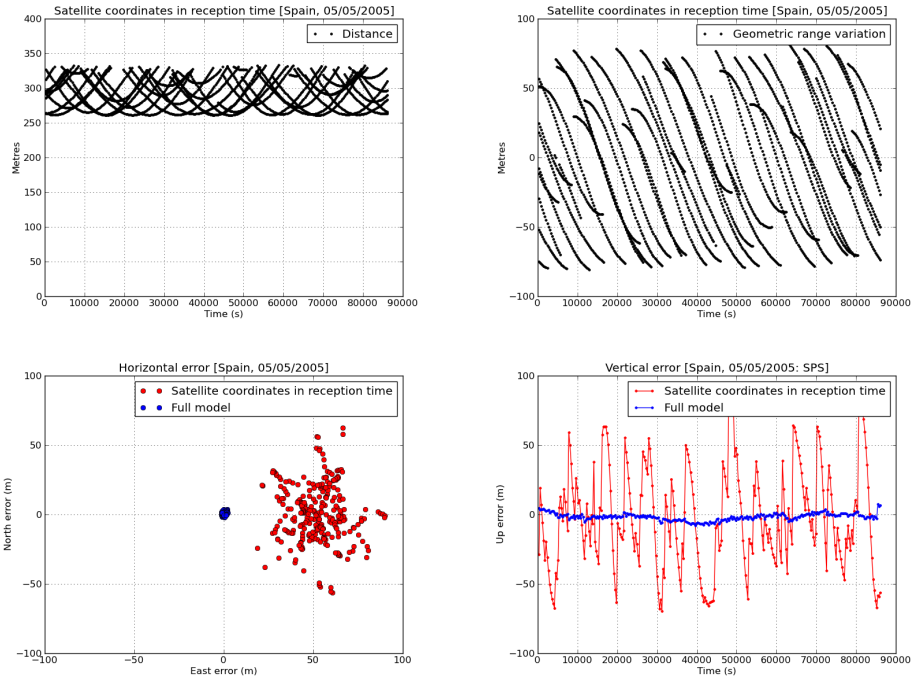


Figure 5.3: Effect of using the satellite coordinates at reception time instead of transmission time in positioning (see exercise 1.h of session 5.2 in Volume II and [Sanz et al., 2010]).

precisely, with respect to the partial derivative of the receiver clock in the design matrix.

Figure 5.3 illustrates the effect of neglecting the travel time in the satellite coordinate computation for positioning. It corresponds to a receiver located in Barcelona, Spain (receiver coordinates $\varphi \simeq 41°$, $\lambda \simeq 2°$). During the 70 to 90 ms of travel time, the satellite moves about 300 m, which leads to ± 60 m in range. The effect on the user's position is up to 50 m or more in the horizontal and vertical components.

5.1.1.2 Computation of Satellite Coordinates

Once the signal transmission time is known, the satellite coordinates may be calculated at that epoch in a given reference frame (either the inertial or ECEF rotating frame). Algorithms (and the associated software code) for computing such coordinates in an ECEF reference frame, from the broadcast navigation message, were given in section 3.3.1 for the GPS, Galileo or Beidou satellites, and in section 3.3.2 for the Glonass satellites. Algorithms using precise orbits can be found in section 3.3.3.

As mentioned previously, it must be taken into account that an ECEF frame is an Earth-fixed system and therefore rotates with Earth. This system must be taken at the signal reception time (in the GNSS time scale), because it is a common reference for all measurements.

An algorithm for computing satellite coordinates at transmission time, but referred to the ECEF tied to Earth at reception time, is as follows:

1. Calculate the satellite coordinates at emission time in the associated ECEF reference frame (i.e. tied to emission time).

Indeed, assume that the software routine `orb` does this computation;[11] then

$$T[emission] \Longrightarrow [\texttt{orb}] \Longrightarrow \widetilde{\mathbf{r}}^{sat} \tag{5.10}$$

Note again that the computed coordinates are given in the ECEF frame tied to Earth at emission time. Thus, Earth's rotation while the signal travels from the satellite to the receiver must be taken into account to transform such coordinates to the adopted common ECEF frame at reception time. This is done in the next step.

2. Transform the satellite coordinates from the system tied to Earth at 'emission time' to the system tied to Earth at 'reception time'(which is common for all measurements). In order to do this, consider Earth's rotation during the time interval Δt that the signal takes to propagate from the satellite to the receiver:

$$\mathbf{r}^{sat} = \mathbf{R}_3 \left(\omega_E \Delta t\right) \cdot \widetilde{\mathbf{r}}^{sat} \tag{5.11}$$

where ω_E is Earth's rotation rate and $\mathbf{R}_3[\theta]$ is a matrix defining a rotation of angle θ around the z-axis:

$$\mathbf{R}_3[\theta] = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.12}$$

and

$$\Delta t = \frac{\left\| \widetilde{\mathbf{r}}^{sat} - \mathbf{r}_{0_{rcv}} \right\|}{c} \tag{5.13}$$

where $\mathbf{r}_{0_{rcv}}$ is an approximate position of the user's receiver.

Note that it is advisable to calculate Δt using the former expression, even when the pseudorange method is used to find the signal propagation time (equation (5.5)). The reason for this is that R/c includes other delays (clock offsets etc.) besides the purely geometric part ρ/c. In other words, R/c establishes a very precise link between reception time (in the receiver time tags) and transmission time (in the satellite clock). Nevertheless, as a geometric distance it can be biased (mainly by the satellite and receiver clock offsets) by hundreds of kilometres.

Figure 5.4 illustrates the effect of neglecting Earth's rotation (equation (5.11)) during the travel time from satellite to receiver for the same example in Fig. 5.3. The effect on range is up to 20 m. As expected, this is seen in the position, basically, as a rotation to the east, moving the receiver about 25 m (receiver coordinates $\varphi \simeq 41°$, $\lambda \simeq 2°$).

[11] For broadcast GPS, Galileo or Beidou orbits, the `orb` program corresponds to `sub_orbit.f`, see section 3.3.1. For the Glonass broadcast orbits, it corresponds to `GLOeph2sp3.f`, see section 3.3.2. For GNSS precise orbits, `orb` is a Lagrange interpolator of satellite coordinates, as well, see section 3.3.3.

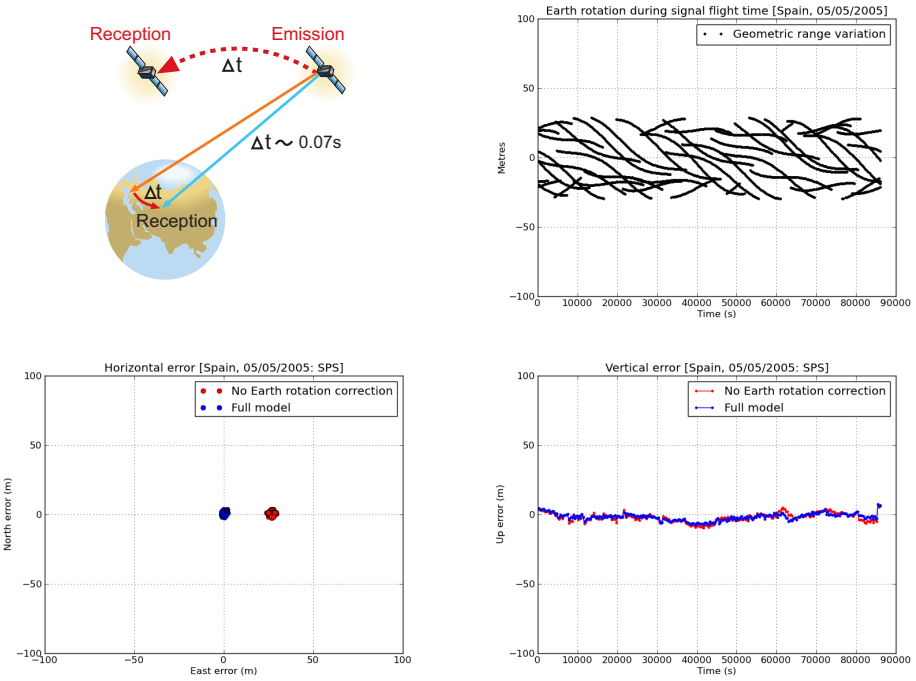


Figure 5.4: Effect of Earth's rotation (see exercise 1 of session 5.2 in Volume II).

5.1.1.3 Satellite Eclipses

High-accuracy GNSS positioning degrades during the GNSS satellites' eclipse seasons. Indeed, once the satellite goes into shadow, the radiation pressure vanishes. This effect introduces errors in the satellite dynamics due to the difficulty of properly modelling the solar radiation pressure. On the other hand, the satellite's solar sensors lose sight of the Sun and the attitude control (trying to keep the panels facing the Sun) is degraded. The yaw attitude of GPS satellites is essentially random during an eclipse and for up to 30 min after exiting the shadow [Bar-Sever, Y.E., 1994]. Every GPS satellite has two eclipse seasons per year, each lasting for about seven weeks [Mervat, 1995].

As a consequence, the orbit during shadow and eclipse periods may be considerably degraded and the removal of satellites under such conditions can improve the high-precision positioning results.

The eclipse condition can be defined by a simple cylinder model for Earth's shadow (see Fig. 5.5) [Mervat, 1995]:

$$\cos\phi = \frac{\mathbf{r}^{sat} \cdot \mathbf{r}_{sun}}{r^{sat}\, r_{sun}} < 0 \quad \text{and} \quad r^{sat}\sqrt{1-\cos^2\phi} < a_E \tag{5.14}$$

where $\mathbf{r}^{sat}$ and r^{sat} are, respectively, the vector from the geocentre to the satellite and the magnitude of that vector; $\mathbf{r}_{sun}$ is the geocentre vector to the Sun, and a_E is the mean equatorial radius of Earth (see Table 3.2 or 3.3).

Note that, as mentioned above, the satellites should be removed during the eclipse condition and also for the first 30 min, at least, after exiting the eclipse.

The eclipse condition is implemented in `gLAB`.

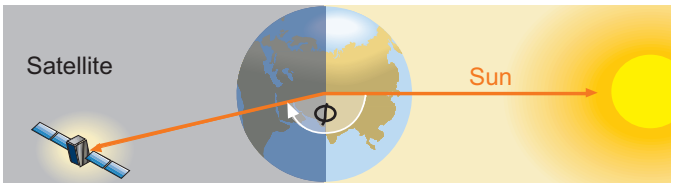


Figure 5.5: Simple cylinder model for the shadow of Earth.

5.1.2 Relativistic Path Range Correction

This is a secondary relativistic effect that might be required only for high-accuracy positioning. Its net effect on range is less than 2 cm and therefore, for most purposes, it can be neglected.

The effect is called the Shapiro signal propagation delay and it introduces a general relativistic correction into the geometric range. Due to the space–time curvature produced by the gravitational field, the Euclidean range computed by equation (5.2) must be corrected by an amount given by the expression

$$\Delta\rho_{rel} = \frac{2\,\mu}{c^2} \ln \frac{r^{sat} + r_{rcv} + r^{sat}_{rcv}}{r^{sat} + r_{rcv} - r^{sat}_{rcv}} \tag{5.15}$$

where r^{sat}, r_{rcv} are the geocentric distances of the satellite and receiver and r^{sat}_{rcv} is the distance between them. The constants c and μ are the speed of light and Earth's gravitational constant, respectively (see Table 3.2).

This correction must be added to the Euclidean distance given by equation (5.2). The effect is implemented in the `gLAB` tool.

Figure 5.6 illustrates an example of the Shapiro signal propagation delays for satellites in view from a receiver in Barcelona, Spain (receiver coordinates $\varphi \simeq 41°$, $\lambda \simeq 2°$).

A very good review of relativistic effects on GPS can be found in [Ashby, N., 2003] (*http://relativity.livingreviews.org/Articles/lrr-2003-1*).

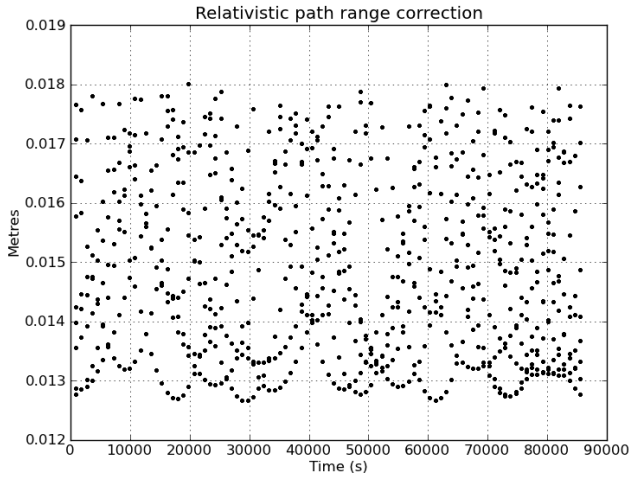


Figure 5.6: Shapiro correction to the geometric range for satellites in view from a receiver at coordinates $\phi \simeq 41°$, $\lambda \simeq 2°$. See exercise 2f of session 5.2 in Volume II.

5.2 Clock Modelling

The clock offsets are due to clock synchronisation errors referring to the GNSS time scale. The modelling of such offsets, as well as its effect on the navigation solution, are described as follows:

Receiver clock offset (δt_{rcv}): This is estimated together with the receiver coordinates, hence no modelling is needed in this case.

Satellite clock offset (δt^{sat}): This can be split into two terms:[12]

$$\delta t^{sat} = \widetilde{\delta t}^{sat} + \Delta_{rel} \tag{5.16}$$

The first term ($\widetilde{\delta t}^{sat}$) can be calculated from values broadcast in the navigation messages (for SPP) or from the precise products available from IGS centres or other providers (for PPP), see section 3.3.3. The second term (Δ_{rel}) is a small relativistic correction caused by the orbital eccentricity, see equation (5.19).

The broadcast navigation message provides the clock information as the coefficients of a polynomial, in a given reference epoch (t_0), to compute the satellite clock offset as

$$\widetilde{\delta t}^{sat} = a_0 + a_1(t - t_0) + a_2(t - t_0)^2 \tag{5.17}$$

In the case of GPS, Galileo or Beidou satellites, the satellite clock offset (a_0), clock drift (a_1) and clock drift rate (a_2) for a second-order polynomial (see Table 3.8) are broadcast. In Glonass, as the message is updated every half hour, only a first-order polynomial is considered, that is ($a_0 = -\tau_n$) the clock offset and ($a_1 = \gamma_n$) the relative frequency offset, see Table 3.9.[13]

The accuracy of broadcast clocks is of the order of some nanoseconds.[14]

The precise GPS and Glonass satellite clocks provided by IGS are accurate to the order of 0.1 ns or better,[15] see Table 3.10.

Figure 5.7 illustrates the effect of neglecting the GPS satellite clock offsets on the user's position, for the same example in Fig. 5.3. As shown, the satellite clocks reach up to more than 150 km, which leads to huge position errors.

[12] Glonass satellites transmit Δ_{rel} within the satellite clock corrections $\widetilde{\delta t}^{sat}$.
[13] Note that, in the RINEX files, the $-\tau_n$ value is given instead of τ_n.
[14] With the GPS selective availability activated they can be degraded by up to more than 100 ns.
[15] Note that 1 ns of error means 30 cm in range.

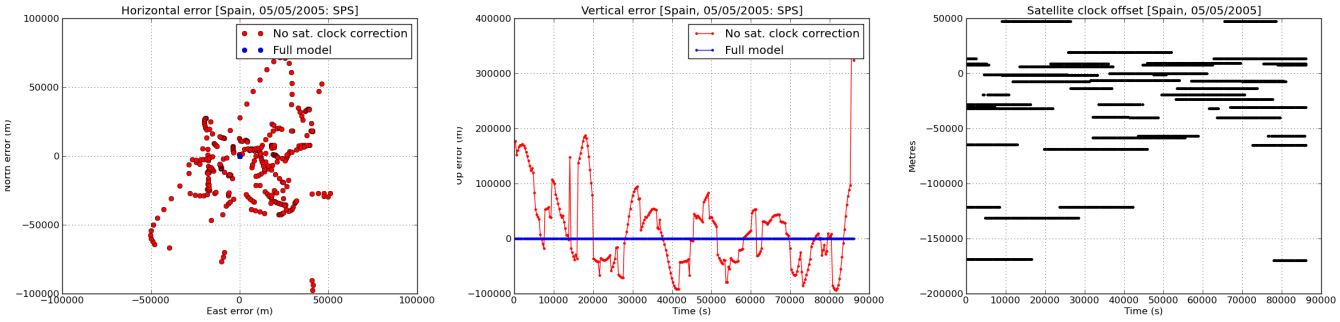


Figure 5.7: Satellite clocks: range and position domain effect. Left and middle panels show the horizontal and vertical positioning errors, respectively, using (blue) or not using (red) the satellite clock offsets. The variation in range is shown in the right panel. See exercise 1.d of session 5.2 in Volume II.

5.2.1 Relativistic Clock Correction

The rate of advance of two identical clocks, one placed in the satellite and the other on the terrestrial surface, will differ due to the difference in the gravitational potential (general relativity) and to the relative speed between them (special relativity). This difference can be split into the following [Ashby, N., 2003]:

- A constant component that depends only on the nominal value of the semi-major axis of the satellite orbit, which is adjusted by modifying (in the factory) the clock oscillating frequency of the satellite:[16]

$$\frac{f_0' - f_0}{f_0} = \frac{1}{2}\left(\frac{v}{c}\right)^2 + \frac{\triangle U}{c^2} \simeq -4.464 \cdot 10^{-10} \tag{5.18}$$

- A periodic component due to the orbital eccentricity (that must be applied by the user's receiver software)

$$\Delta_{rel} = -2\,\frac{\mathbf{r}^{sat} \cdot \mathbf{v}^{sat}}{c^2} \tag{5.19}$$

where $\mathbf{r}^{sat}$ and $\mathbf{v}^{sat}$ are the satellite position (m) and velocity (m/s) vectors in an inertial system.[17] The scalar product $\mathbf{r}^{sat} \cdot \mathbf{v}^{sat}$ can be evaluated either in a CRS or TRS (i.e. ECEF system).

Notice that in an ECEF system, Earth's rotation effect $\boldsymbol{\omega}_E \times \mathbf{r}^{sat}$ *should be discounted from* $\mathbf{v}^{sat}$*, but it cancels in the scalar product with* $\mathbf{r}^{sat}$*.*

[16] Since $f_0 = 10.23\,\text{MHz}$, one has $\Delta f_0 = 4.464 \cdot 10^{-10}\, f_0 = 4.57 \cdot 10^{-3}\,\text{Hz}$, thus the satellite must use $f_0' = 10.229\,999\,995\,43\,\text{MHz}$. Note that f_0' is the frequency 'emitted' by the satellite and f_0 is the one 'received' on the terrestrial surface (i.e. an apparent increase of the frequency is $4.57 \cdot 10^{-3}\,\text{Hz}$). That is, the clock on the satellite appears to run faster ($\simeq 38\,\mu\text{s/day}$) than on the ground (note: $\Delta f/f = \Delta T/T$). This effect is corrected (in the factory) by decreasing the oscillating frequency of the satellite by the amount $4.57 \cdot 10^{-3}\,\text{Hz}$.

[17] Note that, over the osculating orbit (see section 3.2.3), $\sqrt{\mu a}\, e \sin E = \mathbf{r} \cdot \mathbf{v}$. Thus, $\Delta_{rel} = -2\,\mathbf{r} \cdot \mathbf{v}/c^2 = -2\,(\sqrt{\mu a}/c^2)\, e \sin E$, where μ is the gravitational constant, c the speed of light in a vacuum, a and e the semi-major axis and eccentricity of the osculating orbit (see Table 3.2) and E the eccentric anomaly. See exercise 5f of session 5.2 in Volume II.

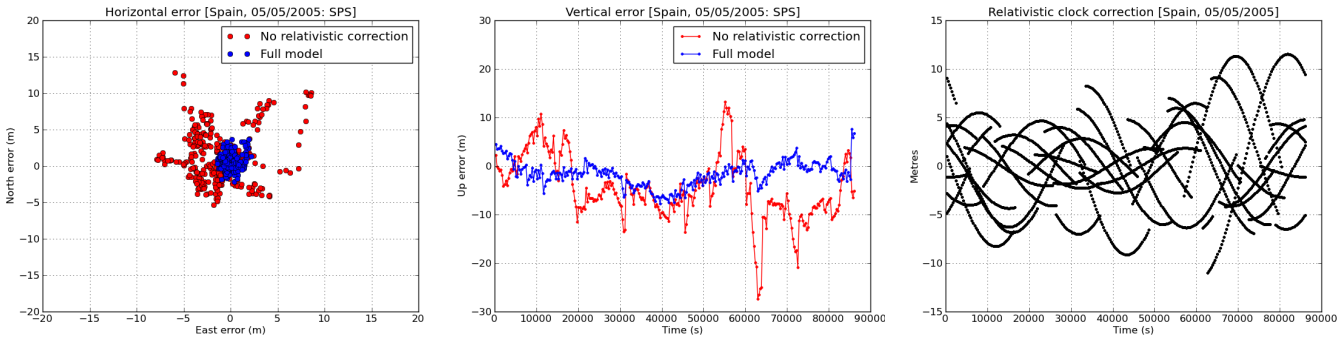


Figure 5.8: Relativistic correction: range and position domain effect. Left and middle panels show the horizontal and vertical positioning error, respectively, using (blue) or not using (red) the relativistic correction (5.19). The variation in range, in metres, is shown in the right panel. See exercise 1b of session 5.2 in Volume II.

Unlike in GPS, Galileo or Beidou relativistic corrections to Glonass orbital eccentricity are transmitted within the navigation message as satellite clock corrections (τ_n, γ_n). Therefore, equation (5.19) is not needed with such a broadcast message [Ashby, N., 2003].

All the previous corrections, that is satellite clock offset from broadcast or precise clocks and relativistic correction (5.19), are implemented in the `gLAB` software.

Figure 5.8 illustrates the effect of neglecting the relativistic correction given by equation (5.19) on the user's position, for the same example as in Fig. 5.3. As was shown, range errors up to 13 m and vertical errors over 20 m can be experienced when neglecting this correction.

5.3 Instrumental Delays

Possible sources of these delays are antennas and cables, as well as different filters used in receivers and satellites. These instrumental delays affect both code and carrier measurements.

The receiver instrumental delay is assimilated in the receiver clock. That is, since the delay is common for all satellites, it is assumed to be zero and is included in the receiver clock estimate.

As explained in section 4.1, the satellite clocks (broadcast or precise) are referred to the ionosphere-free combination of codes (R_C) and hence the instrumental delays cancel in such a combination of two-frequency signals.

For single-frequency users, the satellites broadcast in their navigation messages the Timing Group Delay or Total Group Delay (TGD), which is proportional to the DCB, or interfrequency bias, between the two codes involved in such an R_C combination as $K_{21} \equiv K_2 - K_1$, see equations (4.13) and (4.19). That is,

$$\mathrm{TGD}_1 = \frac{-1}{\gamma_{12} - 1}(K_2^{sat} - K_1^{sat}) = -\tilde{\alpha}_1 \, K_{21}^{sat} \tag{5.20}$$

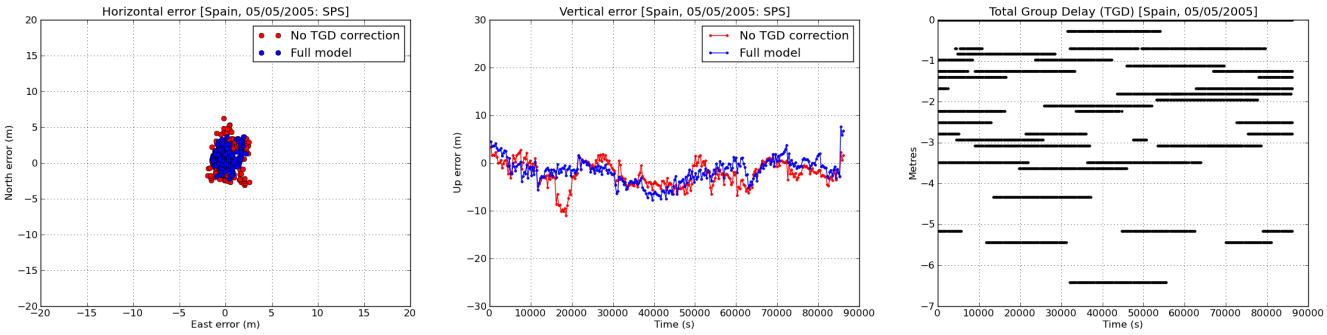


Figure 5.9: Total Group Delay (TGD): range and position domain effect. Left and middle panels show the horizontal and vertical positioning error, respectively, using (blue) or not using (red) the broadcast TGDs (equation (5.20)). The variation in range is shown in the right panel. See exercise 1c of session 5.2 in Volume II.

As the instrumental delays cancel in the R_C combination, the TGDs for the two associated codes are related by the square of their signal frequencies:

$$\mathrm{TGD}_2 = \gamma_{12}\,\mathrm{TGD}_1 \tag{5.21}$$

It must be pointed out that the instrumental delay depends not only on the signal frequency but also on the code. For example, there is a DCB between the C1 and P1 GPS codes, and therefore the DCB between the P2 and P1 codes is different to that between the P2 and C1 codes. The biases between the C1 and P1 codes for the different GPS satellites are provided, for instance, by IGS centres (see exercise 8 of laboratory session 5.1 in Volume II for an example of P1C1 DCB estimation using Ashtech Z-X12 receiver measurements).

The Galileo navigation messages F/NAV and I/NAV broadcast TGDs for the code on frequency E1, associated with the ionosphere-free combinations of codes at the frequencies E5a, E1 and E5b, E1, respectively. Neither TGDs nor any ionospheric model are broadcast in the Glonass navigation message.

Among the navigation messages, DCBs are also provided by IGS centres, together with Global Ionospheric Maps (GIM) files. In this regard, it should be pointed out that, as explained in section 4.1.1.3.3, there is a correlation between the DCBs and the ionospheric estimates. Indeed, the DCBs are associated with the ionospheric model used to compute such values. Thus, the DCBs broadcast in the GPS navigation message must be used with the Klobuchar ionospheric model (see section 5.4.1.2.1), and the DCBs of IGS with its associated GIMs.

Figure 5.9 illustrates the values of the GPS P1 code TGDs for the same example as in Fig. 5.3. Moreover, the effect of neglecting such delays in SPP for the horizontal and vertical error components is depicted by comparing the navigation solution using (blue) and not using (red) such TGDs.

Finally, the carrier phase instrumental delays are assimilated into the unknown ambiguities, which are estimated as real numbers (floating ambiguities) when positioning in PPP (see section 6.2). Figure 6.4 in section 6.3.2 shows the fractional part of the wide-lane and L1 ambiguities (i.e. $b = B - \lambda N$, see equations (4.19)), for a GPS satellite and for a receiver. These carrier instrumental delays cancel in the double differences between

satellites and receivers and thus are not needed for fixing ambiguities in differential mode. Nevertheless, they are needed (the satellite ones) for undifferenced ambiguity fixing, for example ambiguity fixing for PPP (see section 6.3.2).

5.3.1 Code Measurements and Receiver Types

The DCB handling for the C1, P1 and P2 code measurements collected from GPS receivers implementing different tracking technologies is discussed next.

As explained in section 2.2.1, under Anti-Spoofing (A/S) conditions, the GPS P codes (P1, P2) are encrypted to (Y1, Y2) for unauthorised users. Nevertheless, many commercial receivers are able to provide P1, P2 measurements among the C1 ones. The generation of such P1, P2 codes under A/S conditions depends on the code tracking technology, which must be taken into account for correct DCBs handling.

For instance, cross-correlated receivers (e.g. Rogue and Trimble 4000 models) provide a synthetic P2 code generated from C1 code and the cross-correlation of encrypted Y2 − Y1 codes (roughly speaking, P2 := C1 + [Y2 − Y1]). This leads to a P2 code measurement associated with C1, not P1, which then must be corrected by the Differential Code Bias $\mathrm{DCB_{P1\text{-}C1}}$ to get a consistent P2 measurement.

On the other hand, these receivers do not provide the P1 code, just the C1 code. Thus, the $\mathrm{DCB_{P1\text{-}C1}}$ must be added to this C1 code to emulate a consistent P1 code measurement. In this way, the interfrequency bias $\mathrm{DCB_{P2\text{-}P1}}$ will cancel when computing the PC (i.e. the Code ionosphere-free combination) of such P1 and P2 codes.

Three different kinds of receivers are usually considered (for more details see [Schaer, S. and Steingenberger, P., 2006]):

Type 1: Cross-correlated receiver
The C1 and P2 measurements must be corrected by the $\mathrm{DCB_{P1\text{-}C1}}$:

$$\boxed{\begin{aligned} \mathtt{C1}_{\mathtt{raw}} + \mathtt{DCB}_{\mathtt{P1-C1}} &\longrightarrow \mathtt{P1} \\ \mathtt{P2}_{\mathtt{raw}} + \mathtt{DCB}_{\mathtt{P1-C1}} &\longrightarrow \mathtt{P2} \end{aligned}}$$

Type 2: Receivers reporting C1 in place of P1
The C1 must be corrected by the $\mathrm{DCB_{P1\text{-}C1}}$:

$$\boxed{\mathtt{C1}_{\mathtt{raw}} + \mathtt{DCB}_{\mathtt{P1-C1}} \longrightarrow \mathtt{P1}}$$

Type 3: Receivers reporting L1, L2, P1, P2 as a consistent set
No bias removal is needed.

A RINEX conversion utility `cc2noncc.f` is provided by IGS to easily make code measurements consistent with P1/P2 data by applying satellite-dependent P1−C1 bias corrections. This tool is available at *https://goby.nrl.navy.mil/IGStime/cc2noncc/cc2noncc.f*.

These DCBs corrections are only required for geodetic (high-accuracy) surveying, as its impact on coordinate accuracy is at the level of a few millimetres (see exercise 7 of laboratory session 5.1 in Volume II).

This DCBs handling is implemented in `gLAB`, which is able to process the IGS files `P1C1YYMM.DCB` available from the *ftp://ftp.unibe.ch/aiub/CODE* site.

5.4 Atmospheric Effects Modelling

These effects include the excess GNSS signal path due to the variation of the refractive index in the atmosphere and, mainly, the delays/advances in the signal due to its propagation speed in the atmosphere (ionosphere and troposphere).

Electromagnetic Beam Bending

As is well known, the density of the atmospheric gas and plasma is not homogeneous. This causes spatial and temporal variations in the refractive index. Due to the refractive index gradients, the rays have an extended geometric path, regarding the free space path, when travelling through the atmosphere. Snell's law states that if an electromagnetic wave travels from an environment with refractive index n_1 to a second one with refractive index n_2 and crosses the distance between them at an angle φ_1, it will be deviated by an angle φ_2. The relation between these four values is summarised by the expression $n_1 \cdot \sin\varphi_1 = n_2 \cdot \sin\varphi_2$.

Figure 5.10 (left) illustrates this basic postulate: $\varphi_2 < \varphi_1$ because $n_2 > n_1$. The consequences of this law can be seen in Fig. 5.10 (right), where the successive paths through different layers, each with its own refractive index, causes the ray path to bend and thus causes an excess of geometric path as discussed above.

For the troposphere [Parkinson et al., 1996], the effect caused by neglecting path curvature is less than 3 mm for elevations greater than 20°, 2 cm for elevations close to 10°, and increases to 17 cm at elevations close to 5° (mainly due to dry gases).

For the ionosphere, these errors can reach up to a few centimetres for L1 and L2 GPS signals, at low elevations and under high solar activity conditions [Hoque and Jakowsky, 2008]. On the other hand the bending angle is proportional to the inverse of the squared frequency and mostly cancels out in the ionosphere-free combination [Hajj et al., 2002] (see also the tutorial [Sanz et al., 2012]).

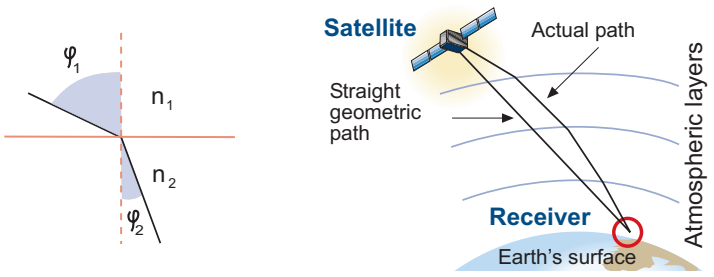


Figure 5.10: Snell's law and its effects over the ray path.

Propagation Speed: Phase and Group Velocity

A brief review of wave propagation concepts is given as follows, in order to introduce the definitions of phase and group velocity.

A pure sinusoidal monochromatic signal travelling in the x direction can be described by (see Fig. 5.11, top)

$$s(x,t) = s_0 \cos(\omega t - k\,x + \phi_0) \tag{5.22}$$

where $\omega = 2\pi/T$, $k = 2\pi/\lambda$ and $T = 1/f = 2\pi/\omega$, with f, T and λ the frequency, period and wavelength, respectively.

Here, $s(x,t)$ is a function that depends on both time and position. At any fixed instant of time t the function varies sinusoidally along the x-axis, whereas at any fixed location of the x-axis the function varies sinusoidally with time.

The speed at which the shape is moving, that is the speed at which any fixed phase of the cycle is displaced, is called the *phase velocity*. This velocity, according to equation (5.22), is given by

$$v_{ph} = \frac{\omega}{k} \tag{5.23}$$

because $\omega t - kx + \phi_0$ remains constant by changing x and t when $x = (\omega/k)t$.

An amplitude-modulated signal can be generated by simply adding two sinusoidal signals at slightly different frequencies:

$$s_1(x,t) = \cos((\omega+\Delta\omega)t-(k+\Delta k)x), \quad s_2(x,t) = \cos((\omega-\Delta\omega)t-(k-\Delta k)x) \tag{5.24}$$

Using simple trigonometric identities, it is found that

$$s_1(x,t) + s_2(x,t) = 2\cos(\Delta\omega t - \Delta k\,x)\,\cos(\omega t - k\,x) \tag{5.25}$$

A plot of an example to illustrate the combination of signals of equation (5.25) is shown in Fig. 5.11.

From equation (5.25), it follows that two signals can appear to travel at two different velocities as a result of the superposition of signals $s_1(x,t)$ and $s_2(x,t)$. That is, in the bottom plot of Fig. 5.11, the internal oscillations (in blue) travel at velocity $v = \omega/k$, whereas the amplitude envelope (in red) travels at velocity $\Delta\omega/\Delta k$.

The velocity of the modulation $\Delta\omega/\Delta k$ when considering the limit $\Delta k \to 0$ can be written as

$$v_{gr} = \frac{d\omega}{dk} \tag{5.26}$$

Equation (5.26) defines the group velocity.

Note that, when the angular frequency ω and the wave number k are proportional, then the phase and group velocities are the same, that is $v_{gr} = v_{ph}$. On the other hand, from equation (5.23) it is obvious that, in this case, the phase velocity v_{ph} does not depend on the frequency. A medium in which this condition is fulfilled is said to be *non-dispersive*.

On the contrary, a medium in which the wave propagation speed varies with the frequency is said to be *dispersive* (from light dispersion by a prism).

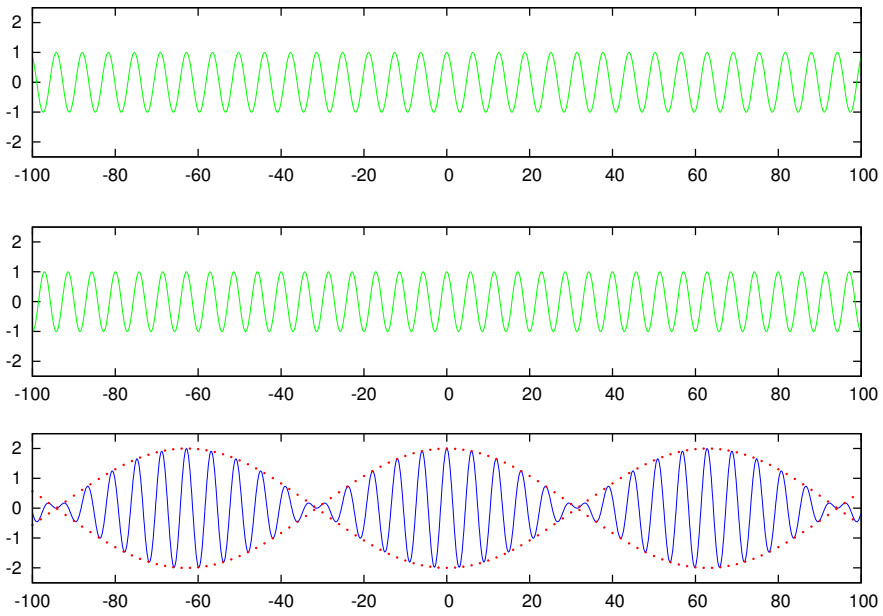


Figure 5.11: The signals of the top two plots, with slightly different frequencies, are added to generate the signal in the bottom plot.

Thus, the envelope travels at different speed v_{gr} than the internal oscillations v_{ph}. In this case, if we were riding along with the envelope in Fig. 5.11, we would observe the internal oscillations moving forward from one group to the next.

The *refractive index* (n) in a medium is the ratio between the propagation speed of a signal in a vacuum (c) to the speed in the medium (v). Hence, two different refractive indices can be defined for the group and phase:

$$n_{ph} = \frac{c}{v_{ph}}\,, \quad n_{gr} = \frac{c}{v_{gr}} \tag{5.27}$$

In the context of GNSS signals, the carrier waves are propagating with the phase velocity and so this is the velocity associated with the carrier measurements. The group velocity is what has to be considered for the code measurements (i.e. pseudoranges obtained from the codes modulated in the carriers).[18]

As explained in the following sections, the troposphere is a non-dispersive medium for radio waves at the GNSS frequencies. Hence, code and carrier phase measurements are associated with the same velocity in the troposphere. By contrast, codes and phases propagate at different speeds in the ionosphere, resulting in a delay in the code and an advance in the carrier measurements, see equation (5.37).

[18] Note that the information and energy are propagated in a wave as changes in the wave, for example by modulating the codes. Thus, the information propagates at the group velocity.

Atmospheric Refraction

The electromagnetic signals experience changes in velocity (speed and direction) when passing through the atmosphere due to the refraction.

According to Fermat's principle, the measured range l is given by the integral of the refractive index n along the ray path from the satellite to receiver:

$$l = \int_{\text{ray path}} n\, dl \tag{5.28}$$

Then, the signal delay can be written as

$$\Delta = \int_{\text{ray path}} n\, dl - \int_{\text{straight line}} dl \tag{5.29}$$

where the second integral is the Euclidean distance between the satellite and receiver.

Note that the previous definition includes both the signal bending and propagation delay. One simplification of the previous expression is to approximate the first integral along the straight line between the satellite and receiver:

$$\Delta = \int_{\text{straight line}} (n-1)\, dl \tag{5.30}$$

From the point of view of signal delay, the atmosphere can be divided into two main components: the neutral atmosphere (i.e. the non-ionised part), which is a non-dispersive medium; and the ionosphere, where the delay experienced by the signals depends on their frequency. The ionospheric and tropospheric delays are discussed in the next two sections.

It must be pointed out that the neutral atmosphere includes the troposphere and stratosphere, but the dominant component is the troposphere, and therefore the name of the delay usually refers only to the troposphere (as tropospheric delay).

5.4.1 Ionospheric Delay

The ionosphere is that part of the terrestrial atmosphere that extends from about 60 km up to more than 2000 km. As its name implies, it contains a partially ionised medium, as a result of solar X- and Extreme UltraViolet (EUV) rays in the solar radiation and the incidence of charged particles.

The propagation speed of GNSS electromagnetic signals in the ionosphere depends on its electron density (see below), which is typically driven by two main processes. During the day, the Sun's radiation ionises neutral atoms to produce free electrons and ions. During the night, the recombination process prevails, where free electrons are recombined with ions to produce neutral particles, which leads to a reduction in the electron density.

As commented above, a medium where the angular frequency ω and the wave number k are not proportional is a dispersive medium (i.e. the wave propagation speed and hence the refractive index depend on the frequency). This is the case with the ionosphere where ω and k are related, to a first

approximation, by [Crawford, 1968], section 4.3, example 7,

$$\omega^2 = c^2k^2 + \omega_p^2 \tag{5.31}$$

where c is the propagation speed of a signal in a vacuum and[19]

$$\omega_p = 2\pi f_p \quad \text{with} \quad f_p = 8.98\sqrt{N_e} \quad \text{in Hz} \tag{5.32}$$

where N_e is the electron density (in e^-/m^3). A complete derivation of this relationship can be found in [Davies, 1989] and the updated higher order terms in [McCarthy, D. and Petit, G., 2009], which is typically less than 0.1% of the total delay.

Equation (5.31) is called the *relation of dispersion* of the ionosphere, and ω_p is called the *critical frequency* of the ionospheric plasma, in the sense that signals with $\omega < \omega_p$ will be reflected and signals with $\omega > \omega_p$ will pass through the plasma [Davies, 1989].

The electron density in the ionosphere changes with height, having a maximum of $N_e \simeq 10^{11}$ to $10^{12}\,\mathrm{e}^-/\mathrm{m}^3$ around $300 - 500\,\mathrm{km}$. Then, according to expression (5.32), electromagnetic signals with $f > f_p \simeq 10^6\,\mathrm{Hz}$ will be able to pass through the ionosphere. This is the case for GNSS signals whose frequencies are of the order of 1 GHz ($= 10^9\,\mathrm{Hz}$). Radio signals with frequencies below f_p will be reflected in the ionosphere.

From equation (5.31), and taking into account that $\omega = 2\pi f$ and the definition in equation (5.23), it follows that

$$v_{ph} = \frac{c}{\sqrt{1 - (f_p/f)^2}} \tag{5.33}$$

Thus, according to equation (5.27), the phase refractive index of the ionosphere can be approximated[20] by

$$n_{ph} = \sqrt{1 - \left(\frac{f_p}{f}\right)^2} \simeq 1 - \frac{1}{2}\left(\frac{f_p}{f}\right)^2 = 1 - \frac{40.3}{f^2}N_e \tag{5.34}$$

At the frequency of GNSS signals, the previous approximation (5.34) accounts for more than 99.9% of the refractivity (first-order ionospheric effect). That is, with less than an error of 0.1%, it can be assumed that

$$n_{ph} = 1 - \frac{40.3}{f^2}N_e \tag{5.35}$$

[19] [Davies, 1989], page 94: $N_e = 1.24 \cdot 10^{10} f_p^2$, with f_p in MHz and N_e in electrons/m^3.

[20] The approximation $\sqrt{1-x^2} \simeq 1 - \frac{1}{2}x^2$ can be used when $|x| \ll 1$. Note that, on taking $f = 10^9\,\mathrm{Hz}$ as the frequency of the GNSS signals, and with $f_p \simeq 10^7\,\mathrm{Hz}$, then $x = f_p/f \simeq 10^{-2}$, which justifies the approximation.

Differentiating equation (5.31) with respect to k, and taking into account (5.23), (5.26), (5.27) and the approximation $(1-\varepsilon^2)^{-1/2} \simeq 1+\frac{1}{2}\varepsilon^2$, yields the group refractive index[21]

$$n_{gr} = 1 + \frac{40.3}{f^2} N_e \tag{5.36}$$

Substituting the phase and group refractive indices (5.35 and 5.36) in (5.30), the difference between the measured range (with frequency f signal) and the Euclidean distance between the satellite and receiver is given by

$$\Delta^{iono}_{ph,f} = -\frac{40.3}{f^2}\int N_e\, dl, \qquad \Delta^{iono}_{gr,f} = +\frac{40.3}{f^2}\int N_e\, dl \tag{5.37}$$

Therefore, phase measurements are advanced on crossing the ionosphere, that is a negative delay, while the code measurements undergo a positive delay.

The differences $\Delta^{iono}_{ph,f}$ and $\Delta^{iono}_{gr,f}$ are called the phase and code ionospheric refraction, respectively, and the integral is defined as the Slant TEC (STEC)

$$\text{STEC} = \int N_e\, dl \tag{5.38}$$

Usually, the STEC is given in TEC units (TECUs), where $1\,\text{TECU} = 10^{16}\,\text{e}^-/\text{m}^2$ and the ionospheric delay I_f (at the frequency f) is written as[22]

$$I_f \equiv \Delta^{iono}_{gr,f} = \alpha_f\,\text{STEC} \quad \text{(units: metres of delay)} \tag{5.39}$$

with

$$\alpha_f = \frac{40.3\cdot 10^{16}}{f^2}\ \text{m}_{\text{signal delay}_{(\text{at frequency}\,f)}}/\text{TECU} \quad \text{(where } f \text{ is in Hz)} \tag{5.40}$$

The TEC, and hence the ionospheric refraction, depend on the geographic location of the receiver, the hour of day and the intensity of the solar activity. Figure 5.12 (left) shows a vertical TEC map of the geographic distribution of the TEC, where the equatorial anomalies are clearly depicted around the geomagnetic equator. The figure on the right shows the 11-year solar cycle with a solar flux plot.

As the ionosphere is a dispersive medium, refraction of the GNSS signals depends on their frequencies (as the squared inverse). This dependence on the signal frequency allows us to remove its effect by up to more than 99.9% using two-frequency measurements (see section 5.4.1.1 next). However, single-frequency receivers have to apply an ionospheric prediction model to remove this effect (as much as possible), which can reach up to several tens of metres in size (see sections 5.4.1.2.1 and 5.4.1.2.2).

Several exercises depicting ionospheric effects on GNSS measurements in different ionospheric conditions are given in session 4 of Volume II. Additional exercises, including electron density profile retrieval and atmospheric bending analysis using radio-occultation measurements from receivers on LEO satellites, can be found in [Sanz et al., 2012].

[21] Notice that $n_{gr} > n_{ph}$ and, thus, from (5.27) it follows that $v_{gr} < v_{ph}$. Moreover, as $n_{ph} < 1$, then $v_{ph} = c/n_{ph} > c$. That is, the phase travels faster than the speed of light, but no information is carried, so no relativity principle is violated.

[22] In equations (5.1) or (4.19) the ionospheric delay I is given in metres of delay in the geometry-free combination Φ_1–Φ_2. That is, $I \equiv (\alpha_2 - \alpha_1)\,\text{STEC}$, see equation (4.16).

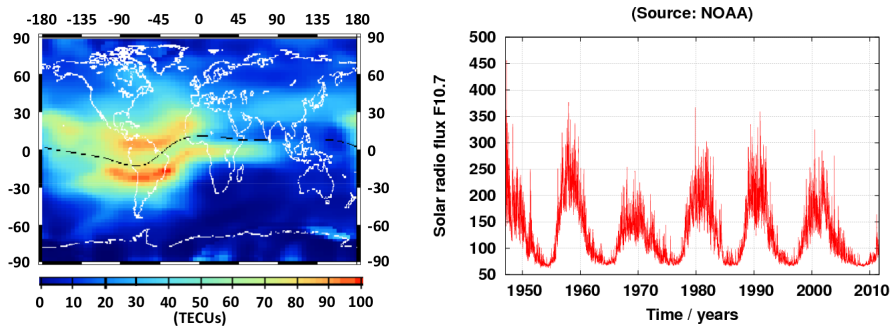


Figure 5.12: The map on the left shows the vertical Total Electron Content in TECUs at 19UT on 26 June 2000 (1 TECU $\simeq$ 16 cm of delay in the GPS L1 signal). The plot on the right shows the evolution of the solar flux during the last solar cycles.

Advanced Comment. Second-Order Ionospheric Effect: The previous results of equation (5.39) are often referred to as the *first-order* ionospheric effect, denoted as $I1_f \equiv I_f$. That is,

$$I1_{ph,f} = -\alpha_f \,\mathrm{STEC} = -I1_{gr,f} \tag{5.41}$$

As commented previously, equation (5.31) is a simplification of the relation of dispersion of the ionosphere given in [McCarthy, D. and Petit, G., 2009], where the dependence of the refractive index on the magnetic field **B** has been neglected. If such terms are taken into account, higher order ionospheric terms appear (but they represent less than 0.1% of the total effect).

The second-order ionospheric term, its implementation and its impact on geodesy are given in [HJS, 2007]. Its effect on carrier ($I2_{ph,f}$) and code ($I2_{gr,f}$) measurements at frequency f is given by

$$I2_{ph,f} = -\frac{7527\,c}{2\,f_1^3}\int_{rec}^{sat} N_e B\cos\theta\, dl \tag{5.42}$$

$$I2_{gr,f} = -2\; I2_{ph,f} \tag{5.43}$$

where c is the speed of light. B is the magnitude of Earth's magnetic field **B** and θ the angle between B and the propagation direction. The units are in the International System of Units (SI).

In global geodetic computations, $I2$ mainly affects the satellite clock estimate (centimetre level) and orbits (a few millimetres), but the impact on receiver position is usually less than 1 mm, see exercises 7 to 10 of session 4.3 in Volume II.

5.4.1.1 Ionosphere-Free Combination for Dual-Frequency Receivers

According to equations (5.37), the first-order ionospheric effects on code R_P and carrier phase Φ_L measurements depend (99.9%) on the inverse of squared signal frequency f. Therefore, the dual-frequency receivers can eliminate these effect through a linear combination of code or carrier measurements:

$$\Phi_{\text{iono-free}} = \frac{f_1^2\,\Phi_1 - f_2^2\,\Phi_2}{f_1^2 - f_2^2}, \qquad R_{\text{iono-free}} = \frac{f_1^2\,R_1 - f_2^2\,R_2}{f_1^2 - f_2^2} \tag{5.44}$$

This combination is called *ionosphere-free* as was presented in section 4.1.

It must be pointed out that the Precise Point Positioning (PPP) uses code and carrier phase measurements in the ionosphere-free combination to remove the ionospheric refraction, because it is one of the effects that is more difficult to model accurately. Moreover, the TGDs of the associated codes also cancel in this combination and are not needed (since the satellite clocks are referred to the ionospheric-free combination of such codes, see section 4.1.1).

5.4.1.2 Ionospheric Models for Single-Frequency Receivers

Single-frequency receivers need to apply a model to remove the ionospheric refraction, which can reach up to few tens of metres, depending on the elevation of rays and the ionospheric conditions.

The models broadcast by the GPS, Galileo and Beidou satellites are described in the next section. No ionospheric model is broadcast by the Glonass satellites, but any of the GPS, Galileo or Beidou models can be used for Glonass signals, by applying a correction factor given by their relative squared frequency ratio.

5.4.1.2.1 Klobuchar Model

GPS and Beidou satellites broadcast the parameters of the Klobuchar ionospheric model for single-frequency users.[23] The Klobuchar model was designed to minimise user computational complexity and user computer storage so as to keep a minimum number of coefficients to be transmitted on the satellite–user link.

This broadcast model, initially developed for GPS, is based on an empirical approach [Klobuchar, 1987] and is estimated to reduce the RMS ionospheric range error by about 50% worldwide. It is assumed that the electron content is concentrated in a thin layer at 350 km in height (375 km is taken in Beidou). Thus, the slant delay is computed from the vertical delay at the Ionospheric Pierce Point (IPP)[24] by multiplying by an obliquity factor (i.e. the mapping function), see Fig. 5.13 and equation (5.53).

Klobuchar Algorithm Equations

The Klobuchar algorithm to run in a single-frequency receiver is provided as follows [Klobuchar, 1987].

Given the user's approximate geodetic latitude φ_u and longitude λ_u, the elevation angle E and azimuth A of the observed satellite and the Klobuchar coefficients α_n and β_n broadcast in the GPS or Beidou satellite navigation message:[25]

[23] Vertical ionospheric delays on a $5^\circ \times 2.5^\circ$ grid covering 70°–145° in longitud and 10°–55° in latitude are also broadcast in the Beidou NAV message [BeiDou-SIS-ICD, 2012].

[24] That is, the intersection of the ray with the ionospheric layer at 350 km in height.

[25] Simplified equations to reduce the computational load are given in [IS-GPS-200, 2010].

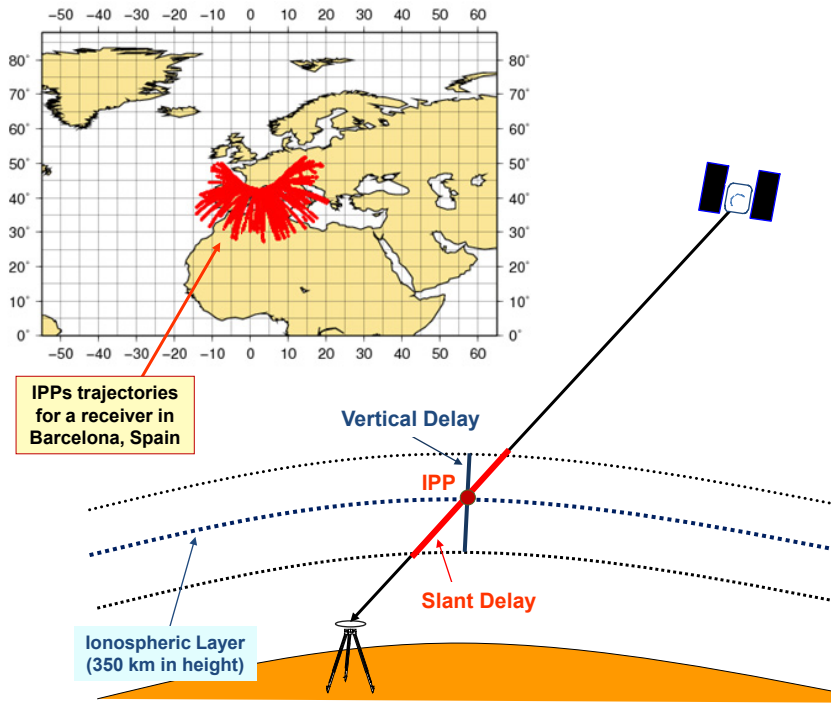


Figure 5.13: Ionospheric Pierce Points (IPPs), vertical and slant delay illustration. The IPP trajectories for the satellites in view from a receiver in Barcelona, Spain, are shown in the map at the top.

1. Calculate the Earth-centred angle[26]

$$\psi = \pi/2 - E - \arcsin\left(\frac{R_E}{R_E + h}\cos E\right) \tag{5.45}$$

2. Compute the latitude of the IPP[27]

$$\phi_I = \arcsin\left(\sin\varphi_u \cos\psi + \cos\varphi_u \sin\psi \cos A\right) \tag{5.46}$$

3. Compute the longitude of the IPP

$$\lambda_I = \lambda_u + \frac{\psi \sin A}{\cos\phi_I} \tag{5.47}$$

4. Find the geomagnetic latitude of the IPP

$$\phi_m = \arcsin\left(\sin\phi_I \sin\phi_P + \cos\phi_I \cos\phi_P \cos(\lambda_I - \lambda_P)\right) \tag{5.48}$$

with $\phi_P = 78.3°$, $\lambda_P = 291.0°$ the coordinates of the geomagnetic pole.

5. Find the local time at the IPP

$$t = 43\,200\,\lambda_I/\pi + t_{GPS} \quad (\lambda_I \text{ in radians, } t \text{ in seconds}) \tag{5.49}$$

where $0 \le t < 86\,400$. Therefore:
If $t \ge 86\,400$, subtract 86 400. If $t < 0$, add 86 400.

6. Compute the amplitude of ionospheric delay

$$A_I = \sum_{n=0}^{3} \alpha_n \left(\phi_m/\pi\right)^n \quad \text{(seconds)} \tag{5.50}$$

If $A_I < 0$, then $A_I = 0$.

[26] The height of the ionospheric layer is taken as 350 km for GPS and 375 km for Beidou. $R_E = 6\,378$ km.
[27] That is, the latitude of the projection on Earth of the IPP with the ionospheric layer.

7. Compute the period of ionospheric delay

$$P_I = \sum_{n=0}^{3} \beta_n \, (\phi_m/\pi)^n \quad \text{(seconds)} \tag{5.51}$$

If $P_I < 72\,000$, then $P_I = 72\,000$.

8. Compute the phase of ionospheric delay

$$X_I = \frac{2\pi(t - 50\,400)}{P_I} \quad \text{(radians)} \tag{5.52}$$

9. Compute the slant factor (ionospheric mapping function)

$$F = \left[1 - \left(\frac{R_E}{R_E + h} \cos E\right)^2\right]^{-1/2} \tag{5.53}$$

10. Compute the ionospheric time delay

$$I_1 = \begin{cases} \left[5 \cdot 10^{-9} + A_I \cos X_I\right] \times F, & |X_I| < \pi/2 \\ 5 \cdot 10^{-9} \times F, & |X_I| \geq \pi/2 \end{cases} \tag{5.54}$$

The delay I_1 is given in seconds and is referred to the GPS L1 or Beidou B1 frequencies.

Note that, I_1 corresponds to the term $\hat{\alpha}_1 \, I$ in equations (4.17), (4.19) or (5.1).

Although this algorithm is provided to estimate the ionospheric delay in the GPS L1 or Beidou B1 signals, it can also be used to estimate the ionospheric time delay in other frequency signals or for the Glonass and Galileo signals, as well. Indeed, taking into account that the ionospheric delay is inversely proportional to the square of the signal frequency, the delay for any GNSS signal transmitted on frequency f_k is given by

$$I_k = \left(\frac{f_1}{f_k}\right)^2 I_1 \tag{5.55}$$

The Klobuchar model is implemented in the `gLAB` software tool for SPP. It is also implemented in program `sub_klob.f` of Volume II.

Figure 5.15 is a layout of the Klobuchar model algorithm, showing that the vertical delays are based on a constant value of 5 ns at night-time and a half-cosine function in daytime, whose amplitude and period are given as a function of the eight parameters (α_i, β_i, $i = 0, \ldots, 3$) broadcast in the GPS and Beidou navigation messages. A map with the geographic variation of the vertical delay is also depicted in this figure.

Figure 5.14 illustrates with an example the values of the GPS delays from the Klobuchar model and shows the effect of neglecting such delays in single-point positioning for the horizontal and vertical error components (same data set as in Fig. 5.3).

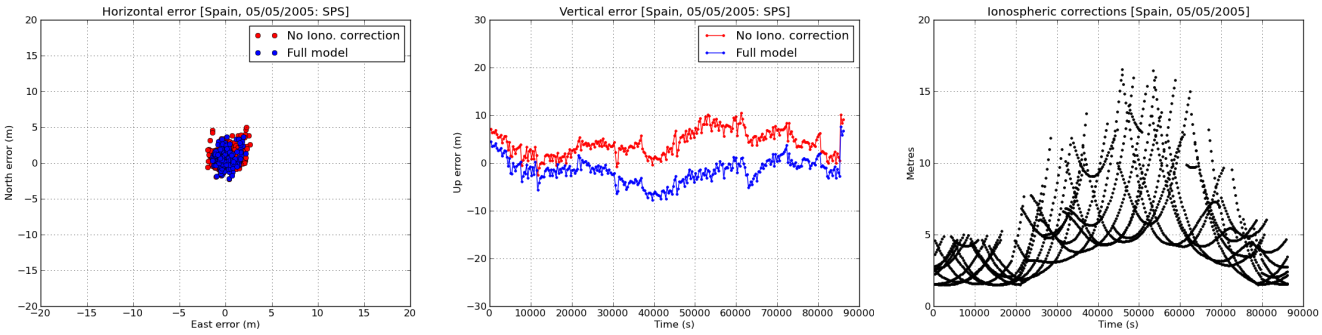


Figure 5.14: Ionospheric correction: range and position domain effect. Left and middle panels show the horizontal and vertical positioning error, respectively, using (blue) or not using (red) the Klobuchar ionospheric correction defined in section 5.4.1.2.1. The variation in range is shown in the right panel. See exercise 1e of session 5.2 in Volume II.

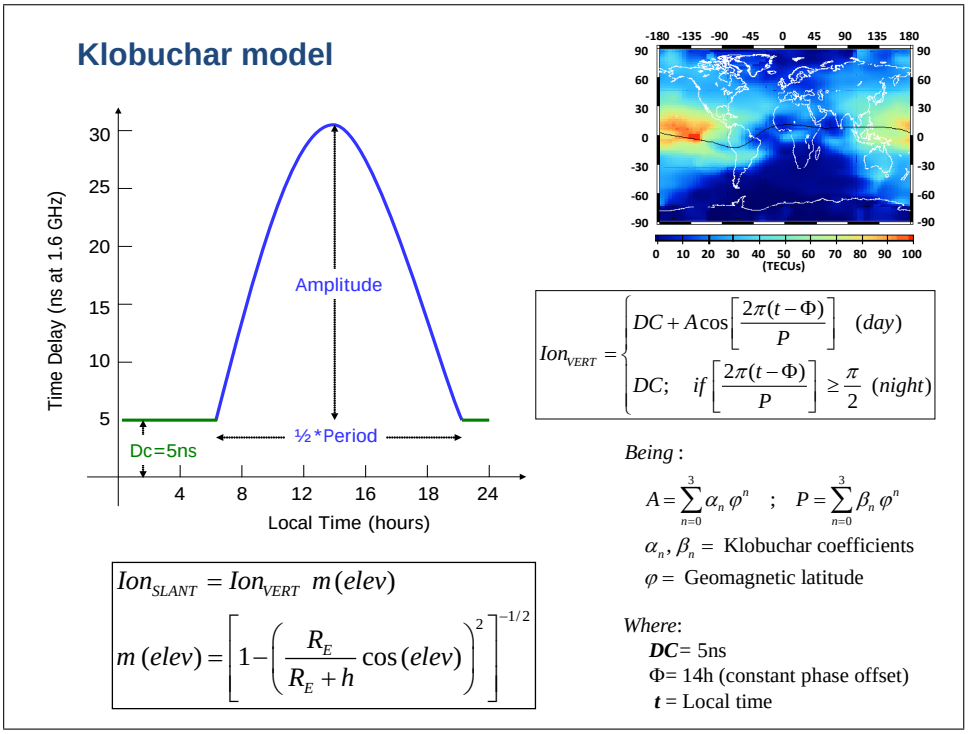


Figure 5.15: Klobuchar ionospheric model: layout of algorithm.

5.4.1.2.2 NeQuick Model

The NeQuick model is the ionospheric model proposed for use in the Galileo single-frequency receiver to compute ionospheric corrections. It is based on the original profiler developed by [Di Giovanni and Radicella, 1990].

NeQuick is a three-dimensional and time-dependent ionospheric electron density model, which provides the electron density in the ionosphere as a function of position and time. Hence, it allows ionospheric delays (TEC or STEC) to be computed as the integrated electron density along any ray path.

NeQuick FORTRAN 77 code was accepted by the International Telecommunications Union Radiocommunication Sector (ITU-R) in 2000 and revised in 2002. It is freely available from *http://www.itu.int/oth/R0A04000018/en*. It is referred to either as version 1 or ITU-R and constitutes the current baseline for Galileo. The package includes a comprehensive description of the implementation as well as numerical integration subroutines allowing vertical and slant TEC to be computed.

The input parameters of the model are the position (longitude, latitude and height), the epoch (month and UT) and the solar activity (expressed[28] as either F10.7 or R12). Other internal parameters are the foF2 and M(3000)F2 values, which can be defined according to the ITU-R, among other options depending on purpose.

The NeQuick model running in the Galileo single-frequency receivers is driven by the *effective ionisation level* Az parameter (replacing the solar flux) that is a function of the receiver's location

$$Az = a_0 + a_1\mu + a_2\mu^2 \tag{5.56}$$

where μ is the MOdified DIP latitude (MODIP),

$$\tan\mu = \frac{I}{\sqrt{\cos\varphi}} \tag{5.57}$$

with I the true magnetic inclination,[29] or *dip* in the ionosphere (usually at 300 km), and φ the geographic latitude of the receiver [Rawer, 1963].

The coefficients a_0, a_1, a_2 will be broadcast to users in the Galileo navigation message and updated at least once a day.[30] Ionospheric Disturbance Flag (IDF) alerts for five predefined regions (see Fig. 5.16) will also be broadcast to alert users when the ionospheric correction coming from the Galileo broadcast message might not meet the specified performance. These disturbance flags will be transmitted continuously and updated with the update rate of the navigation message (every 100 min).

The algorithm for the Galileo single-frequency receiver is based on the following steps (from [Arbesser-Rastburg, B., 2006]):

1. *Az is evaluated using a_0, a_1, a_2 (from navigation message) and MODIP calculated from the data inside the DIPLATS matrix*[31] *from the NeQuick model (which depends on estimated receiver position).*

[28] The F10.7 index is a measure of the solar activity, that is the flux level generated by the Sun at Earth's orbit at a wavelength of 10.7 cm (see Fig. 5.12, right). It has been found to correlate well with the sunspot number (Rz). The sunspot number is defined from counts of the number of individual sunspots as well as the number of sunspot groups. The F10.7 index can be measured relatively easily and quickly and has replaced the sunspot number as an index of solar activity for many purposes. F10.7 and the smoothed sunspot number R12 (12-month moving average) are related by $\mathrm{R12} = \sqrt{167\,273.0 + 1123.6\,(\mathrm{F10.7} - 63.7)} - 408.99$ (ITU-R recommendation).

[29] I is 0° at the magnetic equator and 90° at the magnetic poles.

[30] These parameters will be determined from measured TEC data obtained during the previous 24 h by the globally distributed Galileo Sensor Stations (GSS) network.

[31] Information on Earth's magnetic field will be stored in the receiver firmware as the DIPLATS matrix. It should be updated every five years, to take account of its natural variation. This update should be considered in the design of the Galileo receivers (MODIP table flash-able in firmware). The ITU-R maps for FoF2 and M3000(F2) files will also be stored in 12 ASCII files, one for each calendar month.

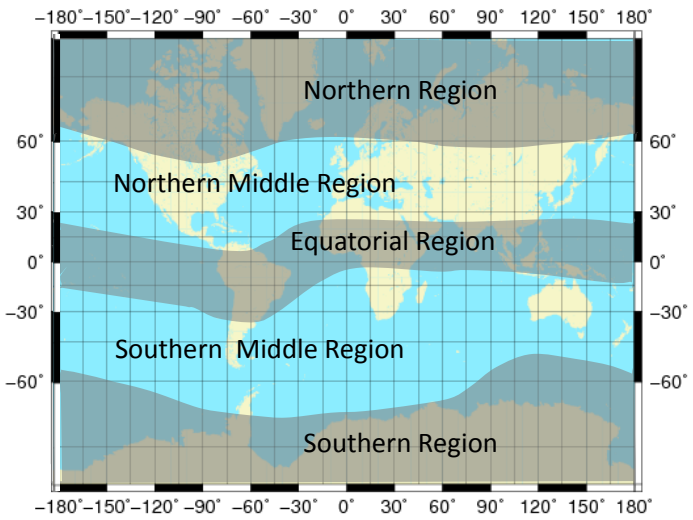


Figure 5.16: The five regions defined for the disturbance flags (in degrees MODIP) −90 to −60, −60 to −30, −30 to 30, 30 to 60, 60 to 90 (source [Arbesser-Rastburg, B., 2006]).

2. *Electron density is calculated for a point along the satellite to receiver path, using the NeQuick model with Az in place of F*10.7.

3. *Steps 1 and 2 are repeated for many discrete points along the satellite receiver path. The number and spacing of the points will depend on the height and they will be a trade-off between integration error and computational time and power.*

4. *All electron density values along the ray are integrated in order to obtain Slant TEC (or STEC).*

5. *STEC, in TECUs, is converted to metres of L1 slant delay for correcting pseudo-ranges, by*

$$I_f = \frac{40.3 \cdot 10^{16}}{f^2}\text{TEC} \quad (\text{where } f \text{ is in Hz}) \tag{5.58}$$

Note that, as with the Klobuchar model, the ionospheric corrections computed by the NeQuick model can be used for any GNSS signal (GPS, Glonass, Galileo, etc.) simply by setting the corresponding frequency in equation (5.58).

5.4.2 Tropospheric Delay

The troposphere is the atmospheric layer between Earth's surface and an altitude of about 60 km.

The effect of the troposphere on the GNSS signals appears as an extra delay in the measurement of the signal travelling from the satellite to the receiver. This delay depends on the temperature, pressure and humidity as well as the transmitter and receiver antenna locations and, according to equation (5.30), can be written as

$$Tr = \int (n-1)\, dl = 10^{-6} \int N\, dl \tag{5.59}$$

where n is the refractive index of air and $N = 10^6\,(n-1)$ is the refractivity. This refractivity can be divided into hydrostatic (i.e. *dry gases*, mainly N_2 and O_2) and wet (i.e. *water vapour*) components $N = N_{hydr} + N_{wet}$.

Each of these components has a different effect on GNSS signals. The main feature of the troposphere is that it is a non-dispersive medium with respect to electromagnetic waves up to 15 GHz; that is, the tropospheric effects are not frequency dependent for GNSS signals. Thus, the carrier phase and code measurements are affected by the same delay.

An immediate consequence of this non-frequency-dependent delay is that the tropospheric refraction cannot be removed by combinations of dual-frequency measurements (as is done with the ionosphere). Hence the only way to mitigate the tropospheric effect is to use models and/or to estimate it from observational data. Nevertheless, fortunately, most of the tropospheric refraction (about 90%) comes from the predictable hydrostatic component [Leick, 1994].

A brief description of the atmospheric dry and wet component effects on GNSS signals is as follows:

- *Hydrostatic component* delay: This is caused by the dry gases present in the troposphere (78% N_2, 21% O_2, 0.9% Ar, etc.). Its effect varies with local temperature and atmospheric pressure in a quite predictable manner, although its variation is less than 1% over a few hours. The error caused by this component is about 2.3 m in the zenith direction and 10 m for lower elevations (10° approximately).

- *Wet component* delay: This is caused by the water vapour and condensed water in the form of clouds and, therefore, it depends on the weather conditions. The excess delay is small in this case, only some tens of centimetres, but this component varies faster than the hydrostatic component and in a quite random way, thus being very difficult to model.

The dry atmosphere can be modelled from the surface pressure and temperature using the laws of ideal gases. The wet component is more unpredictable and difficult to model, so for high-accuracy positioning, this delay is estimated together with the coordinates.

The tropospheric delay depends on the signal path through the neutral atmosphere, and therefore can be modelled as a function of the satellite elevation angle. Due to the differences between the atmospheric profiles of the dry gases and water vapour it is better to use different mappings for the dry and wet components. Nevertheless, simple models such as [RTCA-MOPS, 2006] use a common mapping for both components.

Several nominal tropospheric models are available in the literature, but differ in the assumptions made on the vertical profiles and mappings. Basically, they can be classified in two main groups: geodetic-oriented or navigation-oriented. The first group (Saastamoinen, Hopfield and other models [Xu, 2007]) are more accurate but generally more complex, and need surface meteorological data, since their accuracy is affected by the quality of these data. The second group are less accurate, but meteorological data are not needed.

5.4.2.1 Example of Tropospheric Model for SPP

The tropospheric model presented here was developed by [Collins, 1999] and is the model adopted by the Satellite-Based Augmentation System (SBAS) (WAAS, EGNOS, etc.) [RTCA-MOPS, 2006].

In this case, there is a common mapping function for the wet and dry troposphere:

$$Tr(E) = (Tr_{z,d} + Tr_{z,w})\, M(E) \tag{5.60}$$

where $Tr_{z,d}$ and $Tr_{z,w}$ (in m) are calculated from the receiver's height and estimates of five meteorological parameters: pressure [P (mbar)], temperature [T (K)], water vapour pressure [e (mbar)], temperature 'lapse' rate [β (K/m)] and water vapour 'lapse rate' [λ (dimensionless)]. The obliquity factor $M(E)$ is the [Black and Eisner, 1984] mapping

$$M(E) = \frac{1.001}{\sqrt{0.002\,001 + \sin^2(E)}} \tag{5.61}$$

which is valid for satellite elevation angles E over 5°.

For a given receiver latitude φ and day of year D (i.e. from 1 January), the value of each meteorological parameter is computed from the averaged meteorological parameters of Table 5.1.

Indeed, each parameter value (P, T, e, β, λ) is computed as

$$\xi(\varphi, D) = \xi_0(\varphi) - \Delta\xi(\varphi)\, \cos\left[\frac{2\pi(D - D_{min})}{365.25}\right] \tag{5.62}$$

where $D_{min} = 28$ for northern latitudes and 211 for southern latitudes; $\xi_0(\varphi)$ and $\Delta\xi(\varphi)$ are the average and seasonal variation values at the receiver's latitude φ linearly interpolated from Table 5.1.

The zero-altitude vertical delay terms $Tr_{z,d}$ and $Tr_{z,w}$ are given by

$$Tr_{z_0,d} = \frac{10^{-6}\, k_1\, R_d\, P}{g_m}, \qquad Tr_{z_0,w} = \frac{10^{-6}\, k_2\, R_d}{(\lambda + 1)\, g_m - \beta\, R_d}\frac{e}{T} \tag{5.63}$$

The vertical delay terms $Tr_{z,d}$ and $Tr_{z,w}$ at the receiver's height H are calculated as

$$Tr_{z,d} = \left[1 - \frac{\beta\, H}{T}\right]^{\frac{g}{R_d \beta}} Tr_{z_0,d}\,, \quad Tr_{z,w} = \left[1 - \frac{\beta\, H}{T}\right]^{\frac{(\lambda+1)g}{R_d \beta} - 1} Tr_{z_0,w} \tag{5.64}$$

where H is the height above mean sea level, in m, $k_1 = 77.604\,\mathrm{K/mbar}$, $k_2 = 382\,000\,\mathrm{K^2/mbar}$, $R_d = 287.054\,\mathrm{J/(kg\ K)}$, $g_m = 9.784\,\mathrm{m/s^2}$ and $g = 9.80665\,\mathrm{m/s^2}$.

The computation of vertical delays in equation (5.64) and mapping function (5.60) are implemented in `gLAB` as the nominal model for SPP. They are also implemented in program `sub_trpUNB3.f` of Volume II.

Figure 5.17 illustrates with an example the values of the tropospheric delay and the effect of neglecting such delays in SPP for the horizontal and vertical error components (same data set as in Fig. 5.3).

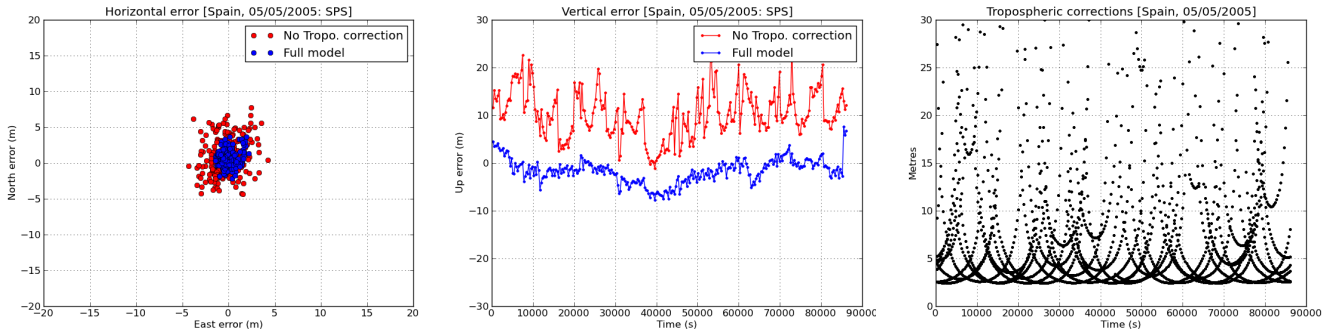


Figure 5.17: Tropospheric correction: range and position domain effect. Left and middle panels show the horizontal and vertical positioning error, respectively, using (blue) or not using (red) the tropospheric correction (equation (5.60)). The variation in range is shown in the right panel. See exercise 1f of session 5.2 in Volume II.

Table 5.1: Meteorological parameters for the tropospheric delay. Parameters for latitudes $15° < |\varphi| < 75°$ must be linearly interpolated between values for the two closest latitudes $[\varphi_i, \varphi_{i+1}]$. Parameters above $|\varphi| \leq 15°$ and $|\varphi| \geq 75°$ are taken directly from the table.

Latitude	Average				
(°)	P_0 (mbar)	T_0 (K)	e_0 (mbar)	β_0 (K/m)	λ_0
15 or less	1013.25	299.65	26.31	$6.30{\cdot}10^{-3}$	2.77
30	1017.25	294.15	21.79	$6.05{\cdot}10^{-3}$	3.15
45	1015.75	283.15	11.66	$5.58{\cdot}10^{-3}$	2.57
60	1011.75	272.15	6.78	$5.39{\cdot}10^{-3}$	1.81
75 or more	1013.00	263.65	4.11	$4.53{\cdot}10^{-3}$	1.55
Latitude	Seasonal variation				
(°)	ΔP (mbar)	ΔT (K)	Δe (mbar)	$\Delta\beta$ (K/m)	$\Delta\lambda$
15 or less	0.00	0.00	0.00	$0.00{\cdot}10^{-3}$	0.00
30	−3.75	7.00	8.85	$0.25{\cdot}10^{-3}$	0.33
45	−2.25	11.00	7.24	$0.32{\cdot}10^{-3}$	0.46
60	−1.75	15.00	5.36	$0.81{\cdot}10^{-3}$	0.74
75 or more	−0.50	14.50	3.39	$0.62{\cdot}10^{-3}$	0.30

5.4.2.2 Example of Tropospheric Model for PPP

The example presented here is one the tropospheric models implemented in the GIPSY-OASIS II software [Webb and Zumberge, 1993], which does not require any surface meteorological data. This model uses the mapping of Niell (see next section), which considers different obliquity factors for the wet and dry components

$$Tr(E) = Tr_{z,dry} \cdot M_{dry}(E) + Tr_{z,wet} \cdot M_{wet}(E) \tag{5.65}$$

In this implementation, the wet tropospheric delay is estimated by the navigation filter together with the receiver's position. This approach allows

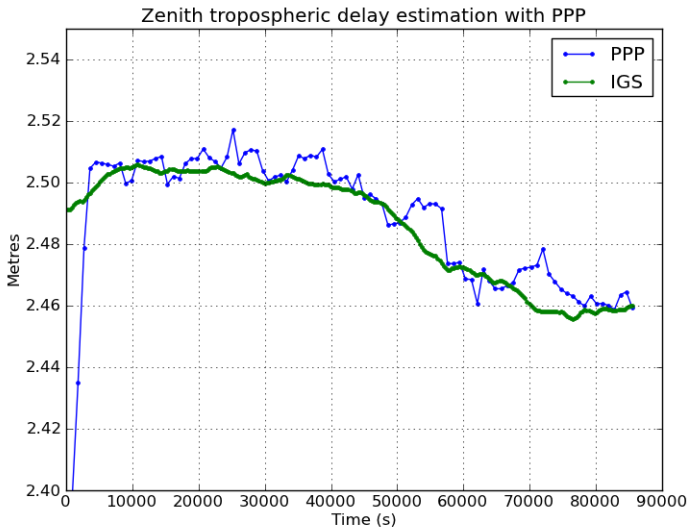


Figure 5.18: Comparison of ZTD (PPP) estimate with IGS determination. See exercise 5b of session 5.1 in Volume II.

a huge simplification of the model for the vertical delays, whose nominal values are

$$\begin{aligned} Tr_{z,dry} &= \alpha \, \mathrm{e}^{-\beta H} \\ Tr_{z,wet} &= Tr_{z_0,wet} + \Delta Tr_{z,wet} \end{aligned} \tag{5.66}$$

where $\alpha = 2.3\,\mathrm{m}$, $\beta = 0.116 \cdot 10^{-3}$ and H is the height above sea level, in m. $Tr_{z_0,wet} = 0.1\,\mathrm{m}$ and $\Delta Tr_{z,wet}$ is estimated[32] as a random walk process (typically 1 cm^2/h of process noise) in the navigation – Kalman – filter together with the coordinates and other parameters, see section 6.2.2. The accuracy of the $\Delta Tr_{z,wet}$ estimates in static positioning is at the centimetre level.

Figure 5.18 shows a comparison of the Zenith Tropospheric Delay (ZTD) estimate (static PPP) with the IGS determination (exercise 6 of laboratory session 5.2 in Volume II).

5.4.2.2.1 The Mapping of Niell

A mapping function that does not require surface meteorology measurements, but provides comparable accuracy and precision to others that require such data, is presented in [Niell, 1996]. This mapping uses only the receiver's geographic location and measurement time as inputs.

The mapping of Niell involves two different obliquity factors, for the dry and wet components, which are computed from equations (5.67) to (5.70), where (E) is the elevation of the ray and (H) is the receiver's height above sea level, in km:

Hydrostatic mapping function:

$$\begin{aligned} &M_{dry}(E,H) = m(E,a_d,b_d,c_d) + \Delta m(E,H) \\ &\text{with} \\ &\Delta m(E,H) = \left[\frac{1}{\sin E} - m(E,a_{ht},b_{ht},c_{ht})\right] H \end{aligned} \tag{5.67}$$

[32] Actually, by using the simplest model, the mismodelling for the dry component is also captured by this parameter estimate.

Wet mapping function:

$$M_{wet}(E) = m(E, a_w, b_w, c_w) \tag{5.68}$$

where $m(E, a, b, c)$ is the [Marini, 1972] mapping normalised to unity at zenith:

$$m(E, a, b, c) = \frac{1 + \dfrac{a}{1 + \dfrac{b}{1+c}}}{\sin E + \dfrac{a}{\sin E + \dfrac{b}{\sin E + c}}} \tag{5.69}$$

The hydrostatic parameters a_d, b_d, c_d are time (t) and latitude (φ) dependent parameters given by

$$\xi(\varphi, t) = \xi_{avg}(\varphi) - \xi_{amp}(\varphi) \cos\left(2\pi \frac{t - T_0}{365.25}\right) \tag{5.70}$$

where t is the time from January 0.0, in days, and T_0 is taken as DoY 28 (i.e. $T_0 = 28$). The parameters $\xi_{avg}(\varphi_i)$ and $\xi_{amp}(\varphi_i)$ are linearly interpolated from Table 5.2 between the nearest $\xi(\varphi_i)$. The a_{ht}, b_{ht}, c_{ht} parameters are taken directly from the same table.

The wet parameters a_w, b_w, c_w are latitude dependent and are linearly interpolated from Table 5.3 between the nearest $\xi(\varphi_i)$.

Table 5.2: Coefficients of the hydrostatic mapping function.

Coefficient	Latitude (φ)				
ξ	15°	30°	45°	60°	75°
	Average				
a	1.2769934e-3	1.2683230e-3	1.2465397e-3	1.2196049e-3	1.2045996e-3
b	2.9153695e-3	2.9152299e-3	2.9288445e-3	2.9022565e-3	2.9024912e-3
c	62.610505e-3	62.837393e-3	63.721774e-3	63.824265e-3	64.258455e-3
	Amplitude				
a	0.0	1.2709626e-5	2.6523662e-5	3.4000452e-5	4.1202191e-5
b	0.0	2.1414979e-5	3.0160779e-5	7.2562722e-5	11.723375e-5
c	0.0	9.0128400e-5	4.3497037e-5	84.795348e-5	170.37206e-5
	Height correction				
a_{ht}	2.53e-5				
b_{ht}	5.49e-3				
c_{ht}	1.14e-3				

Table 5.3: Coefficients of the wet mapping function.

Coefficient	Latitude (φ)				
ξ	15°	30°	45°	60°	75°
a	5.8021897e-4	5.6794847e-4	5.8118019e-4	5.9727542e-4	6.1641693e-4
b	1.4275268e-3	1.5138625e-3	1.4572752e-3	1.5007428e-3	1.7599082e-3
c	4.3472961e-2	4.6729510e-2	4.3908931e-2	4.4626982e-2	5.4736038e-2

This tropospheric model is implemented in the `gLAB` software tool for PPP. It is also provided in program `mapp_niell.f` given in Volume II.

5.5 Carrier Phase Wind-up Effect

Wind-up affects only the carrier phase measurements, not the code pseudo-ranges (see the term $\lambda_i\,\omega$ in equation (5.1)). It is due to the electromagnetic nature of circularly polarised waves, intrinsic to the GNSS (GPS, Glonass, Galileo, Beidou) signals. This correction is required only for high-accuracy positioning (e.g. PPP).

The wind-up effect on phase measurements depends on the relative orientation of the satellite and receiver antennas, and the direction of the line of sight.

For a receiver with fixed coordinates, the wind-up is due to the satellite's orbital motion. As the satellite moves along its orbital path it must perform a rotation to keep its solar panels pointing towards the Sun in order to obtain the maximum energy while its antenna keeps pointing towards the centre of Earth. This rotation causes a phase variation that the receiver misunderstands as a range variation.

The wind-up correction ($\omega \equiv \Delta\Phi$) due to the satellite's motion can be computed from the expressions derived by [Wu et al., 1993] for a crossed dipole antenna:

$$\Delta\Phi = \delta\phi + 2N\pi \tag{5.71}$$

where $\delta\phi$ is the fractional part of a cycle given by

$$\delta\phi = \text{sign}(\zeta)\ \arccos\left(\frac{\mathbf{d}'\cdot\mathbf{d}}{\|\mathbf{d}'\|\,\|\mathbf{d}\|}\right),\ \text{with}\quad \zeta = \hat{\boldsymbol{\rho}}\,\left(\mathbf{d}'\times\mathbf{d}\right) \tag{5.72}$$

and N is an integer number given by

$$N = \text{nint}\left[\frac{\Delta\Phi_{prev} - \delta\phi}{2\pi}\right] \qquad (N \text{ can be initialised as zero}) \tag{5.73}$$

where $\Delta\Phi_{prev}$ is the previous value of the phase correction, nint stands for the *nearest integer* and $\mathbf{d}$, $\mathbf{d}'$ are two effective dipoles for the receiver and transmitter:

$$\begin{aligned}\mathbf{d} &= \hat{\mathbf{a}} - \hat{\boldsymbol{\rho}}(\hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{a}}) + \hat{\boldsymbol{\rho}}\times\hat{\mathbf{b}}\\ \mathbf{d}' &= \hat{\mathbf{a}}' - \hat{\boldsymbol{\rho}}(\hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{a}}') - \hat{\boldsymbol{\rho}}\times\hat{\mathbf{b}}'\end{aligned} \tag{5.74}$$

where the unit vectors $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$, $\hat{\mathbf{a}}'$, $\hat{\mathbf{b}}'$ and $\hat{\boldsymbol{\rho}}$ are defined in Fig. 5.19 and below.

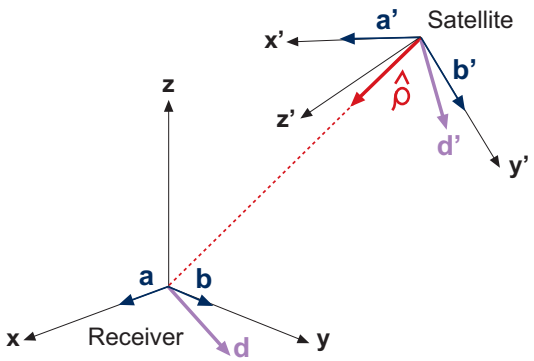


Figure 5.19: Layout of dipole orientation for computing the wind-up.

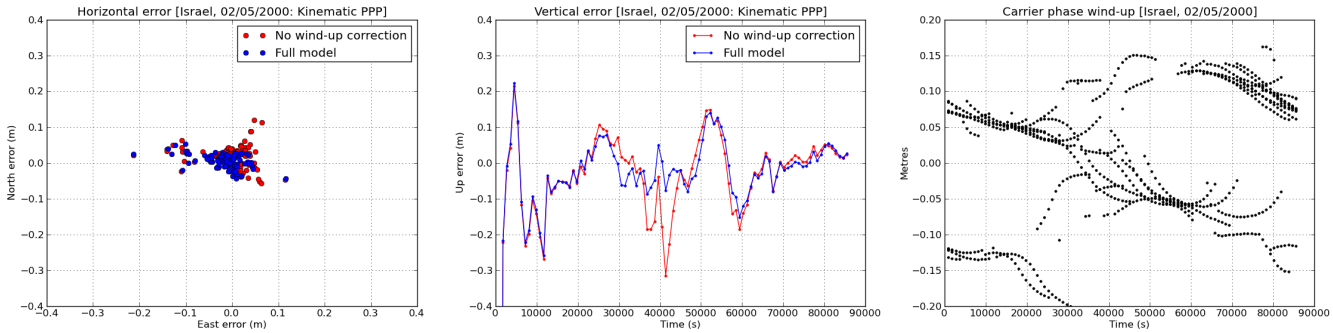


Figure 5.20: Carrier phase wind-up: range and position domain effect. Left and middle panels show the horizontal and vertical positioning error, respectively, using (blue) or not using (red) the carrier wind-up correction, equation (5.71). The wind-up effect on the range is shown in the right panel. See exercise 2b of session 5.2 in Volume II.

Continuity between consecutive phase observations must be ensured by adding full-cycle terms of $\pm 2\pi$ to the correction (5.71).

Receiver coordinate system, $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$: The local East North Up (ENU) coordinates can be used for the receiver, and the orthogonal unit vectors $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ can be defined as pointing to the east and north.

Satellite coordinate system, $\hat{\mathbf{a}}'$, $\hat{\mathbf{b}}'$: The satellite-fixed coordinate (x', y', z') frame of section 5.6.3 can be used for the satellite, and the orthogonal unit vectors $\hat{\mathbf{a}}'$ and $\hat{\mathbf{b}}'$ can be defined as pointing in the $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ directions, defined by equations (5.78) and (5.77), respectively.[33]

Line-of-sight vector $\hat{\boldsymbol{\rho}}$: This is a unit vector pointing from the satellite to the receiver.

Due to the carrier phase ambiguity, the value of N can be chosen arbitrarily at the beginning of a phase tracking session, but usually is taken as zero (i.e. it is assimilated into the unknown carrier phase ambiguity).

The wind-up correction is implemented in `gLAB`.

Figure 5.20 illustrates the effect of the carrier wind-up correction for the same example as in Fig. 5.23 (kinematic PPP). The solution computed using the wind-up correction is shown in blue and the solution without this correction in red. The wind-up effect on the range is shown in the plot on the right.

[33] The planetary ephemerides are usually given in the CRS reference frame (i.e. relative to the Aries direction. Nevertheless, given the low accuracy required for these computations (only to compute the unit vectors), a simple rotation $R_3[-\theta]$ of sidereal time θ is enough to transform the CRS coordinates to the TRS (see section 3.1.2).

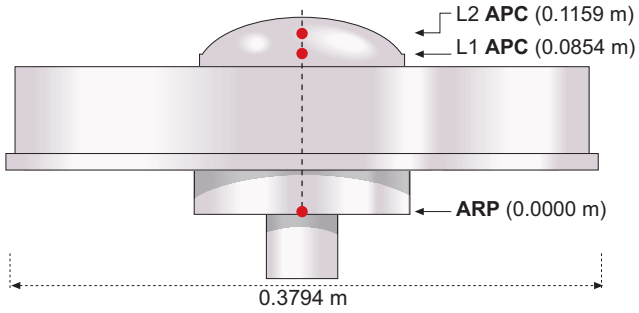


Figure 5.21: Layout of a Dorne Margolin T antenna. The Antenna Phase Centre (APC) offsets for L1 and L2 signals and the Antenna Reference Point (ARP) are indicated in the figure.

5.6 Antenna Phase Centre Correction

As in the previous section, the corrections outlined below are required only for PPP.

5.6.1 Antenna Phase Centre

The GNSS measurements are referred to the so-called antenna phase centre (satellite and receiver). The position of the antenna phase centre is not necessarily the geometric centre of the antenna. Indeed, it is not constant, but depends on the direction of the incoming radio signal.

The phase centre is defined as the *apparent source of radiation*. If the source were ideal it would have a spherical equiphase contour, but the real case is slightly different, because the equiphase contour is irregular and each segment has its own apparent radiation origin.

Moreover, the phase centre of an antenna not only is angle dependent (elevation and azimuth), but also depends on the signal frequency. A simple model is to assume that the phase centres differ only on the vertical axis of the antenna, see Fig. 5.21. Manufacturers include technical sheets indicating the phase centre offsets.

Antenna phase centre corrections for different models of receivers, *relative to the AOA Dorne Margolin T (AOAD/M_T) antenna*,[34] were compiled by IGS and are available from *ftp://igscb.jpl.nasa.gov/igscb/station/general/igs_01.pcv*. These files[35] include North East Up (NEU) offsets of the phase centre location and Phase Centre Variation (PCV) as a function of the elevation angle over 10°. They were determined using short-baseline field measurements taking the AOAD/M_T antenna as a reference [Gendt, G. and Schmid, R., 2005]. Nevertheless, as shown by [Mader, 1999], on long baselines the relative PCVs are not adequate, even for the same antenna model.

On 5 November 2006, IGS switched from relative to *absolute antenna phase centre corrections*, which took place together with the adoption of ITRF2005 [Gendt, G. and Schmid, R., 2005]. The absolute antenna phase centre offsets and PCVs are determined by a robotic system developed

[34] The IGS station network was initially dominated by AOA Dorne Margolin chokering antenna models.

[35] See *http://igscb.jpl.nasa.gov/igscb/station/general/igs_01.txt*.

at the University of Hanover and the company Geo++ (see *http://www.geopp.de*). They include azimuthal values and elevations down to 0°. IGS is compiling a consistent set of absolute antenna phase centre corrections for both receiver networks and satellites, which are provided in so-called Antenna Exchange Format (ANTEX) files.[36]

Since 5 November 2006 (GPS week 1400) the IGS products, that is Standard Product #3 (SP3) files[37] with the GPS and Glonass precise orbits and clocks, etc., and the Solution (Software/technique) INdependent EXchange format (SINEX) files with the computed precise coordinates of the reference station network, have been associated with the ANTEX files for absolute antenna phase centre corrections (offsets and PCVs). It must be pointed out that users should not mix absolute and relative PCVs. Moreover, absolute PCVs require corrections for both satellite and receiver antennas.

5.6.2 Receiver Antenna Phase Centre and Antenna Reference Point

The measurements are referred to the Antenna Phase Centre (APC) position. As this location is frequency dependent, a point tied to the base of the antenna is used as a more suitable reference. This point is called the Antenna Reference Point (ARP). As mentioned above, manufacturers provide technical information on the APC position relative to the ARP. On the other hand, as also already mentioned, relative and absolute APC corrections have been compiled by IGS and are provided in the PCV and ANTEX files respectively for several antenna models.

Finally, in geodetic positioning, the receiver coordinates are referred to a Monument Marker (MM) or an external Benchmark (BM). Figure 5.22 illustrates this.

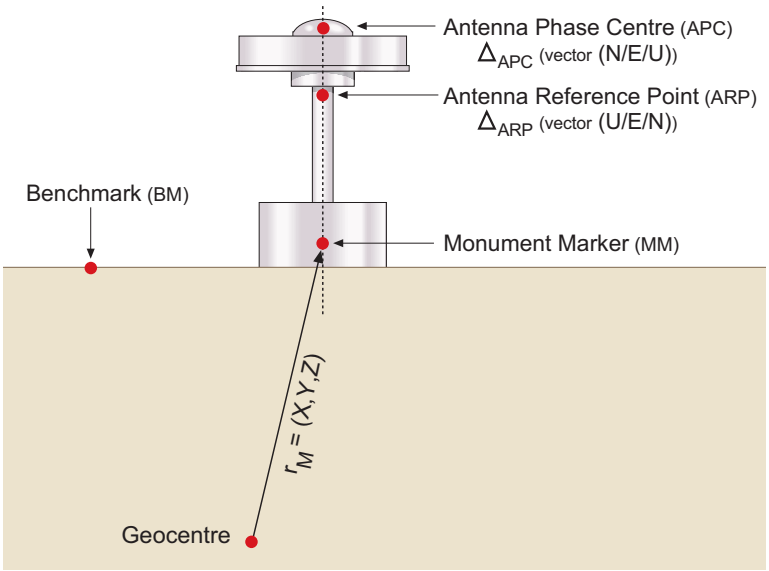


Figure 5.22: Layout of a permanent receiver site showing the MM, ARP and APC.

[36] See ANTEX format at *http://www.epncb.oma.be/ftp/station/general*.

[37] A reference to the satellite antenna phase centres used is indicated in the header of the SP3 files.

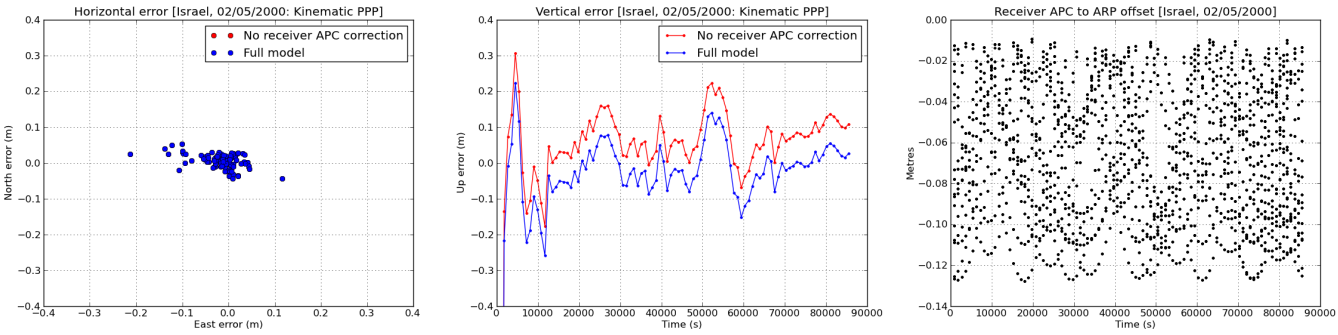


Figure 5.23: Receiver APC: range and position domain effect. Left and middle panels show the horizontal and vertical positioning error, respectively, using (blue) or not using (red) the receiver's APC correction. The variation in range is shown in the right panel. As the APC vector is along the vertical axis, its effect is a displacement in the vertical component. See exercise 2e of session 5.2 in Volume II.

In the IGS SINEX files, the computed coordinates for the MM are given in the block 'SOLUTION/ESTIMATE', in ECEF coordinates, for each processed station. The position of the ARP relative to the MM, or the site eccentricity vector, is given in the block 'SITE/ECCENTRICITY' in UNE (Up, North, East) coordinates. Finally, the APC offsets for the different frequencies and the antenna calibration model (e.g. ANTEX file) are given in the 'SITE/GPS_PHASE_CENTER' block of data.[38]

Let $\mathbf{r}_M$ be the position of the MM in an ECEF reference frame. Let $\boldsymbol{\Delta}_{\text{ARP}}$ be the offset vector defining the ARP position relative to the MM, and $\boldsymbol{\Delta}_{\text{APC}}$ the offset vector defining the APC position relative to the ARP. Thus the receiver's APC position $\mathbf{r}$, in an ECEF reference frame, is given by[39]

$$\mathbf{r} = \mathbf{r}_M + \boldsymbol{\Delta}_{\text{ARP}} + \boldsymbol{\Delta}_{\text{APC}} \tag{5.75}$$

The receiver's phase centre offset correction, using the ANTEX file format, and the ARP correction are implemented in `gLAB`.

Figure 5.23 shows an example of the effect of APC correction in the positioning domain (left and middle plots) and range domain (right plot). The solution computed using the APC correction is shown in blue and that without in red. The projection in range of the APC offset is shown in the plot on the right. The results correspond to an Ashtech-ZXII3 receiver with an ASH70093D_M antenna, located at coordinates $\varphi = 30°36^{\text{m}}$, $\lambda = 34°45^{\text{m}}$ (Israel), on 2 May 2000 and positioned in kinematic PPP mode.

[38] See the SINEX format at *http://www.iers.org/MainDisp.csl?pid=190-1100110*. SINEX files can be found at *ftp://cddis.gsfc.nasa.gov/pub/gps/products*.

[39] Note that, as commented earlier, $\boldsymbol{\Delta}_{\text{APC}}$ is a frequency-dependent correction. Equation (5.81) below gives the correction $\boldsymbol{\Delta}_{\text{APC}_{\text{LC}}}$ from the L1, L2 APC vectors.

5.6.3 Satellite Antenna Phase Centre

The broadcast ephemerides are referred to the satellite's APC (of the ionosphere-free combination signal, see equation (5.81)) in an ECEF reference frame (see [GPS/SPS-SS, 1995], [GLONASS ICD, 2008]) and, therefore, no additional correction is needed when using the navigation message. Nevertheless, the precise orbits and clocks are referred to the Satellite Mass Centre (MC), thus it is necessary to account for the phase centre offset vector when these products are used. This offset is given in a satellite-fixed coordinate frame, defined by the unit vectors $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ as follows:

- $\hat{\mathbf{k}}$ is the unit vector pointing from the MC to the centre of Earth. This vector can be computed as

$$\hat{\mathbf{k}} = -\frac{\mathbf{r}^{sat_{\mathrm{MC}}}}{\| \mathbf{r}^{sat_{\mathrm{MC}}} \|} \tag{5.76}$$

  where $\mathbf{r}^{sat_{\mathrm{MC}}}$ are the MC coordinates in the ECEF reference frame used for the SP3 file, for example IGS05.

- $\hat{\mathbf{j}}$ is the resulting unit vector of the cross-product of the $\hat{\mathbf{k}}$ vector with the unit vector from the satellite to the Sun. That is,

$$\hat{\mathbf{j}} = \hat{\mathbf{k}} \times \hat{\mathbf{e}}, \quad \text{with} \quad \hat{\mathbf{e}} = \frac{\mathbf{r}_{sun} - \mathbf{r}^{sat_{\mathrm{MC}}}}{\| \mathbf{r}_{sun} - \mathbf{r}^{sat_{\mathrm{MC}}} \|} \tag{5.77}$$

  where the $\mathbf{r}_{sun}$ vector can be computed from the planetary ephemerides.[40]

- $\hat{\mathbf{i}}$ completes the right-handed system

$$\hat{\mathbf{i}} = \hat{\mathbf{j}} \times \hat{\mathbf{k}} \tag{5.78}$$

Thus, if $\mathbf{\Delta}_{\mathrm{APC}}$ is the APC offset in the satellite-fixed $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ system, the satellite APC coordinates in the ECEF frame are

$$\mathbf{r}^{sat_{\mathrm{APC}}} = \mathbf{r}^{sat_{\mathrm{MC}}} + \mathbf{R} \cdot \mathbf{\Delta}_{\mathrm{APC}} \tag{5.79}$$

where

$$\mathbf{R} = \begin{bmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \end{bmatrix} \tag{5.80}$$

Note that the APCs ($\mathbf{\Delta}_{\mathrm{APC_{L1}}}$, $\mathbf{\Delta}_{\mathrm{APC_{L2}}}$) are frequency dependent as shown in Fig. 5.21. The APC offset in the ionosphere-free combination (see equation (4.4)) is given by

$$\mathbf{\Delta}_{\mathrm{APC_{LC}}} = \frac{f_1^2 \mathbf{\Delta}_{\mathrm{APC_{L1}}} - f_2^2 \mathbf{\Delta}_{\mathrm{APC_{L2}}}}{f_1^2 - f_2^2} \tag{5.81}$$

From 29 November 1998 (GPS week 986) to 4 November 2006 (GPS week 1400), the IGS products used the APC offsets given in Table 5.4 for the ionosphere-free combination ($\mathbf{\Delta}_{\mathrm{APC_{LC}}}$).

[40] Simplified expressions for solar coordinates can be found, for instance, in [Montenbruck and Eberhard, 2005] or in [GLONASS ICD, 1998]. The equations from Glonass-ICD are implemented in program `sub_Sun_pos_GLO.f` in Volume II.

Table 5.4: $\boldsymbol{\Delta}_{\mathrm{APC_{LC}}}$ of GPS satellite used by IGS until 4 November 2006.

Block	$\hat{\mathbf{i}}$	$\hat{\mathbf{j}}$	$\hat{\mathbf{k}}$
I	0.210 m	0.000 m	0.854 m
II/IIA	0.279 m	0.000 m	1.023 m
IIR	0.000 m	0.000 m	0.000 m

From 5 November 2006, IGS has used different APC vectors ($\boldsymbol{\Delta}_{\mathrm{APC}}$) even within the same GPS satellite block. These values are derived from an absolute calibration of the APC corrections (offset and PCVs, see section 5.6.1), for GPS and also for Glonass satellites.

Table 5.5 shows an example of $\boldsymbol{\Delta}_{\mathrm{APC_{LC}}}$ values for Glonass satellites provided in the IGS ANTEX files.

Table 5.5: Examples of $\boldsymbol{\Delta}_{\mathrm{APC_{LC}}}$ for Glonass satellites.

Orb. slot	#GC	$\hat{\mathbf{i}}$	$\hat{\mathbf{j}}$	$\hat{\mathbf{k}}$
01	796	0.000 m	0.000 m	1.9444 m
04	795	0.000 m	0.000 m	2.0061 m
05	711	0.000 m	0.000 m	1.9141 m
06	701	−0.545 m	0.000 m	2.1947 m
07	712	−0.545 m	0.000 m	2.3232 m
02,03,08–22	...	−0.545 m	0.000 m	2.3000 m
23	714	−0.545 m	0.000 m	2.2772 m
24	713	−0.545 m	0.000 m	2.3253 m

The APC corrections model (i.e. ANTEX file) associated with each precise satellite orbit and clock SP3 file is indicated in the file header.

The APCs associated with the GPS broadcast orbits and clocks can be found at *http://earth-info.nga.mil/GandG/sathtml/*, from the NGA. An example is given in Table 5.6 (see the full table in Appendix F).

The satellite phase centre offset correction, using the ANTEX file format, is implemented in the `gLAB` software tool for PPP.

Figure 5.24 illustrates the effect of the satellite APC correction for the same example as in Fig. 5.23 (kinematic PPP). The navigation solution computed with the APC correction ($\boldsymbol{\Delta}_{\mathrm{APC}}$) is shown in blue and the solution without in red. The projection in range of the satellite APC offset is shown in the second row on the left.

Table 5.6: $\boldsymbol{\Delta}_{\mathrm{APC_{LC}}}$ of GPS satellites used for broadcast orbits and clocks.

Block	PRN	$\hat{\mathbf{i}}$	$\hat{\mathbf{j}}$	$\hat{\mathbf{k}}$
II		0.2794	0.0000	0.9519
IIA		0.2794	0.0000	0.9519
IIR	02	−0.0099	0.0061	−0.0820
IIR	11	0.0019	0.0011	1.5141
IIR	19	−0.0079	0.0046	−0.0180
IIR	23	−0.0088	0.0035	0.0004
IIR-M	01	0.012 45	−0.000 38	−0.022 83
IIR-M	17	−0.009 96	0.005 99	−0.100 60
IIR-M	29	−0.010 12	0.005 91	−0.015 12
IIR-M	31	0.001 60	0.000 33	−0.057 50

5.7 Earth Deformation Effects Modelling

The receiver station coordinates are affected by tidal motions, which must be accounted for when high accuracy is required. It is important to point out that these effects do not affect GNSS signals, but if the tidal effects were not considered, the station coordinates would oscillate in relation to a mean value as a consequence of these effects.

The main reasons why Earth's crust varies and modifies the coordinates of the receiver's location are summarised in the next sections under the headings of solid tides, ocean loading and pole tide. The equations for computing the vector of tidal displacements for each one of these effects ($\Delta\mathbf{r}_{sol}$, $\Delta\mathbf{r}_{ocn}$ and $\Delta\mathbf{r}_{pol}$) are given as follows. After computing each tidal displacement, the receiver's location is given by

$$\mathbf{r}_M = \mathbf{r}_{M_0} + \Delta\mathbf{r}_{sol} + \Delta\mathbf{r}_{ocn} + \Delta\mathbf{r}_{pol} \tag{5.82}$$

where $\mathbf{r}_{M_0}$ are the MM coordinates free of tidal displacements and $\mathbf{r}_M$ is given by equation (5.75).

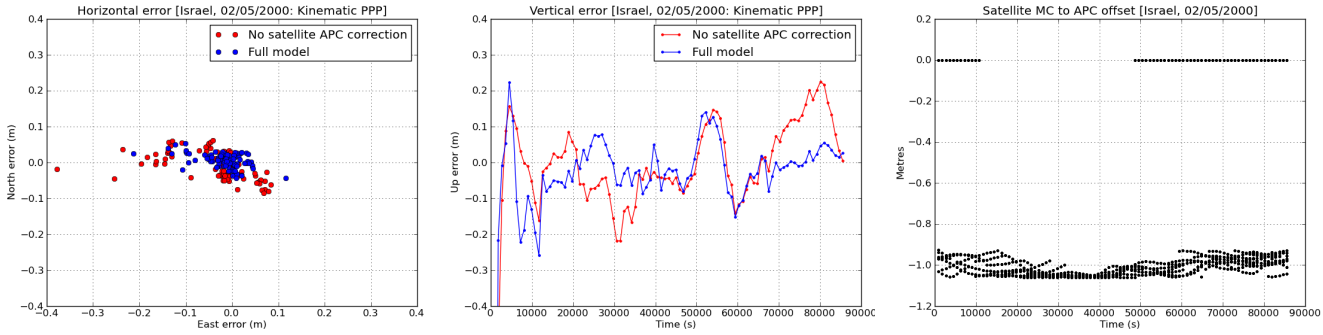


Figure 5.24: Satellite APC: range and position domain effect. Left and middle panels show the horizontal and vertical positioning error, respectively, using (blue) or not using (red) the satellite APC correction. The variation in range is shown in the right panel. The $\boldsymbol{\Delta}_{\mathrm{APC}}$ values of Table 5.4 have been used (data set collected in year 2002). See exercise 2d of session 5.2 in Volume II.

5.7.1 Solid Tides

These concern the movement of Earth's crust (and thus the variation in the receiver's location coordinates) due to gravitational attractive forces produced by external bodies, mainly the Sun and Moon. Solid tides produce vertical and horizontal displacements that can be expressed by the spherical harmonics expansion (m, n), characterised by the Love and Shida numbers h_{mn} and l_{mn}.

A simplified model for the tidal displacement, to a few millimetres of accuracy, is given by the following expression (from the IERS Convention's degree 2 tides displacement model – in-phase corrections; see [McCarthy and Petit, 2004], page 79):

$$\Delta\mathbf{r}_{sol} = \sum_{j=2}^{3} \frac{G\,M_j\,R_E^4}{G\,M_E\,R_j^3} \left\{ h_2\,\hat{\mathbf{r}} \left(\frac{3}{2}(\hat{\mathbf{R}}_j \cdot \hat{\mathbf{r}})^2 - \frac{1}{2} \right) + 3\,l_2\,(\hat{\mathbf{R}}_j \cdot \hat{\mathbf{r}}) \left[\hat{\mathbf{R}}_j - (\hat{\mathbf{R}}_j \cdot \hat{\mathbf{r}})\,\hat{\mathbf{r}} \right] \right\} \tag{5.83}$$

where:

$\Delta\mathbf{r}_{sol}$	is a site displacement vector in ECEF Cartesian coordinates
$G\,M_E$	is the gravitational parameter of Earth
$G\,M_j$	is the gravitational parameter of the Moon $(j = 2)$ and Sun $(j = 3)$ $(M_{sun}/M_E = 332\,946.0,\ M_{moon}/M_E = 0.012\,300\,02)$
$\hat{\mathbf{R}}_j$, R_j	are the unit vector from the geocentre to the Moon or Sun and the magnitude of that vector
R_E	is the equatorial radius of Earth $(R_E = 6\,378\,136.6\,\mathrm{m})$
$\hat{\mathbf{r}}$, r	are the unit vector from the geocentre to the station and the magnitude of that vector
h_2	is the nominal degree 2 Love number $(h_2 = 0.6078)$.
l_2	is the nominal degree 2 Shida number $(l_2 = 0.0847)$.

Note that the radial (not vertical) component $(\hat{\mathbf{r}})$ is proportional to the Love number h_2, while the terms in l_2 correspond to components orthogonal to the radial direction (not the horizontal plane).

A small correction for geocentric latitude ϕ dependence can be considered in the h_2 and l_2 values of equation (5.83), according to the expressions

$$\begin{aligned} h_2 &= 0.6078 - 0.0006\,\left[(3\sin^2\phi - 1)/2\right] \\ l_2 &= 0.0847 + 0.0002\,\left[(3\sin^2\phi - 1)/2\right] \end{aligned} \tag{5.84}$$

An additional refinement is to take into account the extra contribution due to the degree 3 tides ([McCarthy and Petit, 2004], page 80):

$$\begin{aligned} \Delta\mathbf{r}_{sol} = \sum_{j=2}^{3} \frac{G\,M_j\,R_E^5}{G\,M_E\,R_j^4} \Big\{ & h_3\,\hat{\mathbf{r}} \left(\tfrac{5}{2}(\hat{\mathbf{R}}_j \cdot \hat{\mathbf{r}})^3 - \tfrac{3}{2}(\hat{\mathbf{R}}_j \cdot \hat{\mathbf{r}}) \right) \\ & + l_3 \left(\frac{15}{2}(\hat{\mathbf{R}}_j \cdot \hat{\mathbf{r}})^2 - \frac{3}{2} \right) \left[\hat{\mathbf{R}}_j - (\hat{\mathbf{R}}_j \cdot \hat{\mathbf{r}})\,\hat{\mathbf{r}} \right] \Big\} \end{aligned} \tag{5.85}$$

where $h_3 = 0.292$ and $l_3 = 0.015$. Only the Moon's contribution $(j = 2)$ needs to be computed, because the contribution of the Sun $(j = 3)$ is negligible. However, the Moon's contribution to this degree 3 tide in the radial

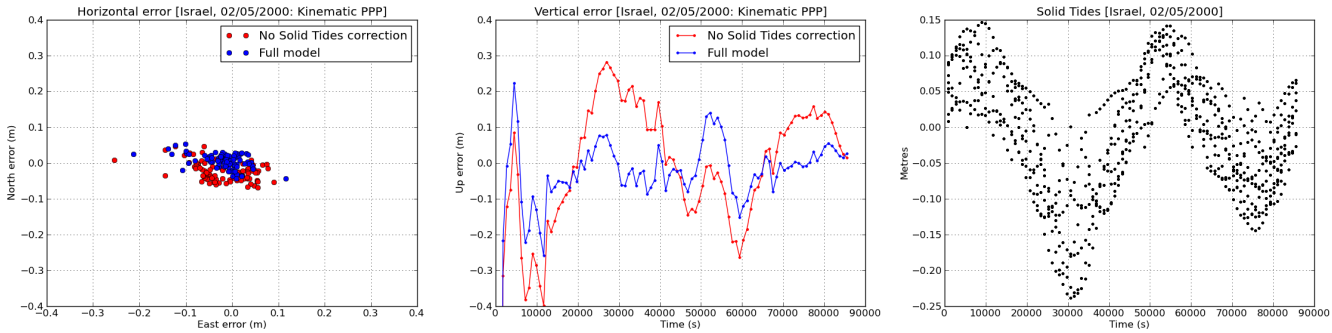


Figure 5.25: Solid tides: range and position domain effect. Left and middle panels show the horizontal and vertical positioning error, respectively, using (blue) or not using (red) the solid tides correction (from equations (5.83) to (5.85)). The effect on range of the solid tides displacement is shown in the right panel. See exercise 2c of session 5.2 in Volume II.

displacement does not exceed 1.7 mm in radial and 0.02 mm in transversal components.

Finally, it must be taken into account that the previous equations provide the correction to obtain the coordinates relative to the 'conventional tide-free' position. To obtain the position relative to the 'mean tide' the following vector must be added:[41]

$$\begin{array}{lll} [-0.1206 + 0.0001\, P_2(\sin\phi)]\, P_2(\sin\phi) & \text{(m)} & \text{radial direction} \\ [-0.0252 + 0.0001\, P_2(\sin\phi)]\, \sin 2\phi & \text{(m)} & \text{north direction} \end{array} \tag{5.86}$$

where $P_2(\sin\phi) = (3\,\sin^2\phi - 1)/2$.

Note that the radial component of this part can amount to about −12 cm at the poles and about +6 cm at the equator.[42]

The previous equations (5.83), (5.84) and (5.85) – but not (5.86) – implementing the in-phase degree 2 and 3 tides are coded in the `gLAB` software tool.

The FORTRAN code `dehanttideinel.f` implementing the full IERS Convention's Solid-Tides Model is available at *http://www.usno.navy.mil/USNO/earth-orientation/eo-info/general/conv1996/dehanttideinel.f*.[43]

Figure 5.25 illustrates the effect of the solid tides correction for the same example as in Fig. 5.23 (kinematic PPP). The navigation solution computed with the solid tides correction is shown in blue and the solution without in red. The effect of the solid tides displacement on range is shown in the second row on the left.

[41] According to the IERS conventions (see [McCarthy and Petit, 2004], pages 9–10 and 83), the (degree 2 zonal) tidal potential contains a time-independent (i.e. permanent) part, which is included in the geoid definition. Thus, the 'mean tide' position is obtained after removing this part from the tidal displacement. This is done by adding the vector expression (5.86) to the tidal displacement computed from equation (5.83).

[42] Resolution 16 of the 18th IAG General Assembly (1983) recommended the use of mean tide values for quantities associated with station displacements. "This recommendation, however, has not been implemented in the algorithms used for tide modelling by the geodesy community in the analysis of space geodetic data in general", see [McCarthy and Petit, 2004] page 10.

[43] The permanent tidal deformation restitution (i.e. equation (5.86) to obtain the 'mean tide position' from the 'conventional tide-free' position) is deactivated in this software code (see details in file `dehanttideinel.f` header).

5.7.2 Ocean Loading

This is a secondary tidal effect, due to the elastic response of Earth's crust to ocean tides, producing deformation of the sea floor and a surface displacement of the adjacent land.

Ocean loading is more localised than solid tides and for convention it does not have a permanent part. For kinematic PPP at an accuracy of a few centimetres, or a few millimetres for static PPP over 24 h and/or far from the oceans, it can be neglected [Kouba and Héroux, 2000].

A model for ocean loading is described in the IERS Convention's document [McCarthy and Petit, 2004], page 73, a simplified version of which can be summarised as[44]

$$\Delta\mathbf{r}_{ocn} = \sum_j f_j A_{cj} \cos\left(\omega_j t + \chi_j + u_j - \Phi_{cj}\right) \tag{5.87}$$

where:

$\Delta\mathbf{r}_{ocn}$	is the site displacement vector in (radial, west, south) coordinates
j	represents 11 tidal waves: M_2, S_2, N_2, K_2, K_1, O_1, P_1, Q_1, M_f, M_m, S_{sa}
f_j, u_j	depend on the longitude of the lunar node
ω_j	is the tidal angular velocity at time $t = 0$h
χ_j	is an astronomical argument at time $t = 0$h
A_{cj}	is a station-specific amplitude
Φ_{cj}	is a station-specific phase.

The FORTRAN code `HARDISP.f` implementing the full IERS Convention's Solid-Tides Model is available at *ftp://tai.bipm.org/iers/convupdt/chapter7/hardisp*. This routine computes time series of tidal displacements from an input file containing the ocean loading coefficients for a given station. These coefficients can be obtained from the ocean loading service by request from the website *http://holt.oso.chalmers.se/loading/index.html*.

This secondary tidal effect is not implemented in the `gLAB` tool.

5.7.3 Pole Tide

As mentioned in section 3.1.2, Earth's instantaneous axis of rotation shifts inside a square of about 20 m side in relation to a point with fixed coordinates on Earth (i.e. the Chandler wobble with a period of 14 months). This entails a varying elastic response in Earth's crust. This has an effect of less than 2.5 cm in the vertical and 0.7 cm in the horizontal direction, but must be taken into account if the observations are carried out over periods longer than two months.

From the IERS Conventions [McCarthy and Petit, 2004], page 84, the following expression[45] can be derived for the displacement at a point of

[44] Additional information describing the model of GIPSY-OASIS II can be found in [Webb and Zumberge, 1993].

[45] Note the use of latitude ϕ in equations (5.88) and (5.89), instead of the co-latitude θ used in the IERS equations.

geocentric latitude ϕ and longitude λ:

$$\begin{aligned}
\delta_{\hat{r}} &= -\frac{\omega_E^2\, R_E}{2\, g}\, h\, \sin 2\phi\, (m_1 \cos\lambda + m_2 \sin\lambda)\, \hat{\mathbf{r}} \\
\delta_{\hat{\lambda}} &= -\frac{\omega_E^2\, R_E}{g}\, l\, \sin\phi\, (-m_1 \sin\lambda + m_2 \cos\lambda)\, \hat{\boldsymbol{\lambda}} \\
\delta_{\hat{\phi}} &= -\frac{\omega_E^2\, R_E}{g}\, l\, \cos 2\phi\, (\, m_1 \cos\lambda + m_2 \sin\lambda)\, \hat{\boldsymbol{\phi}}
\end{aligned} \tag{5.88}$$

where (m_1, m_2) are the displacements (in metres) from the 1903.0 Conventional International Origin (CIO), pole position, and $h = 0.6027$, $l = 0.0836$ are the Love numbers.

Using Earth's angular rotation $\omega_E = 7.29 \cdot 10^{-5}$ rad/s, Earth's equatorial radius $R_E = 6\,378 \cdot 10^3$ m and the gravitational acceleration $g = 9.8\,\mathrm{m/s^2}$, it follows that

$$\begin{aligned}
\delta_{\hat{r}} &= -32\, \sin 2\phi\, (x_1 \cos\lambda + x_2 \sin\lambda)\, \hat{\mathbf{r}} \quad \text{(mm)} \\
\delta_{\hat{\lambda}} &= -9\, \sin\phi\, (-x_1 \sin\lambda + x_2 \cos\lambda)\, \hat{\boldsymbol{\lambda}} \quad \text{(mm)} \\
\delta_{\hat{\phi}} &= -9\, \cos 2\phi\, (\, x_1 \cos\lambda + x_2 \sin\lambda)\, \hat{\boldsymbol{\phi}} \quad \text{(mm)}
\end{aligned} \tag{5.89}$$

where (x_1, x_2) are the displacements in seconds of arc ($m_i = x_i R_E \pi / 648000$). Pole displacements can be found at *ftp://hpiers.obspm.fr/iers/eop*.

The displacement $\boldsymbol{\delta}$ is given by the radial, longitudinal and latitudinal vectors $(\hat{\mathbf{r}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\phi}})$ (positive upwards, eastwards and northwards, respectively). Thus, the displacement vector $\Delta\mathbf{r}_{pol}$ in the (x, y, z) ECEF Cartesian coordinates is given by

$$\Delta\mathbf{r}_{pol} = \mathbf{R}_3(-\lambda) \cdot \mathbf{R}_2(\phi) \cdot \boldsymbol{\delta} \tag{5.90}$$

where $\mathbf{R}_3(-\lambda) \cdot \mathbf{R}_2(\phi)$ are the rotations in latitude (1) and longitude (2) indicated in Fig. 5.26 (see also section B.2.1). This short tidal effect is not implemented in the `gLAB` tool.

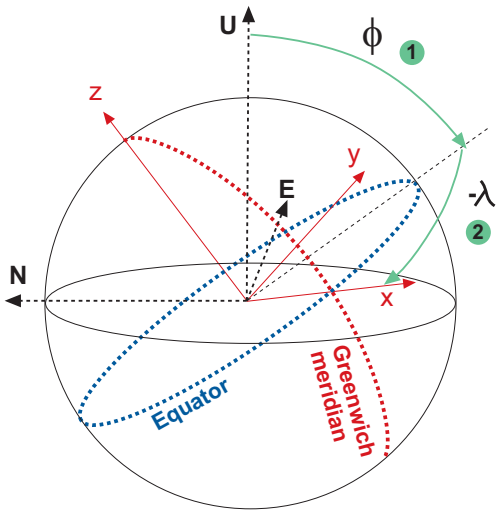


Figure 5.26: Transformation from UEN $(\hat{\mathbf{r}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\phi}})$ to ECEF (x, y, z) coordinates.

6. Solving Navigation Equations

This chapter is devoted to estimating the receiver's position. In particular, it focuses on: (1) single-frequency code-based point positioning, hereafter SPP; and (2) dual-frequency code and carrier-based PPP. These two topics are complemented with a brief section on carrier phase ambiguity fixing in differential and undifferenced mode.

As the measurements are noisy (receiver noise, multipath) and the applied models are not perfect, adjustment and filtering techniques are needed for parameter estimation. It is beyond the scope of this book to go into the fundamentals of parameter estimation, since the least squares solution or Kalman filtering is presented only from 'instrumental usage' point of view.

6.1 Basic Concepts: Code-Based Positioning

The aim is to determine the receiver coordinates $\mathbf{r} = (x, y, z)$ and clock offset δt from pseudorange measurements R of at least four satellites in view. As explained in the first chapter, the positioning principle is based on solving a geometric problem (see Chapter1) from the measured ranges to the satellites, with known coordinates. The satellite coordinates can be computed from the broadcast message, which also provides all the necessary information for modelling the measurements for the Standard Positioning Service (i.e. the SPP).

From the code pseudorange measurements R^j for $n \geq 4$ satellites

$$R^j = \rho^j + c(\delta t - \delta t^j) + Tr^j + \hat{\alpha}_1 I^j + \mathrm{TGD}^j + \mathcal{M}^j + \varepsilon^j, \quad j = 1, \ldots, n \quad (6.1)$$

(see equations (5.1) and (5.20)), the following measurement equation system can be written,[1], neglecting the multipath and receiver noise terms:

$$R^j - D^j \simeq \sqrt{(x^j - x)^2 + (y^j - y)^2 + (z^j - z)^2} + c\,\delta t, \quad j = 1, \ldots, n \quad (6.2)$$

where the left-hand side contains the measurements R^j and all modelled terms $D^j = -c\,\delta t^j + Tr^j + \hat{\alpha}_1 I^j + \mathrm{TGD}^j$, see equations (5.16), (5.60), (5.54) and (5.20). The right-hand side contains the four unknown parameters: the receiver coordinates (x, y, z) and the receiver clock offset δt.

Equations (6.2) define a nonlinear system, whose usual resolution technique consists of linearising the geometric range ρ in the neighbourhood of a point (x_0, y_0, z_0) corresponding to the approximate position of a receiver (see Fig. 6.1).

[1] The unknown receiver DCB $K_{21,rcv}$ is included in the receiver clock term δt.

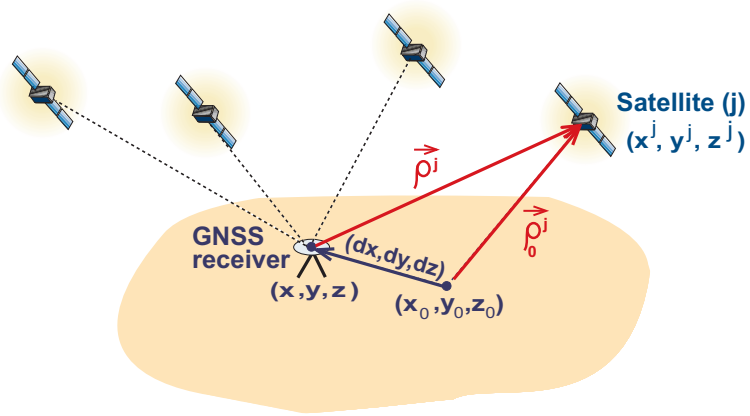


Figure 6.1: Geometric concept of GNSS positioning: Equations are linearised about the approximate receiver coordinates (x_0, y_0, z_0). The correction (dx, dy, dz) is estimated after solving the navigation equations (Eq. 6.6).

Then, linearising the satellite–receiver geometric range

$$\rho^j(x,y,z) = \sqrt{(x^j - x)^2 + (y^j - y)^2 + (z^j - z)^2} \tag{6.3}$$

gives, for the approximate solution $\mathbf{r}_0 = (x_0, y_0, z_0)$,

$$\begin{aligned} \rho^j &= \rho_0^j + \frac{x_0 - x^j}{\rho_0^j}dx + \frac{y_0 - y^j}{\rho_0^j}dy + \frac{z_0 - z^j}{\rho_0^j}dz \\ \text{with}\quad & dx = x - x_0, \quad dy = y - y_0, \quad dz = z - z_0 \end{aligned} \tag{6.4}$$

Substituting (6.4) in (6.2), we can rewrite the measurement equations as a linear system (where R^j can be either smoothed or unsmoothed code)

$$R^j - \rho_0^j - D^j = \frac{x_0 - x^j}{\rho_0^j}dx + \frac{y_0 - y^j}{\rho_0^j}dy + \frac{z_0 - z^j}{\rho_0^j}dz + c\,\delta t, \quad j = 1, \ldots, n \tag{6.5}$$

The previous system for the navigation equations[2] is written in matrix notation as

$$\begin{bmatrix} R^1 - \rho_0^1 - D^1 \\ \vdots \\ R^n - \rho_0^n - D^n \end{bmatrix} = \begin{pmatrix} \frac{x_0 - x^1}{\rho_0^1} & \frac{y_0 - y^1}{\rho_0^1} & \frac{z_0 - z^1}{\rho_0^1} & 1 \\ \vdots & \vdots & \vdots & \\ \frac{x_0 - x^n}{\rho_0^n} & \frac{y_0 - y^n}{\rho_0^n} & \frac{z_0 - z^n}{\rho_0^n} & 1 \end{pmatrix} \begin{bmatrix} dx \\ dy \\ dz \\ c\,\delta t \end{bmatrix} \tag{6.6}$$

In general, an over-determined system is obtained (for $n > 4$), which can be solved using the least squares adjustment. See exercise 7 of session 5.2 in Volume II.

After solving the equation system (6.6), the estimate of the receiver coordinates is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} \tag{6.7}$$

[2]Strictly speaking, this system corresponds to the case where the satellite coordinates at emission time have been calculated using the pseudorange algorithm described on page 98. When the purely geometric algorithm on page 99 is used, the elements of the associated matrix (design matrix or Jacobian) vary slightly (see details in Appendix E, page 200).

Equations (6.2) can be linearised again about these new estimates (6.7) of the receiver's position, and the solution can be iterated until the change between two consecutive iterations is below a given threshold. Typically, the iterations converge quickly, in a few iterations, even if starting with $(x_0, y_0, z_0) = (0, 0, 0)$, that is Earth's centre.

Equations (6.6) will be called the navigation equations system and can be written in compact form as

$$\mathbf{y} = \mathbf{G}\,\mathbf{x} \tag{6.8}$$

where the vectors and matrix involved can be defined as follows:

Prefit residuals: $\mathbf{y}$ is an $(n \times 1)$ vector containing the residuals between the measured and predicted pseudoranges, 'before fitting' the parameters $(dx, dy, dz, \delta t)$ to the linear model.

Geometry matrix: $\mathbf{G}$ is an $(n \times 4)$ matrix containing the receiver–satellite geometry.[3] Notice that the first three elements of each row $(j = 1, \ldots, n)$ are the components of the unitary line-of-sight vector

$$\widehat{\boldsymbol{\rho}}_0^j = -(x_0 - x^j, y_0 - y^j, z_0 - z^j)/||(x_0 - x^j, y_0 - y^j, z_0 - z^j)||$$

Unknown parameters: $\mathbf{x}$ is a (4×1) vector containing the deviation (dx, dy, dz) between the true and approximate coordinates, and the receiver clock offset δt.

6.1.1 Parameter Adjustment

Equation (6.8) was written neglecting the measurement noise and mismodelling (see equation (6.1)). If such errors $(\boldsymbol{\varepsilon})$ are explicitly written, then the linear model is as follows:

$$\mathbf{y} = \mathbf{G}\,\mathbf{x} + \boldsymbol{\varepsilon} \tag{6.9}$$

where the error term $\boldsymbol{\varepsilon}$ is only known from some statistical properties, usually the mean $\mathbf{m} = \mathrm{E}[\boldsymbol{\varepsilon}]$ and covariance matrix $\mathbf{R} = \mathrm{E}[\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}^T]$.

Due to the error term $\boldsymbol{\varepsilon}$, in general equation (6.9) defines an incompatible system (i.e. there is no 'exact' solution fulfilling the system). In this context, the parameters' solution can be taken as the vector $\hat{\mathbf{x}}$ that minimises the discrepancy in the equation system. That is, the vector $\hat{\mathbf{x}}$ provides the 'best fit' of $\mathbf{y} \simeq \mathbf{G}\,\hat{\mathbf{x}}$ in a given sense.

A common criterion used in GNSS is the *least squares* adjustment, which is defined by the condition

$$\min \|\mathbf{y} - \hat{\mathbf{y}}\|^2 = \min \left[\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \right] \tag{6.10}$$

$$\text{where} \quad \hat{\mathbf{y}} = \mathbf{G}\,\hat{\mathbf{x}}$$

[3] The matrix $\mathbf{G}$ can be computed in ENU coordinates instead of XYZ as in equation (6.6). In this local system the rows are $[-\cos el^i \sin az^i, -\cos el^i \cos az^i, -\sin el^i, 1]$, where el^i and az^i are the elevation and azimuth angles of satellite i observed from the receiver's position (see equation (B.14)).

The discrepancy vector between the measurements $\mathbf{y}$ and the fitted model (i.e. $\hat{\mathbf{y}} = \mathbf{G}\,\hat{\mathbf{x}}$) is usually called the *postfit residual* vector:

$$\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{G}\,\hat{\mathbf{x}} \tag{6.11}$$

Thus, the least squares estimator solution defined by equation (6.10) gives the vector $\hat{\mathbf{x}}$ that minimises[4] the residual quadratic norm $||\mathbf{r}||^2$.

From the basic results of linear algebra, it follows that the solution fulfilling condition (6.10) is given by

$$\boxed{\hat{\mathbf{x}} = (\mathbf{G}^T\,\mathbf{G})^{-1}\mathbf{G}^T\,\mathbf{y}} \tag{6.12}$$

Substituting equations (6.12) and (6.9) in (6.11), the postfit residual vector is

$$\mathbf{r} = \left[\mathbf{I} - \mathbf{G}(\mathbf{G^T}\,\mathbf{G})^{-\mathbf{1}}\mathbf{G^T}\right]\,\mathbf{y} = \mathbf{S}\,\mathbf{y} = \mathbf{S}\,\boldsymbol{\varepsilon}$$

where $\mathbf{S}$ is a symmetric and idempotent (*projection*) matrix (6.13)

$$\mathbf{S} = \mathbf{I} - \mathbf{G}(\mathbf{G}^T\,\mathbf{G})^{-1}\mathbf{G}^T,\;\; \mathbf{S}^T = \mathbf{S},\;\; \mathbf{S}^2 = \mathbf{S},\;\; \mathbf{r} = \mathbf{S}\,\mathbf{y} \perp \hat{\mathbf{y}}$$

From equations (6.12) and (6.9) the error of the estimate can be written as

$$\hat{\mathbf{x}} - \mathbf{x} = (\mathbf{G}^T\,\mathbf{G})^{-1}\mathbf{G}^T\,\boldsymbol{\varepsilon} \tag{6.14}$$

Assuming that the prefit residuals have mean zero errors ($\mathrm{E}[\boldsymbol{\varepsilon}] = 0$) and covariance matrix $\mathbf{R}$, then the mean error, covariance matrix and Mean-Square Error (MSE) of the estimate are given by

$$\begin{aligned}
\mathbf{m} &= \mathrm{E}[\hat{\mathbf{x}} - \mathbf{x}] = (\mathbf{G}^T\,\mathbf{G})^{-1}\mathbf{G}^T\,\mathrm{E}[\boldsymbol{\varepsilon}] = 0 \\
\mathbf{P} &= \mathrm{E}[(\hat{\mathbf{x}} - \mathbf{x})\,(\hat{\mathbf{x}} - \mathbf{x})^T] = (\mathbf{G}^T\,\mathbf{G})^{-1}\mathbf{G}^T\,\mathrm{E}[\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}^T]\,\mathbf{G}(\mathbf{G}^T\,\mathbf{G})^{-1} \\
&= (\mathbf{G}^T\,\mathbf{G})^{-1}\mathbf{G}^T\;\mathbf{R}\;\mathbf{G}(\mathbf{G}^T\,\mathbf{G})^{-1} \\
\mathrm{MSE} &= \mathrm{E}[(\hat{\mathbf{x}} - \mathbf{x})^T\,(\hat{\mathbf{x}} - \mathbf{x})] = \mathrm{tr}(\mathbf{P})
\end{aligned} \tag{6.15}$$

The last two expressions become simpler by assuming uncorrelated values with identical variance σ^2. That is, by taking $\mathbf{R} = \mathrm{E}[\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}^T] = \sigma^2\,\mathbf{I}$, then

$$\mathbf{P} = \sigma^2\,(\mathbf{G}^T\,\mathbf{G})^{-1}\,, \quad \mathrm{MSE} = \sigma^2\,\mathrm{tr}[(\mathbf{G}^T\,\mathbf{G})^{-1}] \tag{6.16}$$

6.1.1.1 Weighted Least Squares (WLS) Solution

In the least squares fit of equation (6.10) all measurements have been treated in the same way. Nevertheless, in a real scenario not all the measurements have the same quality, that is the same error. For instance, measurements at low elevation have larger multipath error than measurements at high elevation (see Fig. 4.5), or the modelling errors due to the tropospheric or ionospheric mapping functions are larger at lower elevations, as well.

[4] Equation (6.10), in which a quadratic sum is minimised, could be interpreted in physical terms as minimising the energy of the error fit. Thus the estimate $\hat{\mathbf{x}}$ can be seen as an equilibrium solution.

The measurement quality can be incorporated into the fitting criterion (6.10) by introducing a symmetric, positive-definite weighting matrix $\mathbf{W}$ and redefining the 'best fit' of equation (6.10) as the condition[5]

$$\min \|\mathbf{y} - \hat{\mathbf{y}}\|_{\mathbf{W}}^2 \tag{6.17}$$

where the norm of the residual vector $\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}}$ to be minimised is, now, associated with the scalar product defined by the weighting matrix $\mathbf{W}$.

With this weighting criterion, estimator $\hat{\mathbf{x}}_{\mathbf{W}}$ and its covariance matrix $\mathbf{P}_{\mathbf{W}}$ become

$$\boxed{\hat{\mathbf{x}}_{\mathbf{W}} = (\mathbf{G}^T \mathbf{W}\, \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}\, \mathbf{y}} \tag{6.18}$$

$$\boxed{\mathbf{P}_{\mathbf{W}} = (\mathbf{G}^T \mathbf{W}\, \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}\, \mathbf{R}\, \mathbf{W}\, \mathbf{G} (\mathbf{G}^T \mathbf{W}\, \mathbf{G})^{-1}} \tag{6.19}$$

6.1.1.2 Best Linear Unbiased Minimum Variance Estimator

The weighting matrix $\mathbf{W}$ was introduced in the previous section as a way to account for the different quality of the data in the adjustment problem, but no criteria were given to define this matrix.

To this end, we can note that equations (6.18) and (6.19) are simplified when we take the weighting matrix $\mathbf{W}$ as the inverse of the covariance matrix $\mathbf{R}$. That is, when

$$\mathbf{W} = \mathbf{R}^{-1} \tag{6.20}$$

Then equations (6.18) and (6.19) become

$$\boxed{\hat{\mathbf{x}} = (\mathbf{G}^T \mathbf{R}^{-1}\, \mathbf{G})^{-1} \mathbf{G}^T \mathbf{R}^{-1}\, \mathbf{y}} \tag{6.21}$$

$$\boxed{\mathbf{P} = (\mathbf{G}^T \mathbf{R}^{-1}\, \mathbf{G})^{-1}} \tag{6.22}$$

Using a different approach, it can be shown (see for instance [Tapley et al., 2004] or [Bierman, 1976]) that this solution corresponds to the Best Linear Unbiased Minimum Variance Estimator (BLUE).

The *minimum variance criterion* is widely used because of its simplicity. It has the advantage that a complete statistical description of the random errors is not required. Rather, only the first- and second-order statistics for the measurement error are needed[6] (i.e. $\mathrm{E}[\boldsymbol{\varepsilon}] = 0$, $\mathbf{R} = \mathrm{E}[\boldsymbol{\varepsilon}\, \boldsymbol{\varepsilon}^T]$).

Unfortunately, the characterisation of measurement error is very difficult and even the covariance matrix is not usually known. A simplification commonly used is to assume that the measurements (prefit residuals

[5] Note that, if $\mathbf{W}$ is a diagonal matrix, then the norm $\|\mathbf{y} - \hat{\mathbf{y}}\|_{\mathbf{W}}^2 = \sum w_i (y_i - \hat{y}_i)^2$, where the terms in the sum are weighted by the diagonal elements w_i. Thus, it follows that the measurements associated with large w_i values will contribute more to the adjustment.

[6] The minimum variance estimate (6.21) gives the maximum likelihood estimate when the observation errors are assumed to be distributed normally with zero mean and covariance $\mathbf{R}$.

$\mathbf{y} = (y_1, \ldots, y_n)^T$) from the different satellites are uncorrelated. Then, the weighting matrix $\mathbf{W}$ becomes

$$W = \mathbf{R}^{-1} = \begin{bmatrix} 1/\sigma_{y_1}^2 & & \\ & \ddots & \\ & & 1/\sigma_{y_n}^2 \end{bmatrix} \tag{6.23}$$

where $\sigma_{y_i}^2$ comes from the uncertainty of the different error sources (satellite clocks, ephemeris, ionosphere, troposphere, multipath and receiver noise)

$$\sigma_{y_i}^2 \equiv \sigma_{\mathrm{UERE}_i}^2 = \sigma_{clk_i}^2 + \sigma_{eph_i}^2 + \sigma_{iono_i}^2 + \sigma_{tropo_i}^2 + \sigma_{mp_i}^2 + \sigma_{noise_i}^2 \tag{6.24}$$

(see for instance [RTCA-MOPS, 2006]), where UERE stands for User Equivalent Range Error. A simpler model for equation (6.24) is to take just an elevation-dependent function according to the expression

$$\sigma_{y_i} = a + b\, \mathrm{e}^{-elev/c} \tag{6.25}$$

6.1.1.3 Block-wise WLS

Consider two linear $[m_1 \times n], [m_2 \times n]$ equation systems sharing the same unknown parameter vector $\mathbf{x}$:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{G}_1\, \mathbf{x}, \qquad \mathbf{R}_1 \\ \mathbf{y}_2 &= \mathbf{G}_2\, \mathbf{x}, \qquad \mathbf{R}_2 \end{aligned} \tag{6.26}$$

where $\mathbf{R}_1$ and $\mathbf{R}_2$ are the covariance matrices of measurement vectors $\mathbf{y}_1$ and $\mathbf{y}_2$.

The two systems can then be combined into a common $[(m_1 + m_2) \times n]$ system as

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix} \mathbf{x}\,, \qquad \mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 \end{bmatrix} \tag{6.27}$$

where no correlation between the two measurement vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ is assumed in matrix $\mathbf{R}$.

From equations (6.21) and (6.22), it is easy to show that, by taking the corresponding augmented matrices $\mathbf{y}$ and $\mathbf{G}$, the WLS solution of the system (6.27) yields

$$\hat{\mathbf{x}} = \left[\mathbf{G}_1^T\, \mathbf{R}_1^{-1}\, \mathbf{G}_1 + \mathbf{G}_2^T\, \mathbf{R}_2^{-1}\, \mathbf{G}_2\right]^{-1} \left[\mathbf{G}_1^T\, \mathbf{R}_1^{-1}\, \mathbf{y}_1 + \mathbf{G}_2^T\, \mathbf{R}_2^{-1}\, \mathbf{y}_2\right] \tag{6.28}$$

$$\mathbf{P} = \left[\mathbf{G}_1^T\, \mathbf{R}_1^{-1}\, \mathbf{G}_1 + \mathbf{G}_2^T\, \mathbf{R}_2^{-1}\, \mathbf{G}_2\right]^{-1} \tag{6.29}$$

Recursive computation: From the previous approach, the recursive computation of estimate $\mathbf{x}$ can be written as

$$\begin{aligned} \mathbf{P}_1 &= \left[\mathbf{G}_1^T\, \mathbf{R}_1^{-1}\, \mathbf{G}_1\right]^{-1} \\ \hat{\mathbf{x}}_{(1)} &= \mathbf{P}_1 \cdot \left[\mathbf{G}_1^T\, \mathbf{R}_1^{-1}\, \mathbf{y}_1\right] \\ \mathbf{P}_2 &= \left[\mathbf{P}_1^{-1} + \mathbf{G}_2^T\, \mathbf{R}_2^{-1}\, \mathbf{G}_2\right]^{-1} \\ \hat{\mathbf{x}}_{(2)} &= \mathbf{P}_2 \cdot \left[\mathbf{P}_1^{-1}\, \mathbf{x}_{(1)} + \mathbf{G}_2^T\, \mathbf{R}_2^{-1}\, \mathbf{y}_2\right] \end{aligned} \tag{6.30}$$

Note that, if only the final estimate is desired, it is best not to process data sequentially using equation (6.30), but instead to apply equations (6.21) and (6.22) and accumulate the equations without solving them until the end [Bierman, 1976]. This could be especially useful in the case of numerical instabilities, because it avoids the propagation of numerical inaccuracies during the recursive steps.

Constraints: A priori information can be added to the linear system (6.26) as *constraining equations* $\boldsymbol{\lambda} = \mathbf{A}\mathbf{x}$ with a given weight $\mathbf{W} = \mathbf{R}_{\boldsymbol{\lambda}}{}^{-1}$. Indeed,

$$\begin{aligned} \mathbf{y} &= \mathbf{G}\ \mathbf{x}, & \mathbf{R} \\ \boldsymbol{\lambda} &= \mathbf{A}\ \mathbf{x}, & \mathbf{R}_{\boldsymbol{\lambda}} \end{aligned} \tag{6.31}$$

6.1.2 Kalman Filter

The principle of Kalman filtering can be roughly summarised as the WLS solution of the linearised observation system augmented with a prediction of the estimate as additional equations. The predicted estimate and the weighted solution are given as follows.

Predicted estimate (from a simple linear model): Let $\widehat{\mathbf{x}}(n-1)$ be the estimate for the $(n-1)$th epoch; then a prediction for the next epoch $\widehat{\mathbf{x}}^{-}(n)$ is computed according to the model[7]

$$\boxed{\begin{aligned} \widehat{\mathbf{x}}^{-}(n) &= \boldsymbol{\Phi}(n-1)\ \widehat{\mathbf{x}}(n-1) \\ \mathbf{P}^{-}_{\widehat{\mathbf{x}}(n)} &= \boldsymbol{\Phi}(n-1)\ \mathbf{P}_{\widehat{\mathbf{x}}(n-1)}\ \boldsymbol{\Phi}^{T}(n-1) + \mathbf{Q}(n-1) \end{aligned}} \tag{6.32}$$

where $\boldsymbol{\Phi}$ is called the transition matrix and defines the propagation of the vector parameter estimate $\widehat{\mathbf{x}}$, and $\mathbf{Q}$ is the process noise matrix. Matrix $\mathbf{Q}$ allows us, in particular, to add some additional noise to account for possible mismodelling due to the simple prediction model used or, almost the same, to an inexact description of the problem in general.

Weighted solution (from measurements and predicted estimate): According to the approach in section 6.1.1.3, the measurements (i.e. the linearised observation equations) are combined with the estimate of the predicted parameters as follows:

$$\begin{bmatrix} \mathbf{y}(n) \\ \widehat{\mathbf{x}}^{-}(n) \end{bmatrix} = \begin{bmatrix} \mathbf{G}(n) \\ \mathbf{I} \end{bmatrix} \mathbf{x}(n) \qquad \mathbf{P} = \begin{bmatrix} \mathbf{R}(n) & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{-}_{\widehat{\mathbf{x}}(n)} \end{bmatrix} \tag{6.33}$$

which is solved like equation(6.27), being the WLS estimate:

$$\boxed{\begin{aligned} \mathbf{P}_{\widehat{\mathbf{x}}(n)} &= \left[\mathbf{G}^{T}(n)\ \mathbf{R}^{-1}(n)\ \mathbf{G}(n) + \left(P^{-}_{\widehat{\mathbf{x}}(n)}\right)^{-1}\right]^{-1} \\ \widehat{\mathbf{x}}(n) &= \mathbf{P}_{\widehat{\mathbf{x}}(n)} \cdot \left[\mathbf{G}^{T}(n)\ \mathbf{R}^{-1}(n)\ \mathbf{y}(n) + \left(\mathbf{P}^{-}_{\widehat{\mathbf{x}}(n)}\right)^{-1} \widehat{\mathbf{x}}^{-}(n)\right] \end{aligned}} \tag{6.34}$$

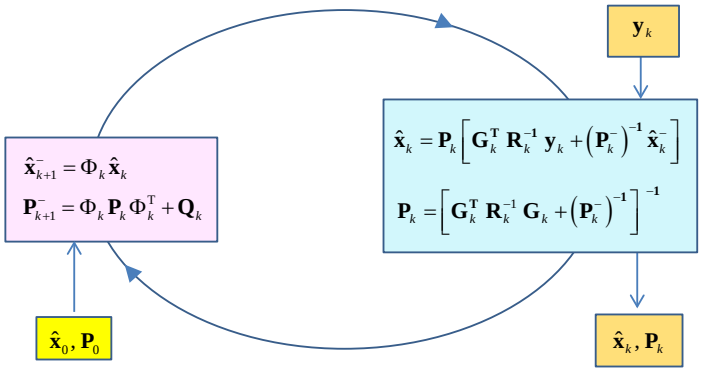


Figure 6.2: Kalman filter diagram. Notation: $\mathsf{R}_k = \mathsf{R}(k)$, $\mathsf{P}_k = \mathsf{P}_{\hat{\mathsf{x}}(k)}$.

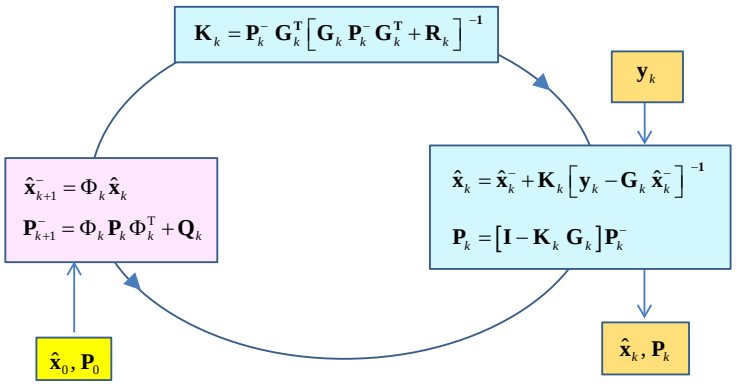


Figure 6.3: Classical formulation of Kalman filter.

The algorithm can be summarised in the scheme shown in Fig. 6.2.[8] See exercise 8 of session 5.2 in Volume II.

Using the relations [Bierman, 1976]

$$\begin{aligned}(ACB + D)^{-1} &= D^{-1} - D^{-1}AMBD^{-1}\\ M &= (BD^{-1}A + C^{-1})^{-1}\end{aligned} \tag{6.35}$$

it can be shown that the previous formulation is algebraically equivalent to the classical formulation of the Kalman filter given in Fig. 6.3.

6.1.2.1 Some Elemental Examples of Matrix Definitions Φ and Q

The determination of the state transition matrix $\mathbf{\Phi}$ and process noise matrix $\mathbf{Q}$ is usually based on physical models describing the estimation problem. For instance, for satellite tracking or orbit determination, they are derived from the orbital motion equations. Nevertheless, for the purpose of this book (i.e. SPP and PPP), only very simple formulations are needed. Such elemental formulations are covered by the following.

[7] This is a first-order Gauss–Markov model. The dynamic character is established through the state transition matrix $\mathbf{\Phi}$ and the noise matrix of the process $\mathbf{Q}$.

[8] For readers who wish to go deeper into this theme, the book by [Bierman, 1976] is recommended, especially the chapters relating to the U-D covariance filter and SRIF.

6.1.2.1.1 Static Positioning

The state vector to be determined is given by $\widehat{\mathbf{x}} = (dx, dy, dz, \delta t)$ where the coordinates[9] are considered as *constants* (because the receiver is kept fixed) and the clock offset can be modelled as *white noise* with zero mean. Under these conditions matrices $\mathbf{\Phi}$ and $\mathbf{Q}$ are given by

$$\mathbf{\Phi}(n) = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{bmatrix} \qquad \mathbf{Q}(n) = \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & \sigma^2_{\delta t} \end{bmatrix} \tag{6.36}$$

where σ_{dt} is the uncertainty in the clock prediction model (for instance, $\sigma_{dt} = 1\,\text{ms} = 300\,\text{km}$ for an unknown clock (i.e. 1 leap millisecond). Note that the prediction model for the coordinates is exact and, therefore, the associated elements in matrix $\mathbf{Q}$ are null.

6.1.2.1.2 Kinematic Positioning

1. If a vehicle is moving at a high velocity, the coordinates can be modelled as *white noise* with zero mean, the same as the clock offset:

$$\mathbf{\Phi}(n) = \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix} \quad \mathbf{Q}(n) = \begin{bmatrix} \sigma^2_{dx} & & & \\ & \sigma^2_{dy} & & \\ & & \sigma^2_{dz} & \\ & & & \sigma^2_{\delta t} \end{bmatrix} \tag{6.37}$$

2. If a vehicle is moving at a low velocity, the coordinates can be modelled as a *random walk* process with its uncertainty growing with time:

$$\mathbf{\Phi}(n) = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{bmatrix} \quad \mathbf{Q}(n) = \begin{bmatrix} Q'_{dx}\Delta t & & & \\ & Q'_{dy}\Delta t & & \\ & & Q'_{dz}\Delta t & \\ & & & \sigma^2_{\delta t} \end{bmatrix} \tag{6.38}$$

6.1.3 Positioning Error

The formal, predicted and measured accuracy concepts are discussed in this section.

The formal accuracy is a measure of the uncertainty of the estimates, according to a statistical characterisation of the errors and linear model used for the position estimate.

The predicted accuracy provides the expected position accuracy based on a simple statistical description of the measurement errors. Its computation does not require the measurements, just the standard deviation σ, and the approximated satellites' and user's location coordinates. Thus, it

[9] We are referring to deviations from nominal values (x_0, y_0, z_0); that is, what is estimated from the navigation equations.

can be computed at any point and at any time with the almanac, without needing any measurements.

The measured accuracy is the measure of the true error and must be assessed with the actual measurements.

6.1.3.1 Formal accuracy

Denoting as P_{xx}, P_{yy}, P_{zz}, P_{tt} the diagonal elements of matrix $\mathbf{P}$ of equation (6.22), then the standard deviations are the *formal errors* of the estimated x, y, z and t components

$$\sigma_x = \sqrt{P_{xx}}, \quad \sigma_y = \sqrt{P_{yy}}, \quad \sigma_z = \sqrt{P_{zz}}, \quad \sigma_t = \sqrt{P_{tt}} \tag{6.39}$$

These expressions provide a characterisation of the quality of the coordinates and clock estimates (linked to the measurement error model assumed). Thus, they are not the actual errors, just a measure of the uncertainty of the error estimates.

The previous expressions give the errors in ECEF XYZ coordinates. Nevertheless, it is usually more meaningful to a user to think in terms of horizontal and vertical position error, or ENU coordinates.

Let $\mathbf{R}$ be the transformation matrix (6.40) of ENU coordinates to XYZ (i.e. whose columns are the unit orthogonal vectors $\{\hat{\mathbf{e}}, \hat{\mathbf{n}}, \hat{\mathbf{u}}\}$ as expressed in the XYZ coordinate system at a point of latitude φ and longitude λ), see equation (B.8):

$$\mathbf{R} = \begin{bmatrix} -\sin\lambda & -\sin\varphi\cos\lambda & \cos\varphi\cos\lambda \\ \cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi\sin\lambda \\ 0 & \cos\varphi & \sin\varphi \end{bmatrix} \tag{6.40}$$

Hence, $\mathbf{P}_{enu} = \mathbf{R}^T \mathbf{P}_{xyz} \mathbf{R}$ is given, where $\mathbf{P}_{xyz}$ is the submatrix of $\mathbf{P}$ containing solely the geometric components.[10] From this covariance matrix, they can be defined as

$$\sigma_{East} = \sqrt{P_{ee}}, \quad \sigma_{North} = \sqrt{P_{nn}}, \quad \sigma_{Up} = \sqrt{P_{uu}} \tag{6.41}$$

For the horizontal error, the following are usually also defined:

$$\sigma_{Horizontal} = \sqrt{P_{ee} + P_{nn}} \tag{6.42}$$

It is usual to express the horizontal error in two main directions where the covariance matrix $\mathbf{P}_{enu}$ is diagonal. They define two orthogonal axes, associated with the minor and major axis of the error ellipse. The major axis is given by

$$\sigma_{H_{major}} = \sqrt{\frac{P_{ee} + P_{nn}}{2} + \sqrt{\left(\frac{P_{ee} - P_{nn}}{2}\right)^2 + P_{en}^2}} \tag{6.43}$$

[10] The matrix $\mathbf{P}_{enu}$ can be obtained directly in ENU coordinates, using the geometry matrix $\mathbf{G}$ computed in ENU coordinates. In this local system the rows of matrix $\mathbf{G}$ (of linear model (6.6)) are $[-\cos el^i \sin az^i, -\cos el^i \cos az^i, -\sin el^i, 1]$, where el^i and az^i are the elevation and azimuth angles of satellite i observed from the receiver's position, see equation (B.14) in Appendix B.

6.1.3.2 Predicted Accuracy: Dilution Of Precision

As explained earlier, assuming the simple statistical model

$$\mathbf{m} = \mathrm{E}[\boldsymbol{\varepsilon}] = 0, \qquad \mathbf{R} = \mathrm{E}[\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}^T] = \sigma^2\,\mathbf{I} \tag{6.44}$$

the covariance matrix of the error of the estimate is given by equation (6.16) as

$$\mathbf{P} = \sigma^2\,(\mathbf{G}^T\,\mathbf{G})^{-1}$$

This expression depends upon two factors: (1) the variance of the measurements (prefit residuals) σ^2; and (2) the geometry matrix ($\mathbf{G}$), which is linked only to the receiver–satellite ray geometry, see equation (6.6).

The Root Mean Square Error (RMSE) is given by

$$\mathrm{RMSE} = \sqrt{\mathrm{tr}(\mathbf{P})} = \sigma\sqrt{\mathrm{tr}\,[(\mathbf{G}^T\,\mathbf{G})^{-1}]} \tag{6.45}$$

which means that the trace of matrix $(\mathbf{G}^T\,\mathbf{G})^{-1}$ is a scale factor on σ for the RMSE.

Note that, as the matrix $\mathbf{G}$ does not depend on the measurements, but only on the geometry, it can be computed from the almanac (because accurate satellite positions are not needed); that is, it does not require receiver measurements.

On the basis of this simple approach, the following Dilution Of Precision (DOP) parameters are defined:

$$\mathbf{Q} \equiv (\mathbf{G}^T\,\mathbf{G})^{-1} = \begin{bmatrix} q_{xx} & q_{xy} & q_{xz} & q_{xt} \\ q_{xy} & q_{yy} & q_{yz} & q_{yt} \\ q_{xz} & q_{yz} & q_{zz} & q_{zt} \\ q_{xt} & q_{yt} & q_{zt} & q_{tt} \end{bmatrix} \tag{6.46}$$

- *Geometric Dilution Of Precision:*

$$\mathrm{GDOP} = \sqrt{q_{xx} + q_{yy} + q_{zz} + q_{tt}} \tag{6.47}$$

- *Position Dilution Of Precision:*

$$\mathrm{PDOP} = \sqrt{q_{xx} + q_{yy} + q_{zz}} \tag{6.48}$$

- *Time Dilution Of Precision:*

$$\mathrm{TDOP} = \sqrt{q_{tt}} \tag{6.49}$$

As in the previous case, using equation (6.40) the submatrix $\mathbf{Q}_{xyz}$ of $\mathbf{Q}$ can be transformed to ENU coordinates as $\mathbf{Q}_{enu} = \mathbf{R}^T\,\mathbf{Q}_{xyz}\,\mathbf{R}$, in order to define the following:

- *Horizontal Dilution Of Precision:*

$$\mathrm{HDOP} = \sqrt{q_{ee} + q_{nn}} \tag{6.50}$$

- *Vertical Dilution Of Precision:*

$$\text{VDOP} = \sqrt{q_{uu}} \tag{6.51}$$

Hence, estimations of the expected accuracy are given by

GDOP σ	geometric precision in position and time
PDOP σ	precision in position
TDOP σ	precision in time
HDOP σ	precision in horizontal positioning
VDOP σ	precision in vertical positioning

where, basically, DOP represents an approximate ratio factor between the precision in the measurements (σ) and that in positioning. This ratio is computed only from the satellite–receiver geometry. See exercises 6 and 7 of session 5.2 in Volume II.

6.1.3.3 Measured Accuracy

Let ΔE_i, ΔN_i and ΔU_i be the errors in the east, north and vertical components of the ith position estimate sample. The RMS vertical, horizontal (2D) and 3D errors are defined as[11]

$$\text{RMS vertical error} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\Delta U_i^2} \tag{6.52}$$

$$\text{2D-RMS horizontal error} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\Delta E_i^2 + \Delta N_i^2\right)^2} \tag{6.53}$$

$$\text{3D-RMS error} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\Delta E_i^2 + \Delta N_i^2 + \Delta U_i^2\right)^2} \tag{6.54}$$

Other measures of the quality of the position estimates are:

- 50th or 95th percentiles of horizontal, vertical and 3D errors;
- CEP (Circular Error Probable), as the 50th percentile of horizontal error
  (CEP is defined thus for this chapter only);
- SEP (Spherical Error Probable), as the 50th percentile of 3D error; and
- 2drms: calculated as twice the 2D-RMS horizontal error given by (6.53).

Assuming that the position estimates follow a multivariate normal distribution centred at the true position and the errors in east, north and

[11] Note that the RMS equals the standard derivation s only when the mean position error is zero. Indeed, let $m = (1/n)\sum_{i=1}^{n} x_i$, $\text{RMS} = \sqrt{(1/n)\sum_{i=1}^{n} x_i^2}$ and $s = \sqrt{(1/n)\sum_{i=1}^{n}(x_i - m)^2}$; then $\text{RMS}^2 = s^2 + m^2$.

up components, (E, N, U), are uncorrelated, and assumig $\sigma_E \simeq \sigma_N$ and $\sigma_U \simeq 2(\sigma_E^2 + \sigma_N^2)^{1/2}$, then the following relations can be considered [Misra and Enge, 2001]:

RMS vertical error	$\simeq 0.5 \times$ vertical error (95%)
	$\simeq$ 2drms
	$\simeq 0.9 \times$ 3D-RMS error
2D RMS error	$= 0.5 \times$ 2drms
	$\simeq 0.6 \times$ horizontal error (95%)
	$\simeq 1.2 \times$ CEP
3D RMS error	$\simeq 2.2 \times$ 2D-RMS error
	$\simeq 1.2 \times$ horizontal error (95%)
	$\simeq 1.3 \times$ SEP

A wider discussion of accuracy equivalences can be found in [Diggelen, 2007].

6.2 Code and Carrier-Based Positioning

For high-accuracy positioning, the carrier phase must be used, besides the code pseudorange. As commented in section 4.2, the carrier measurements are very precise, typically at the level of a few millimetres, but contain unknown ambiguities which change every time the receiver locks the signal after a cycle slip. Nevertheless, such ambiguities can be estimated in the navigation solution, together with the coordinates and other parameters.

6.2.1 Precise Modelling Terms for PPP

The PPP technique allows centimetre-level accuracy to be achieved for static positioning and decimetre level, or better, for kinematic positioning.[12] This high accuracy requires accurate measurement modelling, where all model terms described in the previous chapter must be taken into account (up to centimetre level or better).

This modelling involves consideration of the following terms:

Precise satellite orbits and clocks:

The precise orbits and clock files (see section 3.3.3) must be used instead of the broadcast ones used in the SPP. The polynomial in equation (3.25) can be applied to interpolate the precise orbits. Note that the orbits are referred to the satellite mass centre, hence equation (5.79) must be applied to compute the $\boldsymbol{\Delta}_{\mathrm{APC}}$ vector offset.

The satellite clocks should not be interpolated and thus only epochs having clocks available must be used (see section 3.3.3).

Relativistic effects: The gravitational path range correction, equation (5.15), can be added to the satellite clock correction due to the orbital eccentricity, equation (5.19), considered in the SPP.

[12] That is, the coordinates are treated as white-noise parameters in the navigation filter.

Atmospheric effects:

The ionospheric refraction and DCBs are removed using the ionosphere-free combination of measurements (see details in section 5.4.1.1).

The tropospheric refraction can be modelled by equation (5.65), where the mapping of Niell is used (see section 5.4.2.2.1). The dry and wet tropospheric delays are given by equation (5.66). The deviation of the zenith tropospheric delay $\Delta T_{z,wet}$ with respect to the nominal $T_{z_0,wet}$ must be estimated in the Kalman filter, together with the coordinates, clock and carrier phase biases.

Antenna biases and orientation:

The satellite and receiver antenna phase centres can be found in the IGS ANTEX files, after GPS week 1400 (see section 5.6.1). Before this week, the values of Table 5.4 can be used.

The carrier phase wind-up effect due to the satellite's motion is given by equation (5.71).

Satellites under eclipse conditions should be removed from the computation due to the large orbit error. The eclipse condition is given by equation (5.14).

Earth deformation effects:

Solid tides can be modelled by equations (5.83), (5.84) and (5.85).

Ocean loading and pole tides are second-order effects and can be neglected for PPP accuracies at the centimetre level (see comments in sections 5.7.2 and 5.7.3).

6.2.2 Linear Observation Model for PPP

Based on equations (4.19), code and carrier measurements in the ionosphere-free combination are modelled as follows:

$$\begin{aligned} R_C^j &= \rho^j + c(\delta t - \delta t^j) + Tr^j + \mathcal{M}_C^j + \varepsilon_C^j \\ \Phi_C^j &= \rho^j + c(\delta t - \delta t^j) + Tr^j + \lambda_N\, w^j + B_C^j + m_C^j + \epsilon_C^j \end{aligned} \tag{6.55}$$

where R_C^j is the unsmoothed code pseudorange measurement for the jth satellite in view and Φ_C^j is the corresponding carrier measurement. Remark: ρ is referred to the APCs in the ionosphere-free combination (see equation (5.81)).

Following the same procedure as in section 6.1, the linear observation model $\mathbf{y} = \mathbf{G}\,\mathbf{x}$ for the code and carrier measurements can be written as follows.

Prefit residuals:

$$\mathbf{y} = \begin{bmatrix} R_C^1 - \rho_0^1 + c\,\delta t^1 - Tr_0^1 \\ \Phi_C^1 - \rho_0^1 + c\,\delta t^1 - Tr_0^1 - \lambda_N\, w^1 \\ \vdots \\ R_C^n - \rho_0^n + c\,\delta t^n - Tr_0^n \\ \Phi_C^n - \rho_0^n + c\,\delta t^n - Tr_0^n - \lambda_N\, w^n \end{bmatrix} \tag{6.56}$$

Note: The satellite clock offset δt^j includes the satellite clock relativistic correction due to the orbital eccentricity, equation (5.19). The relativistic path range correction (5.15) is included in the geometric range ρ_0^j.

The term Tr_0 is the nominal value for the tropospheric correction. Note that, according to equation (5.66), the tropospheric delay in equation (6.55) can be decomposed into a nominal term $Tr_0(E)$ and the deviation from this nominal term $M_{wet}(E)\,\Delta Tr_{z,wet}$. That is,

$$\begin{aligned} Tr(E) &= Tr_0(E) + M_{wet}(E)\,\Delta Tr_{z,wet} \\ Tr_0(E) &= Tr_{z,dry}\,M_{dry}(E) + Tr_{z_0,wet}\,M_{wet}(E) \end{aligned} \tag{6.57}$$

The mapping factor $M_{wet}(E)$ is an element of the design matrix (6.58) and $\Delta Tr_{z,wet}$ is a component of the parameter vector (6.59), as follows:

Design matrix

$$\mathbf{G} = \begin{bmatrix} \frac{x_0-x^1}{\rho_0^1} & \frac{y_0-y^1}{\rho_0^1} & \frac{z_0-z^1}{\rho_0^1} & 1 & M_{wet}^1 & 0 & \dots & 0 & \dots & 0 \\ \frac{x_0-x^1}{\rho_0^1} & \frac{y_0-y^1}{\rho_0^1} & \frac{z_0-z^1}{\rho_0^1} & 1 & M_{wet}^1 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ \frac{x_0-x^k}{\rho_0^k} & \frac{y_0-y^k}{\rho_0^k} & \frac{z_0-z^k}{\rho_0^k} & 1 & M_{wet}^k & 0 & \dots & 0 & \dots & 0 \\ \frac{x_0-x^k}{\rho_0^k} & \frac{y_0-y^k}{\rho_0^k} & \frac{z_0-z^k}{\rho_0^k} & 1 & M_{wet}^k & 0 & \dots & \underbrace{1}_{k} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ \frac{x_0-x^n}{\rho_0^n} & \frac{y_0-y^n}{\rho_0^n} & \frac{z_0-z^n}{\rho_0^n} & 1 & M_{wet}^n & 0 & \dots & 0 & \dots & 0 \\ \frac{x_0-x^n}{\rho_0^n} & \frac{y_0-y^n}{\rho_0^n} & \frac{z_0-z^n}{\rho_0^n} & 1 & M_{wet}^n & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \tag{6.58}$$

Vector parameters (to be estimated)

$$\mathbf{x} = \left[dx, dy, dz, c\,\delta t, \Delta Tr_{z,wet}, B_C^1, \dots, B_C^k, \dots, B_C^n\right]^T \tag{6.59}$$

6.2.3 Parameter Adjustment for PPP

The linear observation model $\mathbf{y} = \mathbf{G}\,\mathbf{x}$ can be solved using the Kalman filter, considering the carrier phase biases B_C^i as 'constant' along continuous phase arcs, and as 'white noise' at those instants when cycle slips occurs.

The following stochastic model can be used for the filter:[13]

Carrier phase biases (B_C) are taken as 'constant' along continuous phase arcs, and as 'white noise' when a cycle slip happens ($\sigma = 10^4$ m can be taken, for instance), see section 6.1.2.1.

Wet tropospheric delay ($\Delta Tr_{z,wet}$) is taken as a random walk process (a process noise with $d\sigma^2/dt = 1\,\text{cm}^2/\text{h}$, initialised with $\sigma_0^2 = 0.25\,\text{m}^2$ can be used for most of the applications), see section 6.1.2.1.2.

Receiver clock ($c\,\delta t$) is taken as a white-noise process (with $\sigma = 3 \cdot 10^5$ m, i.e. 1 ms for instance), see section 6.1.2.1.

Receiver coordinates (dx, dy, dz)

For static positioning the coordinates are taken as constants, see section 6.1.2.1.1.

For kinematic positioning the coordinates are taken as white noise or a random walk process as in section 6.1.2.1.2.

This solution procedure is called *floating* ambiguities – floating in the sense that the ambiguities are estimated by the filter 'as real numbers'. The bias estimations B^i will converge to a solution after a transition time that depends on the observation geometry, model quality and data noise. In general, one must expect errors at the decimetre level, or better, in pure kinematic positioning (after the best part of one hour) and at the centimetre level in static PPP. See exercises 3 and 4 of laboratory session 1.1 in Volume II.

6.3 Carrier Phase Ambiguity Fixing

As already mentioned, the carrier phase measurements are much more precise than the code pseudorange measurements (typically, about two orders of magnitude), but they contain the unknown ambiguities (B), see equation (6.55). If such ambiguities are fixed, then the carrier phase measurements become unambiguous pseudoranges, and accurate at the level of a few millimetres.

6.3.1 Double-Differenced Ambiguity Fixing

In the discussion presented below, the carrier ambiguities will be considered as double differences between pairs of receivers and satellites. This is done in order to cancel out the fractional part of the ambiguities (b_{rcv}, b^{sat}), leaving the remaining ambiguities as an integer number of wavelengths. That is, given

$$B_{rcv}^{sat} = \lambda\, N_{rcv}^{sat} + b_{rcv} + b^{sat} \tag{6.60}$$

the double differences, regarding a reference receiver and satellite, yield

$$\Delta\nabla B_{rcv}^{sat} = B_{rcv}^{sat} - B_{rcv_0}^{sat} - (B_{rcv}^{sat_0} - B_{rcv_0}^{sat_0}) = \lambda\, \Delta\nabla N_{rcv}^{sat} \tag{6.61}$$

where the satellite and receiver ambiguity terms (b_{rcv}, b^{sat}) cancel out.

[13] See the default configuration parameters in the `gLAB` tool.

6.3.1.1 Carrier Phase Ambiguity Fixing with Two Frequencies

A simple approach to ambiguity fixing is presented briefly as follows:[14]

Wide-lane ambiguity fixing: Double-differenced wide-lane ambiguity $\Delta\nabla N_W$ can be computed from the Melbourne–Wübbena combination, by rounding the average in time[15] (see equations (4.19)):

$$\Delta\nabla N_W = \left[\frac{\Delta\nabla\Phi_W - \Delta\nabla R_N}{\lambda_W}\right]_{\text{roundoff}} \tag{6.62}$$

This ambiguity estimate has the advantage that it can easily be obtained separately for each measurement thanks to the enlargement of the ambiguity spacing. Under moderate receiver noise and multipath conditions, a few minutes should be enough to fix the wide-lane ambiguity.

L1 ambiguity fixing: After fixing $\Delta\nabla N_W$, the double-differenced L1 ambiguity ($\Delta\nabla N_1$) can be fixed from the expression

$$\Delta\nabla N_1 = \left[\frac{1}{\lambda_N}\left(\Delta\nabla\hat{B}_C\right) - \frac{\lambda_W}{\lambda_2}\Delta\nabla N_W\right]_{\text{roundoff}} \tag{6.63}$$

when a sufficiently accurate estimate of the ambiguity $\Delta\nabla\hat{B}_C$ is available. This estimate $\Delta\nabla\hat{B}_C$ can be computed as in the previous section (PPP), by floating the B_C ambiguities in the Kalman filter.

As these ambiguities are determined 'by floating' them in the navigation filter, some time is needed (almost an hour) for the filter to converge. Indeed, roughly speaking, the ambiguity B_C is mostly estimated from $R_C - \Phi_C$, where the R_C code noise is about three times noisier than the code measurement in frequency f_1, see Fig. 4.3.

There are more complex methods available, such as the LAMBDA method ([Teunissen, 1996]; [Teunissen et al., 1997]), the Null Space method [Martin-Neira et al., 1995] or others [Kim and Langley, 2000], where the ambiguities are fixed 'as a set', and decorrelation and search (on integers) techniques are applied to enhance resolution of the ambiguity.[16]

Ionosphere-free bias fixed: After fixing the $\Delta\nabla N_W$ and $\Delta\nabla N_1$ ambiguities, the double-differenced ionosphere-free bias is fixed by

$$\Delta\nabla B_C = \lambda_N\left(\Delta\nabla N_1 + \frac{\lambda_W}{\lambda_2}\Delta\nabla N_W\right) \tag{6.64}$$

[14] The expressions below can be easily derived from equations (4.19) and (4.20).

[15] Note that the code noise is reduced in the narrow-lane combination R_N by a factor of $\sqrt{2}$ (see page 76), but on the other hand the ambiguity spacing is wider in the wide-lane combination (see section 4.1 and Figs 4.3 and 4.10). For instance, for the Galileo E1, E5b signals, $\lambda_W = 0.81$ cm, see Table 6.1; or, for the L1, L2 GPS signals, $\lambda_W = 0.86$ cm, see Table 4.1.

[16] MATLAB source code implementing the LAMBDA method is available on request at *http://www.citg.tudelft.nl/en/about-faculty/departments/geoscience-and-remote-sensing/research-themes/gps/lambda-method/*. See also [Strang and Borre, 1997].

Once the $\Delta\nabla B_C$ ambiguity is fixed, a double-differenced measurement two orders of magnitude more accurate than the code (typically) is available. This allows centimetre-level positioning accuracy to be achieved.

An alternative way to avoid the time span needed for the floated $\Delta\nabla\hat{B}_C$ term to converge is to use the following equation for fixing the Φ_1 ambiguity:

$$\Delta\nabla N_1 = \left[\frac{\Delta\nabla\Phi_1 - \Delta\nabla\Phi_2 - \Delta\nabla I - \lambda_2\Delta\nabla N_W}{\lambda_2 - \lambda_1}\right]_{\text{roundoff}} \tag{6.65}$$

However, there is a problem here, mainly with the ionospheric refraction term $\Delta\nabla I$.

For short baselines (up to 10–15 km, depending on the ionospheric conditions) it can be assumed that the double-differenced ionospheric refraction cancels out (i.e. $\Delta\nabla I \simeq 0$), and thus the $\Delta\nabla N_1$ ambiguity can be fixed by rounding the expression

$$\Delta\nabla N_1 = \left[\frac{\Delta\nabla\Phi_1 - \Delta\nabla\Phi_2 - \lambda_2\Delta\nabla N_W}{\lambda_2 - \lambda_1}\right]_{\text{roundoff}} \tag{6.66}$$

This is the approach applied in the Real-Time Kinematics (RTK) ambiguity fixing technique. Executable software implementing the RTK is available at *http://gpspp.sakura.ne.jp/rtklib/rtklib.htm*.

For long baselines, an accurate ionospheric correction estimate $\Delta\nabla\widehat{I}$ is needed to allow the user to fix the ambiguity by rounding equation (6.65). Note that, once the ambiguity N_W is fixed, since the carrier measurement noise is only a few millimetres, the main contribution to the rounding error is from the $\Delta\nabla\widehat{I}$ term. Note also that its accuracy must be better than $(\lambda_2 - \lambda_1)/2$ to allow integer rounding. In the case of the Galileo E1 and E5b signals, this accuracy threshold is $(\lambda_2 - \lambda_1)/2 = 2.9$ cm, see Table 6.1. For the L1 and L2 GPS signals, it is $(\lambda_2 - \lambda_1)/2 = 2.7$cm, see Table 4.1.

The Wide-Area Real-Time Kinematics (WARTK) technique has proved that accurate ionospheric corrections can be computed from the measurements gathered from permanent reference station networks with similar baselines, such as the SBAS (WAAS, EGNOS, etc.) networks. Such highly accurate ionospheric corrections allow carrier phase ambiguity fixing over a wide area. In this way, the WARTK technique can extend the RTK ambiguity fixing from a local to a wide area; that is, at a continental scale. For more details see [HJS et al., 2010] and [Juan et al., 2012b].

6.3.1.2 Carrier Phase Ambiguity Fixing with Three Frequencies

With two-frequency signals, in RTK or WARTK, the convergence time is linked to the wide-lane ambiguity ($\Delta\nabla N_W$) fixing. With three-frequency systems, using two close frequency signals it is possible to generate an extra-wide-lane signal to enable the single-epoch wide-lane ambiguity fixing. That is, without needing the time span required by equation (6.62) to smooth the code noise up to the roundoff threshold.

For a clearer presentation of concepts, the following definitions and notation will be used:

$$\Phi_{EW} = \Phi_{W_{23}}, \quad \Phi_{W} = \Phi_{W_{12}}, \quad R_{EN} = R_{N_{23}} \tag{6.67}$$

where the wide-lane and narrow-lane combinations are computed by combining the respective frequencies according to equations (4.6) and (4.7). Moreover, the frequencies and values of Table 6.1 will be taken as references for the numerical application.

With three-frequency signals, the following ambiguity fixing procedure can be considered.[17]

Wide-lane ambiguity fixing: Equation (6.62) used in the previous section is replaced by the two following equations (6.68) and (6.69).

For the extra-wide-lane ambiguity fixing,

$$\Delta\nabla N_{EW} = \left[\frac{\Delta\nabla\Phi_{EW} - \Delta\nabla R_{EN}}{\lambda_{EW}}\right]_{\text{roundoff}} \tag{6.68}$$

As shown in Fig. 4.3 on page 76, the code noise in the extra-narrow-lane combination $\Delta\nabla R_{EN}$ is reduced by a factor[18] of about $\sqrt{2}$ but, on the other hand, the ambiguity spacing is enlarged by the extra-wide-lane combination. If the frequencies in Table 6.1 are used, then $\lambda_{EW} = 9.678\,\text{m}$, which means about 10 times the λ_W and about 50 times the λ_1 ambiguity spacing.

Thus, since the measurement noise is far from the rounding threshold, the extra-wide-lane ambiguity ($\Delta\nabla N_{EW}$) can be fixed in a single epoch.

For the wide-lane ambiguity fixing,

$$\Delta\nabla N_{W} = \left[\frac{\lambda_{EW}\Delta\nabla N_{EW} + (\tilde{\alpha}_{EW} - \tilde{\alpha}_{W})\Delta\nabla I - (\Delta\nabla\Phi_{EW} - \Delta\nabla\Phi_{W})}{\lambda_{W}}\right]_{\text{roundoff}} \tag{6.69}$$

Using the values of Table 6.1, the coefficient of the ionospheric term $\Delta\nabla I$ is $\tilde{\alpha}_{EW} - \tilde{\alpha}_{W} = 0.63$. This allows an error up to 0.65 cm in the $\Phi_1 - \Phi_2$ delay (or 6.5 TECUs) for the ionospheric correction to fit into the rounding threshold (i.e. $\lambda_W/2$). That is, more than 20 times the error threshold allowed by equation (6.65).

Therefore, with the $\Delta\nabla N_{EW}$ ambiguity fixed, the limiting factor in this equation is the carrier phase multipath. Note that the carrier phase noise in the extra-wide-lane combination is several times larger than in the original signals. That is,

$$\sigma_{\Phi_{EW}} \simeq (\sqrt{\gamma_{23}+1}/(\sqrt{\gamma_{23}}-1))\sigma \simeq 55\,\sigma \tag{6.70}$$

where $\sigma \equiv \sigma_{\Phi_{E5b}} \simeq \sigma_{\Phi_{E5a}}$.

[17] Equations (6.68) and (6.69) are derived from equations (4.6) and (4.7).

[18] $\sigma_{R_{EN}} \simeq (\sqrt{\gamma_{23}+1}/(\sqrt{\gamma_{23}}+1))\sigma \simeq 1/\sqrt{2}\,\sigma$, where uncorrelated errors are assumed and with the same noise $\sigma \equiv \sigma_{R_{\text{E5b}}} \simeq \sigma_{R_{\text{E5a}}}$, see Table 6.1.

Narrow-lane ambiguity fixing: As in the two-frequency case, once the wide-lane ambiguity is fixed, the narrow-lane ambiguity can be fixed by either equation (6.63) or (6.65). After fixing these ambiguities, the unambiguous ionosphere-free bias $\Delta\nabla B_C$ can be obtained from equation (6.64).

If single-epoch ambiguity fixing is desired, then equation (6.65) must be used to fix $\Delta\nabla N_1$. That is,

$$\Delta\nabla N_1 = \left[\frac{\Delta\nabla\Phi_1 - \Delta\nabla\Phi_2 - \Delta\nabla I - \lambda_2\Delta\nabla N_W}{\lambda_2 - \lambda_1}\right]_{\text{roundoff}} \tag{6.71}$$

Note that, as in the two-frequency case studied in the previous section, this equation is affected by the ionospheric refraction decorrelation with distance, limiting the ambiguity fixing to short baselines, as in the Three Carrier Ambiguity Resolution (TCAR) technique, if an accurate prediction of $\Delta\nabla I$ is not available for the user. Thus, as in the previous case, WARTK allows the baselines to be extended by up to hundreds of kilometres as a result of the accurate ionospheric modelling.

Table 6.1: Values illustrating the ambiguity fixing concepts (taken from Tables 4.1 and 4.2). Note that the factors $\tilde{\alpha}_i$ are normalised with respect to $\alpha_2 - \alpha_1$, associated with the frequencies f_1, f_2 used for the ionosphere-free combination.

Signals	Wavelengths (m)	Combinations (m)	$\gamma_{ij} = (f_i/f_j)^2$
$f_1 \equiv f_{E1}$	$\lambda_1 = 0.190$	$\lambda_2 - \lambda_1 = 0.058$	$\gamma_{12} = (77/59)^2$
$f_2 \equiv f_{E5b}$	$\lambda_2 = 0.248$	$\lambda_W \equiv \lambda_{W_{12}} = 0.814$	$\gamma_{13} = (154/115)^2$
$f_3 \equiv f_{E5a}$	$\lambda_3 = 0.255$	$\lambda_{EW} \equiv \lambda_{W_{23}} = 9.768$	$\gamma_{23} = (118/115)^2$

Factors

$\tilde{\alpha}_W = \dfrac{\alpha_W}{\alpha_2 - \alpha_1} = \dfrac{\sqrt{\gamma_{12}}}{\gamma_{12} - 1}$	$\tilde{\alpha}_{EW} = \dfrac{\alpha_{EW}}{\alpha_2 - \alpha_1} = \sqrt{\gamma_{13}}\,\tilde{\alpha}_W$	$\tilde{\alpha}_{EW} - \tilde{\alpha}_W$
1.856	2.485	0.629

6.3.2 Undifferenced Ambiguity Fixing

As commented previously, the double-differenced ambiguities between pairs of satellites and receivers are integer numbers of wavelengths, see equation (6.61). Indeed, the fractional part cancels in such double differences:

$$\nabla\Delta b_{rcv}^{sat} = 0 \tag{6.72}$$

An immediate consequence of this equation is the separability of the fractional part of the ambiguities (for each satellite–receiver arc) in two independent terms, one of them linked only to the receiver and the other only to the satellite:[19]

$$\nabla\Delta b_{rcv}^{sat} = 0 \Longleftrightarrow b_{rcv}^{sat} = b_{rcv} + b^{sat} \tag{6.73}$$

[19]The proof of equation (6.73) is left as exercise for the reader. Note that the separability has been assumed de facto since the beginning of this book, see for instance equations (4.3), (4.19) or (6.61).

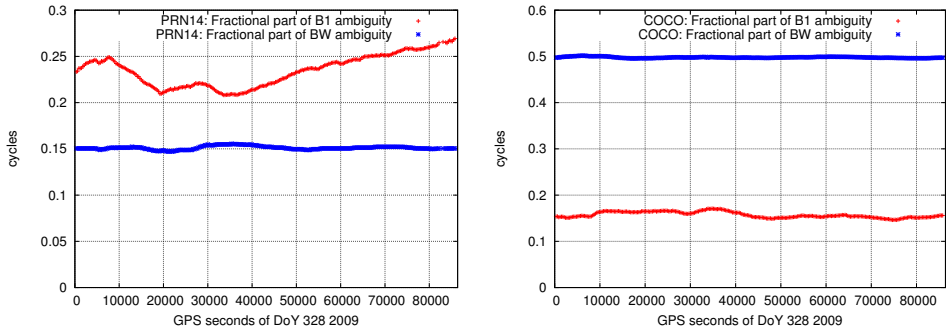


Figure 6.4: Carrier phase instrumental delays or the fractional part of ambiguities for the GPS satellite PRN14 (*left*) and receiver COCO (*right*). The wide lane is shown in blue and L1 in red. The vertical axis is in cycles and the horizontal one in seconds of day. The pattern in the figures is due to correlations with other parameters in the filter, mainly the satellite clocks.

This equation means that fractional parts of the ambiguity (b_{rcv} and b^{sat}) are not linked to a specific satellite–receiver arc, but b^{sat} depends only on the satellite (and it is common to all carrier measurements of receivers tracking this satellite) and b_{rcv} depends only on the receiver (and it is common to all satellites tracked by a given receiver).

The fractional part of the wide-lane ambiguity can be easily estimated from the MW combination, because it is uncorrelated with other parameters in the navigation filter (see Fig. 6.4, in blue),[20] while the fractional part of the short-lane ambiguity (i.e., b_1) can be added to the satellite and receiver clock. This approach leads to so-called *phase clocks* [Laurichesse and Mercier, 2007], which are different from the code clock (i.e. not consistent with the code clock).[21]

Another, more straightforward approach is to consider a fractional part of the ambiguities such as *carrier phase instrumental delays*,[22] and to remove them using accurate determinations of values computed from a receiver network [Juan et al., 2010]. Indeed, the carrier instrumental delays or fractional part of the ambiguities (b_{rcv} and b^{sat}) can be estimated from a global network of permanent stations, after fixing the double-differenced ambiguities $\nabla\Delta B^{sat}_{rcv} = \lambda\,\nabla\Delta N^{sat}_{rcv}$ between the satellites and receivers. As shown in Fig. 6.4, these estimates are stable enough to be broadcast to users as low-varying parameters.

Thus, with accurate determinations of such parameters, the user can remove the fractional part[23] and fix the remaining ambiguity as an integer number. This can be done in the same way as explained in previous sections, but in undifferenced mode, allowing, in this way, *real-time ambiguity fixing for PPP*.

[20] The fractional parts of wide-lane (*blue*) and short-lane (*red*) fractional ambiguity estimates are shown in the same figure. The larger variability seen for the short lane is due to the correlations with the other parameters in the filter, mainly the clocks.

[21] Nevertheless, this is not a critical issue for fixing the ambiguities, because the code is used mainly to initialise the filter (and to help its convergence).

[22] To be more precise, the residual value after removing the integer ambiguities. That is, $b = B - \lambda N$, see equations (4.19).

[23] Only the fractional ambiguities for the satellites must be broadcast to users, because those of the receiver are common to all satellites.

Note that a PPP-based ambiguity fixing approach can allow worldwide ambiguity fixing to be performed, because no baseline limitations apply,[24] thus improving the PPP accuracy. Nevertheless, as with PPP, a large convergence time is required for the filter to start the ambiguity fixing (see section 6.3.1.1).

Some works focusing on the undifferenced ambiguity fixing are as follows: [Juan et al., 2010], [Laurichesse and Mercier, 2007], [Banville et al., 2008], [HJS, 2010] and [Ge et al., 2008], among others.

6.3.3 Accelerating the Filter Convergence: Fast PPP

Accurate ionospheric corrections computed by a wide-area network can be used to accelerate the filter convergence in both undifferenced and differential positioning (in double differences). As is shown in [Juan et al., 2012a], accuracies better than 10 cm are achieved in a few minutes (typically less than 5 min) with two-frequency signals, while almost an hour is needed when the ionospheric corrections are not used (i.e. in classical PPP mode). Moreover, as in section 6.3.1.2, with three-frequency signals, accurate ionospheric corrections allow the single-epoch solution to be achieved.

A simple explanation of how the accurate ionospheric corrections (together with the differential code biases) help the filter convergence is given by the equation

$$\Phi_I - I - K_{21} = \frac{1}{\tilde{\alpha}_W}(B_W - B_C) \tag{6.74}$$

Indeed, this equation defines a link between the geometry-free bias ($B_I = \Phi_I - I - K_{21}$), the ionosphere-free bias B_C, and the wide-lane bias B_W, see equations (4.19), and can be used as a constraint in the user navigation filter.

As mentioned above, the wide-lane bias B_W can be estimated quickly (in a few minutes with two-frequency or in a single epoch with three-frequency signals). Thus, the ionosphere (I), together with the differential code bias (K_{21}), form a *bridge* to transfer this accuracy quickly achieved in the wide-lane bias estimate (B_W) to the ionosphere-free bias (B_C).

In other words, when the ionosphere (I) is available at the user location, then an ionosphere-free bias (B_C) estimate can be computed from equation (6.74) and used by the navigation filter as additional data to help estimate this carrier phase ambiguity (B_C). But these data are several times more accurate than the ionosphere-free code measurement, which is driving filter convergence during the first epochs in the classical PPP approach.

This new approach is called Fast Precise Point Positioning (F-PPP) and is patent-protected [HJS et al., 2011].

An example of the F-PPP performance based on actual GPS data is shown in Fig. 6.5, which summarises the results obtained for users MLVL, EUSK and EIJS (at 252, 170 and 94 km, respectively, from the nearest reference receiver BRUS). In this example, the full user state was reset every 2 h to better characterise the convergence process, and the position

[24] Unlike in the differential positioning regarding a reference station (which requires, at least, common view satellites).

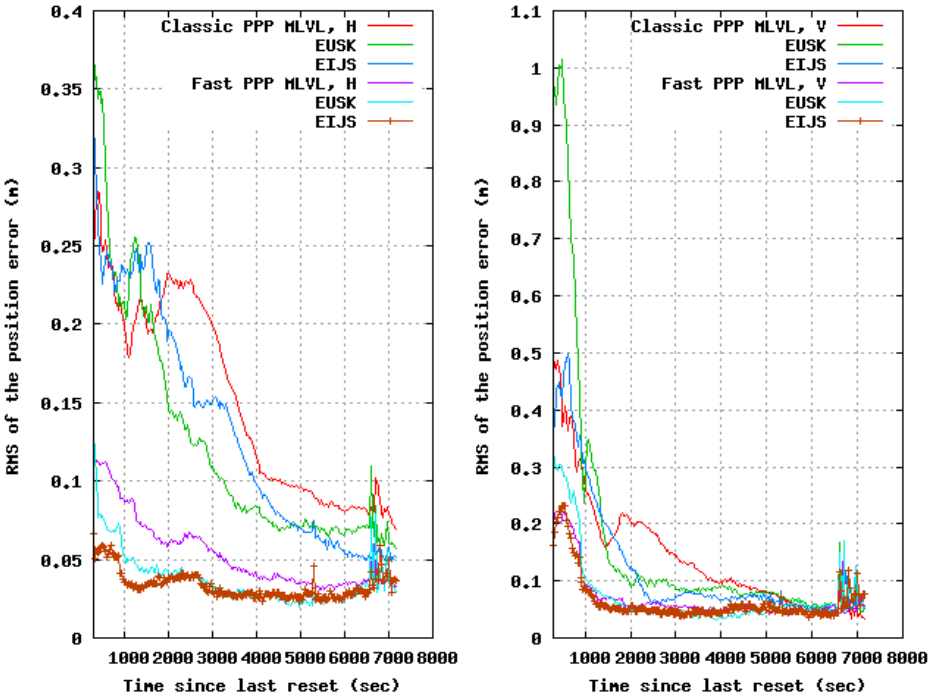


Figure 6.5: RMS of the positioning error for the horizontal component (*left*) and vertical component (*right*). The classical PPP for the rovers MLVL (red), EUSK (green) and EIJS (blues) is compared with the fast PPP for MLVL (violet), EUSK (light blue) and EIJS (brown). Data set of 24 November 2009.

error RMS of the resulting 12 time windows is depicted. The horizontal and vertical errors are shown (left- and right-hand side, respectively) demonstrating the advantage of using precise, real-time ionospheric corrections to speed up the PPP convergence (Fast PPP). It can be seen that the convergence time (to achieve, for instance, a 10 cm error level) is reduced from about 1 h (without ionospheric corrections) to a few minutes. For details see [Juan et al., 2012a].

It must be pointed out that the accurate ionospheric corrections accelerate the filter convergence even without fixing the carrier ambiguities, because equation (6.74) does not require the wide-lane ambiguity to be fixed. Only a sufficiently accurate wide-lane ambiguity estimate is needed to improve the estimate for B_C, which can always be achieved quickly.

This technique has been developed by the Research group of Astronomy and Geomatics (gAGE) of the Technical University of Catalonia (UPC) to provide a reliable High-Precision Positioning Service (HPPS), with a low-bandwidth requirement. Indeed, only 10% of additional bandwidth is needed for the classical PPP, because the additional data to be broadcast (i.e. ionospheric corrections and code and carrier instrumental delays) are slowly varying parameters that can be transmitted at a low rate.

This system approach, protected by several ESA-funded patents,[25] is able to provide an HPPS at a global scale, with worldwide undifferenced ambiguity fixing capability. Wide-area enhancement to allow this high accuracy to be achieved *rapidly* over continental areas is done with the help of accurate ionospheric corrections computed from a sparse reference station network (e.g. EGNOS reference stations (RIMS) in Europe). More details, including integrity issues, can be found in [Juan et al., 2012a].

[25] See the patent references and ESA- and GSA-funded projects to develop this technique at *http://www.gage.upc.edu*.

A. Time and Coordinate Frame Transformations

A brief description of time and coordinate systems was presented in section 3.1. This appendix contains an extended explanation and the necessary equations to perform accurate transformations between the different time and coordinate systems.

To simplify the presentation, the equations given in this appendix are based on the IERS Conventions of 1996 (IERS96) [McCarthy, 1996], although continuous references to IERS03 [McCarthy and Petit, 2004] are given when readers need to be aware of the changes.

A.1 Time Systems

As commented in section 3.1, several types of time systems are defined, according to the different references or periodic process involved (see Table 3.1). The definition, properties and relations between such kinds of time systems are presented below.

A.1.1 Earth's Rotation Time

A time linked to Earth's rotation can be measured from the angle between a reference meridian (a local or the Greenwich one) and the meridian containing a celestial reference. As already explained in section 3.1, solar time takes the Sun as a reference and sidereal time takes the vernal equinox (the Aries point).

A.1.1.1 Solar Time

Our time keeping was initially based on the motion of the Sun, but the way that this time flows is affected by two main causes:

- The orbit of Earth is elliptic. Thus, according to Kepler's second law,[1] the orbital speed is not constant.
- Earth's axis of rotation is not perpendicular to the plane of Earth's orbit around the Sun,[2] hence the angular rate is not constant. It moves fastest at the end of December and slowest in mid-September.

[1] The radius vector from the Sun to Earth sweeps out equal areas in equal times.

[2] The Sun moves around the ecliptic not the equator. The ecliptic is the apparent circular path of the Sun on the celestial sphere during the course of a year. The ecliptic is inclined at an angle of about $23°26'$ with respect to the equator, see equation (A.25).

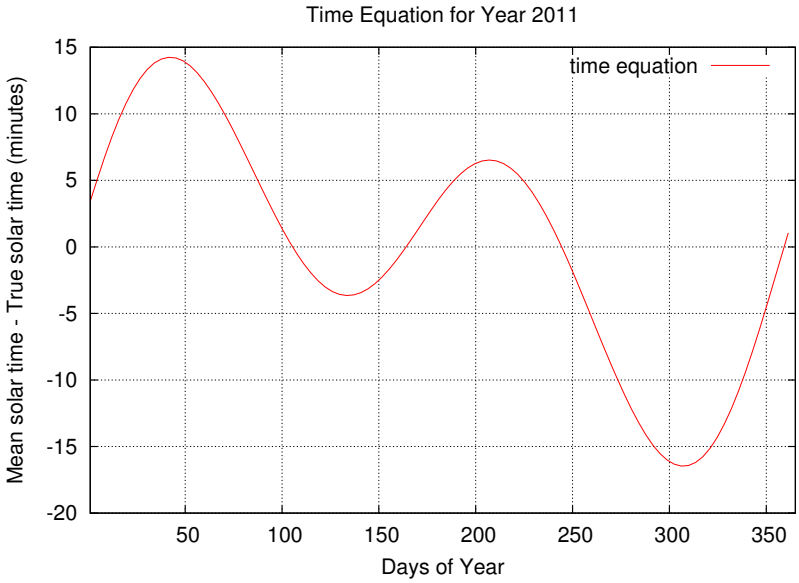


Figure A.1: Equation of time.

To get a more uniform time, a fictitious *mean Sun* is defined, which moves along the (celestial) equator of Earth with uniform speed (mean velocity of the actual Sun). Using this mean Sun, one defines *mean solar time* as the hour angle[3] of the centre of the mean Sun.

The difference between mean solar time and *true solar time* is reflected in the *equation of time*, which is a periodic function with a one-year period. Figure A.1 shows a plot of the equation of time. This evolution is a direct consequence of the two effects on the apparent solar rotation rate described above.

Civil time is defined from mean solar time as *mean solar time* augmented in 12 hours, so that each day begins at midnight.[4] Civil time is local (i.e. associated with local meridians). In order to get a global time that is not linked to these local meridians, Universal Time (UT) is defined as the civil time at the Greenwich meridian.

On the other hand, Earth's rotation rate is not uniform. It is affected by secular variations, mainly due to tidal friction, seasonal changes and other irregular or random effects, producing variations in Earth's distribution of mass and moment of inertia (see page 44). To deal which such effects, the following times (or refinements of UT) have been introduced:

UT0 is the mean solar time at the Greenwich meridian and is determined at a particular observatory by astronomical observations. As this time is based on Earth's instantaneous rotation, it is affected by both Earth's irregular spin rate and polar motion.

UT1 is obtained by correcting UT0 for the effect of polar motion on the location of the observing site (i.e. deducting the CTP pole, see equation (A.11)). UT1 is the same around the world (i.e. it does not depend

[3] *Sun hour angle* is the angular displacement of the Sun (measured over the equator and in a retrograde sense) from the local meridian. Thus, it is zero when the Sun is on the local meridian (solar noon), negative before local noon and positive in the afternoon.

[4] Midnight is the time when the hour angle of the mean Sun's centre is at a longitude of 180°.

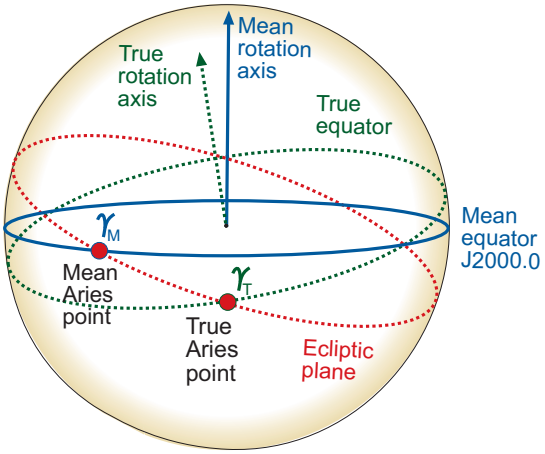


Figure A.2: Mean and true equinox.

on the observatory's location). This time is fundamental in geodetic astronomy, because it defines the orientation of the conventional terrestrial system in space (see section A.2.2).

UT2 is obtained by removing periodic seasonal variations from UT1 time, see equation (A.13), but it is not uniform enough due to the other effects on Earth's rotation that still remain. Nowadays it is considered obsolete.

A.1.1.2 Sidereal Time

The reference is the vernal equinox, or the *Aries point*, which is defined as the intersection of the equator with the ecliptic plane. Two types of Aries point can be considered depending on which equator plane is considered (i.e. the mean equator or the true equator):

- **Mean Aries point (γ_M)**: This is the intersection of the *mean equator* plane at epoch J2000.0 (the effect of precession on Earth's axis of rotation is taken into account, see section 3.1.2.1) with the ecliptic,[5] see Fig. A.2.
- **True Aries point (γ_T)**: This is defined as the intersection of the *true equator*[6] (the effects of precession and nutation of Earth's pole of rotation have been taken into account, see section 3.1.2.1) with the ecliptic,[7] see Fig. A.2.

[5] Actually, the mean ecliptic. Like Earth's pole of rotation, the ecliptic pole undergoes precession and nutation effects due to the perturbation of the Moon and major planets on Earth's orbit. Nevertheless, the amplitude of this pole is 50 times shorter than Earth's pole of rotation and, at the level of accuracy required here, we will not distinguish between the mean or true ecliptic.

[6] This equator is defined as the plane that contains the geocentre and is orthogonal to the instantaneous daily rotation axis.

[7] Due to the accuracies required, it is sufficient to compute the true Aries point using the mean ecliptic plane. We will always refer to this plane as 'the ecliptic', without distinguishing between the mean or the true one.

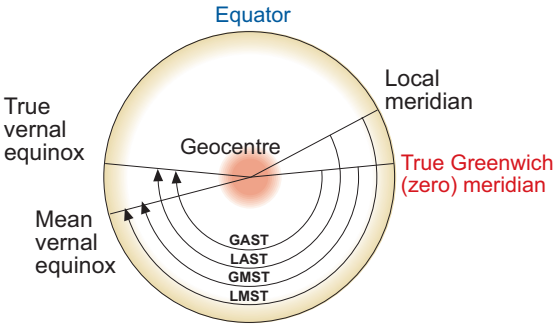


Figure A.3: Different sidereal times.

According to the previous definitions, four classes of sidereal times can be introduced:

- **Greenwich Apparent Sidereal Time (GAST)** (Θ_G) is the hour angle of the 'true' Aries point, from the true Greenwich meridian.
- **Local Apparent Sidereal Time (LAST)** (Θ) is the hour angle of the 'true' Aries point, from the local meridian.[8]
- **Greenwich Mean Sidereal Time (GMST)** (θ_G) is the same as GAST but with the mean equinox.
- **Local Mean Sidereal Time (LMST)** (θ) is the same as LAST but with the mean equinox.

Figure A.3 summarises these four sidereal times. GAST and GMST are given by equations (A.33) and (A.34), respectively. The local and Greenwich sidereal times differ by the longitude λ of the local meridian. The difference between the apparent and mean sidereal times is called the *equation of equinoxes* (where α_E is given by equation (A.37)), that is

$$\begin{aligned} \text{GMST} - \text{LMST} &= \text{GAST} - \text{LAST} = \lambda \\ \text{GMST} - \text{GAST} &= \text{LMST} - \text{LAST} = \alpha_E \end{aligned} \tag{A.1}$$

A.1.1.3 The Relationship between Solar and Sidereal Time

There is a slight difference between a *sidereal day* and a *solar* (or *synodic*) day due to the relative movement between the Sun and Earth as a consequence of its annual translation. Figure A.4 illustrates the concept.

As shown in this figure, the Aries reference point (γ), as seen from Earth, apparently moves clockwise during the course of one year (with respect to the Sun as reference). After one year, the directions of the Sun and Aries coincide again, but the number of laps relative to the Sun (solar days) is one less than those relative to Aries (sidereal days).

[8] *Apparent* in astronomy refers to what is seen from an ideal centre of Earth, without the atmosphere and rotation. Under these conditions the effects of the refraction of light, aberration and parallax are suppressed.

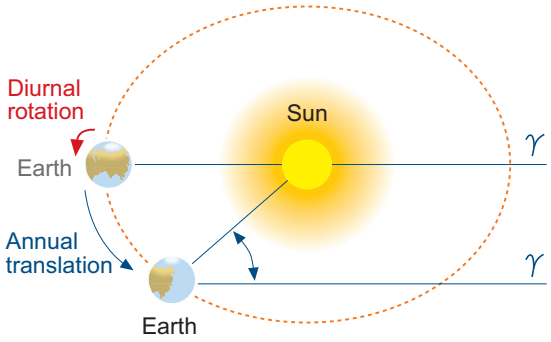


Figure A.4: Solar and sidereal times.

The difference between the *mean sidereal day* and *mean solar day* can be derived approximately as follows. As in a solar tropical year,[9] the mean solar time is delayed by one day with respect to the mean sidereal time, hence

$$\frac{24^{\mathrm{h}}}{365.2422} \simeq 3^{\mathrm{m}}56^{\mathrm{s}} \tag{A.2}$$

Thus, a sidereal day is shorter than a solar day for about $3^{\mathrm{m}}56^{\mathrm{s}}$.

The conversion between UT1 and apparent and mean sidereal time is given by equations (A.33) and (A.35) in section A.2.5.2.

A.1.2 Earth's Revolution Time

A.1.2.1 Ephemeris Time

Ephemeris Time (ET) was adopted in 1952 as the (conceptually uniform) time of Newtonian mechanics. It was defined in terms of the Earth's orbital revolution around the Sun, and it was declared to be free of the effects of irregularity in the unpredictable polar motion and Earth's rate of rotation.

The ET second was established as the SI second between 1956 and 1967 and defined as the tropical year fraction

$$\text{ET second} = \frac{1}{31\,556\,925.9747} \tag{A.3}$$

In 1976, in Grenoble, the IAU resolved that ET would be replaced by two relativistic time scales intended to constitute dynamical time scales (derived from planetary motions in the Solar System): Temps Dinamique Baricentrique (TDB) and Temps Dinamique Terrestre (TDT):

- **TDB** is an inertial time in the Newtonian sense and provides the time variable in the equations of motion for the ephemerides related to the centre of gravity of the Solar System (i.e. for computing planetary orbits around the Sun).

[9] A tropical year is the elapsed time between two successive culminations of the Sun by the mean equinox. It has a duration of 365.2422 mean solar days.

- **TDT** is a quasi-inertial time in the Newtonian sense and is the independent variable of the satellite's equations of motion around Earth.[10]

In the general terminology of gravity, TDB is a *coordinate time* and TDT is a *proper time*.

IAU Resolution A4 (1991) established the framework presently used to define the Barycentric Reference System (BRS) and Geocentric Reference System (GRS), respectively, and associated with these systems were the defined Barycentric Coordinate Time, from the French Temps Coordonné Barycentrique (TCB) and Geocentric Coordinate Time, from the French Temps Coordonné Géocentrique (TCG). An additional time for the GRS system named Terrestrial Time (TT) was also defined. TT corresponds to TDT in the former definitions (TDT and TDB used before 1991) and is introduced as the theoretical counterpart of the realised time scale TAI + 32.184 s (see the IERS Conventions of 2003 [McCarthy and Petit, 2004]).

A.1.3 Atomic Time

A.1.3.1 TAI

TAI stands for 'International Atomic Time' and was established as a reference time by the Bureau International de l'Heure. Its initial epoch was matched to the $0^h0^m0^s$ of the UT2 scale of 1 January 1958, so the difference between TAI and UT2 was zero in this epoch.

The duration of the TAI second was defined in 1967 as the same as the ET second, according to the following extract:

> *The TAI second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the Caesium 133 atom.*

It is realised by several high-precision atomic clocks held at standards institutes in various countries; it is, therefore, a statistical time. There is an elaborate process of continuous intercomparison, leading to a weighted average of all the clocks involved.

Following this definition, the TAI second was substituted for the ET second as the SI second in 1967. Bearing in mind the next three points, one can obtain the relationship between ET and TAI:

- The TAI second equals the ET second.
- The difference between ET time and UT2 time at $0^h0^m0^s$ of the UT2 scale of 1 January 1958 was 32.184 s.
- TAI time equals UT2 time at $0^h0^m0^s$ of the UT2 scale of 1 January 1958.

Obviously the relationship between the TAI scale and ET scale (or TDT scale subsequently) is a constant offset of 32.184 s.

[10] It is quasi-inertial because Earth is subjected to an acceleration caused by its annual translation around the Sun.

A.1.3.2 UTC

UTC stands for Coordinated Universal Time and is a compromise between TAI and UT1. In fact UTC, as an atomic time, is as uniform as the TAI scale is, but it is always kept closer than 0.9 s with respect to UT1, in order to follow the variations in Earth's rotation. This is accomplished by adding (or subtracting) a certain number of *leap seconds* to TAI. This number, which is refreshed periodically, is provided by the IERS.

This time is very important since most of the currently used time signals are synchronised with UTC.

Official time is the one used by all nations in the world. It usually differs by an integer number of hours or half an hour with regard to UTC. This difference is given by time zones and the proper adjustments in summer and winter.

UTC is generated at the Bureau International des Poids et Mesures (BIPM), located near Paris (*http://www.bipm.fr*). It is obtained on the basis of times kept by about 250 caesium clocks and hydrogen masers located at some 65 different laboratories, distributed around the world, and applying a set of algorithms to ensure a uniform time. It is not determined in real time, but generated with a delay of about half a month.

Real-time estimates of UTC are computed and provided by different centres, such as: UTC(USNO), from the United States Naval Observatory (USNO); UTC(NIST), from the National Institute of Standards and Technology (NIST); and UTC(SU) from Russia (SU stands for the former Soviet Union); or, in general, UTC(k) as a realisation of UTC by a given laboratory k. For more details see [Lewandowski et al., 2006].

GNSS satellites broadcast in their navigation messages the necessary parameters to compute real-time estimates of UTC. Indeed, these parameters allow a GPS receiver clock to calculate accurate estimates of UTC(USNO) at the level of 25 ns RMS, and remote clocks can be compared with an accuracy of 5 ns. On the other hand, one of the requirements of Glonass updates is to keep the UTC − UTC(SU) difference within 10 ns, as an accurate means of UTC(SU) dissemination [Januszewski, 2010].

A.1.4 Julian Date

In order to facilitate calculations for long time intervals, the Julian date (JD) is used (after Julio Scalier).

The JD defines the number of days[11] elapsed since *4713* BC, *January* $1^d.5$ (i.e. 12^h of 1 January). This day is the zero epoch for the JD and, after that time, the days are counted without interruption. Every Julian day begins at 12 hours UT of the civil day and ends at 12 hours UT of the following day. This kind of calendar does not use the hour:minute:second format, but fractions of the day.

[11] The Julian century is defined as 36 525 days.

The conversion between calendar year(Y):month(M):day(D) of UT and JD is given by the following expressions [Hofmann-Wellenhof et al., 2008]:

$$\mathrm{JD} = \mathrm{int}[365.25\ y] + \mathrm{int}[30.6001\ (m+1)] + D + \frac{\mathrm{UT\_hours}}{24.0} + 1\,720\,981.5 \tag{A.4}$$

where

$$\begin{array}{lll} y = Y - 1 & \text{and } m = M + 12; & M \leq 2 \\ y = Y & \text{and } m = M; & M > 2 \end{array} \tag{A.5}$$

The JD relative to J2000 is obtained by subtracting 2 451 545.0 days from the JD (i.e. the number of days elapsed since 12^h on 1 January 2000).

The inverse transformation is carried out by computing the next values:

$$\begin{aligned} a &= \mathrm{int}[\mathrm{JD} + 0.5] \\ b &= a + 1537 \\ c &= \mathrm{int}\left[\frac{b - 122.1}{365.25}\right] \\ d &= \mathrm{int}[365.25\ c] \\ e &= \mathrm{int}\left[\frac{b-d}{30.6001}\right] \end{aligned} \tag{A.6}$$

Afterwards the civil date is obtained from the following expressions:

$$\begin{aligned} D &= b - d - \mathrm{int}[30.6001\ e] + \mathrm{frac}[\mathrm{JD} + 0.5] \\ M &= e - 1 - 12 \times \mathrm{int}\left[\frac{e}{14}\right] \\ Y &= c - 4715 - \mathrm{int}\left[\frac{7+M}{10}\right] \end{aligned} \tag{A.7}$$

where int and frac denote the integer and fractional part of a real number, respectively.

A.1.5 Transformations

A summary of transformations between different times is presented next.

A.1.5.1 TAI – UTC

The difference between these times is given by the accumulated leap seconds. These leap seconds are provided by the IERS. The IGS centres provide files of Earth's rotation parameters and time information, which are actualised periodically (see *http://acc.igs.org*):

$$\mathrm{UTC} = \mathrm{TAI} - \text{leap seconds} \tag{A.8}$$

With the introduction of the leap second, UTC (which is an atomic time) does not differ from Earth's rotational time (UT1) by more than 0.9 s.

A.1.5.2 TAI – UT1

An a priori UT1 ($\widetilde{\text{UT1}}$) can be computed from

$$\widetilde{\text{UT1}} = \text{UT1R} + \Delta\text{UT1} \tag{A.9}$$

hence

$$\text{TAI} - \widetilde{\text{UT1}} = [\text{TAI} - \text{UT1R}] - \Delta\text{UT1}$$

where UT1R is a smoothed a priori measurement of Earth's orientation for which the short-period ($t < 35$ days) tidal effects have been removed before smoothing [Webb and Zumberge, 1993].

TAI−UT1R or UTC−UT1R are provided in files for Earth's rotation parameter.[12]

According to [Yoder et al., 1981], the periodic variations of UT1 due to tidal deformations of the polar moment of inertia are given by

$$\Delta\text{UT1} = \sum_{i=1}^{N} \left[A_i \ \sin \left(\sum_{j=1}^{5} k_{ij} \ \alpha_j \right) \right] \tag{A.10}$$

The number N is chosen to include all terms with a period of less than 35 days. The α_j are functions of time given by equations (A.26) to (A.30) below. The parameters A_i, k_{ij} are given in Table A.1 for periods up to 41 days. A table with periods up to 18.6 years (62 terms) based on [Yoder et al., 1981] is available at *http://hpiers.obspm.fr/eop-pc/models/UT1/UT1R_tab.html*. An extended model with sine and cosine terms is given in the IERS03 Conventions (see [McCarthy and Petit, 2004], page 92) with the associated table of coefficients.

A.1.5.3 UT0, UT1, UT2

As commented above, UT0, UT1 and UT2 are universal times linked to Earth's rotation.

UT1 and UT0 are related by the following expression, where Δl is a correction of longitude due to the effect of polar motion:

$$\text{UT1} = \text{UT0} + \Delta l \tag{A.11}$$

where

$$\Delta l = \frac{1^{\text{s}}}{15} \ (x_p \ \sin\lambda - y_p \ \cos\lambda) \ \tan\varphi \tag{A.12}$$

In this expression x_p and y_p are the instantaneous pole coordinates (see Fig. A.7 below) while λ and φ are the site's latitude and longitude respectively.[13] With this rotation axis correction the pole movement in longitude

[12] Earth's orientation parameters and time information are provided from different servers and with different format files: namely, 'erp' files, the GIPSY format 'PNML' and 'tpeo.nml' files. See *ftp://igscb.jpl.nasa.gov/pub/product/iers/* or *http://acc.igs.org/* or the JPL sites *ftp://sideshow.jpl.nasa.gov/pub/gipsy_products/2009/orbits/* and *ftp://sideshow.jpl.nasa.gov/pub/gipsy_products/RapidService/orbits/*.

[13] The x_p and y_p pole displacements are given in the previously mentioned files for Earth's rotation parameters (e.g. the JPL TPNML files).

(Δl) is removed and one passes from the instantaneous pole (CEP) to the Conventional Terrestrial Pole (CTP), see Fig. 3.3, on the right.

UT2 is obtained by correcting UT1 for the periodic seasonal variation of Earth's rotation

$$\mathrm{UT2} = \mathrm{UT1} + \Delta S \tag{A.13}$$

where

$$\Delta S = 0^{\mathrm{s}}.022 \sin 2\pi t - 0^{\mathrm{s}}.120 \cos 2\pi t - 0^{\mathrm{s}}.0060 \sin 4\pi t + 0^{\mathrm{s}}.0070 \cos 4\pi t \tag{A.14}$$

with ΔS in seconds and t being the date in Besselian years:

$$t = 2000.00 + \frac{\mathrm{MJD} - 51\,544.03}{365.2422} \tag{A.15}$$

A.1.5.4 TAI – TDT, TCG, TT

As mentioned previously, following the IAU Resolution of 1991, the TDT is substituted by TCG and TT, where TT corresponds to the TDT in the former definition (before 1991). Thus

$$\mathrm{TDT} \equiv \mathrm{TT} = \mathrm{TAI} + 32.184\,\mathrm{s} \tag{A.16}$$

where TAI is a realisation of TT, apart from a constant offset of 32.184 s. On the other hand, TCG and TT are related by [McCarthy and Petit, 2004]

$$\mathrm{TCG} - \mathrm{TT} = L_G\,(\mathrm{MJD} - 43\,144.0)\,86\,400\,\mathrm{s} \tag{A.17}$$

with $L_G = 6.969\,290\,134 \cdot 10^{-10}$.

A.1.5.5 TDT – TDB, TCB

The relationship between these times is given by

$$\mathrm{TDB} = \mathrm{TDT} + 0^{\mathrm{s}}.001\,658 \sin(g + 0.0167 \sin g) \tag{A.18}$$

with

$$g = \frac{2\pi(357^\circ.528 + 35\,999^\circ.050\,T)}{360^\circ}$$

where T is in centuries from J2000 of TDT:[14]

$$\mathrm{Centuries} = \frac{\mathrm{JD} - 2\,451\,545.0}{36\,525} \tag{A.19}$$

As in the case of TDT, after the IAU Resolution of 1991, TDB was substituted by TCB, which is related to TDB by [McCarthy and Petit, 2004]

$$\mathrm{TCB} - \mathrm{TDB} = L_B(\mathrm{MJD} - 43\,144.0)\,86\,400\,\mathrm{s} + P_0 \tag{A.20}$$

with $L_B = 1.550\,519\,767\,72 \cdot 10^{-8}$ and $P_0 \simeq 6.55 \cdot 10^{-5}$ s.

[14] T can be taken here as centuries from J2000 of TDT, without incurring significant error.

A.2 Coordinate Systems

Although the CRS and TRS were defined in section 3.1.2, a brief review is given here to help clarify the concepts introduced in this appendix.

A.2.1 Conventional Celestial Reference System (CRS)

This reference system is also known as Earth-Centred Inertial (ECI)[15] and is mainly used for describing satellite motion. CRS has its origin in Earth's centre of mass or geocentre, its fundamental plane is the mean equatorial plane (containing the geocentre) of the epoch J2000.0, and the principal axis x points to the mean vernal equinox of epoch J2000.0 mean Aries point.

The three axes defining this coordinate are shown in Fig. A.5 and are as follows:

- x_{CRS} axis: Its origin is the geocentre, Earth's centre of mass, and its direction is towards the mean equinox at J2000.0 (i.e. the intersection of the J2000.0 mean equatorial plane with the ecliptic plane).
- z_{CRS} axis: This is defined by the direction of Earth's mean rotation pole at J2000.0 (orthogonal to the mean equatorial plane at J2000.0 epoch).
- y_{CRS} axis: This is orthogonal to the other axes, so the system is right handed.

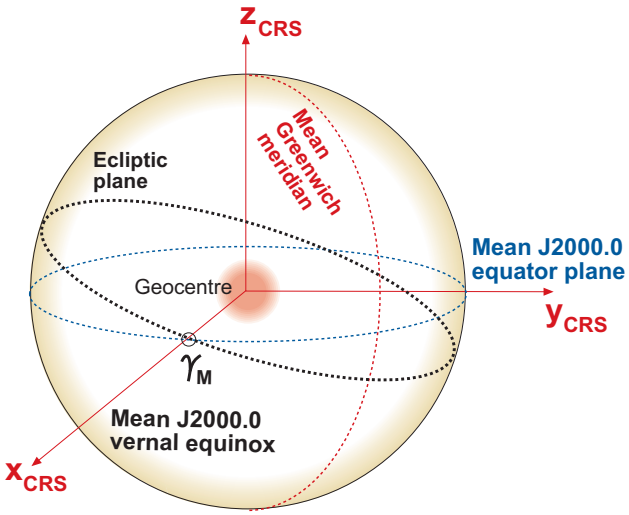


Figure A.5: Conventional Celestial Reference System.

[15] Strictly speaking this is a quasi-inertial system because of the annual motion of Earth around the Sun. Thus it is subjected to a certain acceleration but can be thought as inertial over short periods of time.

A.2.2 Conventional Terrestrial Reference System (TRS)

This reference system is also known as Earth-Centred, Earth-Fixed (ECEF). It is a rotating reference system (not a space-fixed system like CRS). Its origin is Earth's centre of mass; the fundamental plane contains this origin and it is perpendicular to Earth's CTP – defined as the average of the poles from 1900 to 1905. Its principal axis points to the intersection of the mean Greenwich meridian with the equator. Since this coordinate system follows the diurnal rotation of Earth, this is not an inertial reference system.

The three axes that define this system are shown in Fig. A.6 and are as follows:

- z_{TRS} axis: This is defined by the CTP.
- x_{TRS} axis: This is defined as the intersection of the equatorial plane with the mean Greenwich meridian. The equatorial plane is orthogonal to the CTP. The mean Greenwich meridian was established by the Bureau International de l'Heure (BIE) Observatory.
- y_{TRS} axis: This is orthogonal to the other axes, so the system is right handed.

A.2.3 Celestial Ephemeris Pole (CEP)

The CEP is Earth's instantaneous rotation axis. This pole allows us to define a quasi-inertial system, like the CRS, but linked with this true pole. Its origin is Earth's centre of mass and its fundamental plane is the true equator, as it is orthogonal to the true pole; and its principal direction is towards the true vernal equinox (i.e. the intersection between the ecliptic plane and the true equatorial plane), see Fig. 3.3.

If all external forces were null, the CEP would be a fixed pole in space (i.e. it would not precess or nutate). Nevertheless the position of the CEP relative to Earth's crust would change due to the irregular mass distribution of Earth and its variation. But if Earth were spherical and homogeneous,

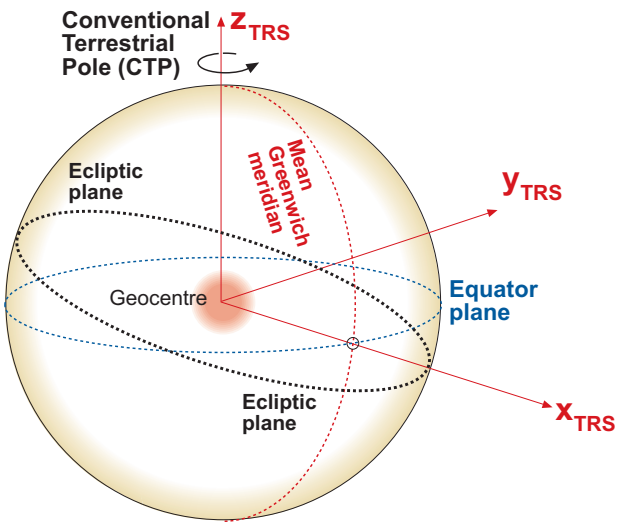


Figure A.6: Conventional Terrestrial Reference System.

the CEP would be a fixed pole in space, whether the external forces are null of not, and also on the surface of Earth. This pole is a suitable reference for an external observer.

IAU Resolution B1.7 recommended that, after 1 January 2003, the CEP should be replaced by the Celestial Intermediate Pole (CIP). *The definition of the CIP is an extension of that of the CEP in the high-frequency domain and coincides with that of the CEP in the low-frequency domain* [McCarthy and Petit, 2004]. It is consistent with the IAU2000A model for precession and nutation.

A.2.4 Reference Systems and Frames

Reference *system* and reference *frame* are different concepts. The first is understood as 'a theoretical definition', including models and standards for its implementation. The second is a 'practical implementation' through observations and a set of reference coordinates, for example a set of fundamental stars, for a celestial reference frame, or fiducial stations for a terrestrial reference frame.

The International Celestial Reference System (ICRS) was proposed by the IERS and formally accepted by the IAU in 1997. A realisation of the ICRS is the International Celestial Reference Frame (ICRF). On the other hand, IERS is in charge of defining, realising and promoting the International Terrestrial Reference System (ITRS). Realizations of ITRS are the ITRFs,[16] with ITRF2008 being the current reference realisation of ITRS at the time of writing.

A.2.5 Transformations between Celestial and Terrestrial Frames

The procedure for performing an accurate transformation between the coordinate frames of the ICRF and ITRF is described in the following sections (see Fig. 3.3).

To simplify the explanations a little, the equations and parameters provided in this appendix implement the IAU 1976 and 1980 precession and nutation models, respectively (see the IERS Conventions 1996 [McCarthy, 1996]). Nevertheless, IAU Resolution B16 recommended that, commencing 1 January 2003, the previous models should be replaced by the new IAU 2000 model. Procedures and equations implementing such transformations are described in the IERS Conventions 2003 [McCarthy and Petit, 2004]. The IERS documents and software routines can be found at *ftp://maia.usno.navy.mil/conv2003/chapter5/*.

[16] See *http://itrf.ensg.ign.fr/ITRF_solutions/index.php*.

A.2.5.1 ICRF to CEP

As indicated above, to transform coordinates from the ICRF to the CEP it is necessary to correct the precession and nutation effects. This transformation is performed by a composite of rotations that can done by two orthogonal matrices $\mathbf{P}$ and $\mathbf{N}$ as follows:

$$\mathbf{r}_{\mathrm{CEP}} = \mathbf{N}\,\mathbf{P}\,\mathbf{r}_{\mathrm{ICRF}} \tag{A.21}$$

Precession Matrix $\mathbf{P}$

Mean coordinates are transformed from the reference epoch J2000.0 to the observation epoch t using the precession matrix

$$\mathbf{P} = \mathbf{R}_3(-z)\ \mathbf{R}_2(\vartheta)\ \mathbf{R}_3(-\zeta)$$

$$= \begin{bmatrix} \cos(-z) & \sin(-z) & 0 \\ -\sin(-z) & \cos(-z) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\vartheta & 0 & -\sin\vartheta \\ 0 & 1 & 0 \\ \sin\vartheta & 0 & \cos\vartheta \end{bmatrix} \begin{bmatrix} \cos(-\zeta) & \sin(-\zeta) & 0 \\ -\sin(-\zeta) & \cos(-\zeta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos z\ \cos\vartheta\ \cos\zeta & -\cos z\ \cos\vartheta\ \sin\zeta & -\cos z\ \sin\vartheta \\ -\sin z\ \sin\zeta & -\sin z\ \cos\zeta & \\ \sin z\ \cos\vartheta\ \cos\zeta & -\sin z\ \cos\vartheta\ \sin\zeta & -\sin z\sin\vartheta \\ +\cos z\ \sin\zeta & +\cos z\ \cos\zeta & \\ \sin\vartheta\ \cos\zeta & -\sin\vartheta\ \sin\zeta & \cos\vartheta \end{bmatrix} \tag{A.22}$$

where the angles are given by the expressions (IAU 1976 Precession Model)[17]

$$\begin{aligned} z &= 2306\overset{''}{.}2181\,T + 1\overset{''}{.}094\,68\,T^2 + 0\overset{''}{.}018\,203\,T^3 \\ \vartheta &= 2004\overset{''}{.}3109\,T - 0\overset{''}{.}426\,65\,T^2 - 0\overset{''}{.}041\,833\,T^3 \\ \zeta &= 2306\overset{''}{.}2181\,T + 0\overset{''}{.}301\,88\,T^2 + 0\overset{''}{.}017\,998\,T^3 \end{aligned} \tag{A.23}$$

with T the time expressed in Julian centuries of TDB (see equation (A.19)) between the reference epoch J2000.0 and the epoch t of observation. Note that the time t is expressed in the TDB scale because, as in previous sections, all computations related to Solar System planetary orbits and mechanics must be given in this scale.

[17] IAU Resolution B16 recommended that, beginning on 1 January 2003, the IAU 1976 Precession Model and IAU 1980 Theory of Nutation should be replaced by the IAU 2000 models. Equations, parameters, procedures and software routines implementing the IAU 2000 model can be found at *ftp://maia.usno.navy.mil/conv2003/chapter5/*. (IERS Recommendations 2003).

Nutation Matrix $\mathbf{N}$

Mean coordinates at the observation epoch t are transformed to true coordinates referred to the instantaneous true equator and equinox using the nutation matrix:

$$\mathbf{N} = \mathbf{R}_1\left(-(\varepsilon')\right)\,\mathbf{R}_3\left(-\Delta\psi\right)\,\mathbf{R}_1\left(\varepsilon\right)$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(-\varepsilon') & \sin(-\varepsilon') \\ 0 & -\sin(-\varepsilon') & \cos(-\varepsilon') \end{bmatrix} \begin{bmatrix} \cos(-\Delta\psi) & \sin(-\Delta\psi) & 0 \\ -\sin(-\Delta\psi) & \cos(-\Delta\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \times$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varepsilon & \sin\varepsilon \\ 0 & -\sin\varepsilon & \cos\varepsilon \end{bmatrix}$$

$$= \begin{bmatrix} \cos\Delta\psi & -\cos\varepsilon\sin\Delta\psi & -\sin\varepsilon\sin\Delta\psi \\ -\cos\varepsilon'\sin\Delta\psi & \cos\varepsilon'\cos\varepsilon\cos\Delta\psi \\ & -\sin\varepsilon'\sin\varepsilon \\ \sin\varepsilon'\sin\Delta\psi & -\sin\varepsilon'\cos\varepsilon\cos\Delta\psi & -\sin\varepsilon'\sin\varepsilon\cos\Delta\psi \\ & +\cos\varepsilon'\sin\varepsilon & +\cos\varepsilon'\cos\varepsilon \end{bmatrix} \tag{A.24}$$

where the angles are given by the expressions (1980 IAU Nutation model)

$$\begin{aligned} \varepsilon' &= \varepsilon + \Delta\varepsilon \\ \varepsilon &= 23^\circ 26' 21\overset{''}{.}448 - 46\overset{''}{.}8150\,T - 0\overset{''}{.}000\,59\,T^2 + 0\overset{''}{.}001\,813\,T^3 \\ \Delta\psi &= \textstyle\sum_{j=1}^{N}\left[(A_{0j} + A_{1j}\,T)\,\sin\left(\sum_{i=1}^{5} k_{ji}\,\alpha_i\right)\right] \\ \Delta\varepsilon &= \textstyle\sum_{j=1}^{N}\left[(B_{0j} + B_{1j}\,T)\,\cos\left(\sum_{i=1}^{5} k_{ji}\,\alpha_i\right)\right] \end{aligned} \tag{A.25}$$

the parameters A_{0j}, A_{1j}, B_{0j}, B_{1j} and k_{ji} being listed in Table A.2. The parameters α_i are time dependent (just like ε) and are computed as

- Moon's mean anomaly

$$\alpha_1 = l = 485\,866\overset{''}{.}733 + (1325^r + 715\,922\overset{''}{.}633)T + 31\overset{''}{.}310T^2 + 0\overset{''}{.}064T^3 \tag{A.26}$$

- Sun's mean anomaly

$$\alpha_2 = l' = 1\,287\,099\overset{''}{.}804 + (99^r + 1\,292\,581\overset{''}{.}224)T - 0\overset{''}{.}577T^2 - 0\overset{''}{.}012T^3 \tag{A.27}$$

- Moon's mean argument of latitude

$$\alpha_3 = F = 335\,778\overset{''}{.}877 + (1342^r + 295\,263\overset{''}{.}137)T - 13\overset{''}{.}257T^2 + 0\overset{''}{.}011T^3 \tag{A.28}$$

- Moon's mean elongation from the Sun

$$\alpha_4 = D = 1\,072\,261.''307 + (1236^r + 1\,105\,601.''328)T - 6.''891T^2 + 0.''019T^3 \tag{A.29}$$

- Mean longitude of the ascending lunar node

$$\alpha_5 = \Omega = 450\,160.''280 - (5^r + 482\,890.''539)T + 7.''455T^2 + 0.''008T^3 \tag{A.30}$$

The value $1^r = 360° = 1\,296\,000''$ and T is the same as the one used to compute the precession matrix.

A.2.5.2 CEP to ITRF

To transform the coordinates from the instantaneous space-fixed equatorial system CEP to a conventional terrestrial frame ITRF it is necessary to correct the polar motion and sidereal time (because the CEP is a non-rotating system while ITRF is tied to Earth's rotation). This transformation is given by the composite of rotation matrices $\mathbf{R}_S$ and $\mathbf{R}_M$ as follows:

$$\mathbf{r}_{\text{ITRF}} = \mathbf{R}_M\,\mathbf{R}_S\,\mathbf{r}_{\text{CEP}} \tag{A.31}$$

Earth Rotation Matrix $\mathbf{R}_S$

This matrix defines a rotation around the CEP of angle Θ_G. That is,

$$\mathbf{R}_S = \mathbf{R}_3(\Theta_G) = \begin{bmatrix} \cos\Theta_G & \sin\Theta_G & 0 \\ -\sin\Theta_G & \cos\Theta_G & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{A.32}$$

where Θ_G is the Greenwich Apparent (=true) Sidereal Time (GAST)[18] of the date, which is given by

$$\Theta_G = \theta_G + \alpha_E \tag{A.33}$$

where θ_G is the Greenwich Mean Sidereal Time[19] (GMST) and α_E is the *equinox equation* at the observation time UT1, and where[20]

$$\theta_G = 1.002\,737\,909\,350\,795 \cdot \text{UT1} + \theta_{G_0} \tag{A.34}$$

with θ_{G_0} the GMST at 0^h UT1:

$$\theta_{G_0} = 6^{\text{h}}41^{\text{m}}50.^{\text{s}}548\,41 + 864\,018\,4.^{\text{s}}812\,866T_u + 0.^{\text{s}}093\,104T_u^2 - 6.^{\text{s}}2\,10^{-6}T_u^3 \tag{A.35}$$

and T_u expressed in centuries of Julian UT1 date:

$$T_u = \frac{(\text{Julian UT1 date}) - 2\,451\,545.0}{36\,525} \tag{A.36}$$

[18] That is, the hour angle of the true equinox, from the true Greenwich meridian (see section A.1.1.2).

[19] The hour angle of the mean equinox of the date, from the Greenwich meridian.

[20] The ratio of universal to sidereal time 1.002 737 909 350 795 can be approximately computed as the ratio between 24^h and $24^{\text{h}} - 3^{\text{m}}56^{\text{s}}$. A very accurate value is $1.002\,737\,909\,350\,795 + 5.9006 \cdot 10^{-11}T_u - 5.9 \cdot 10^{-15}T_u^2$.

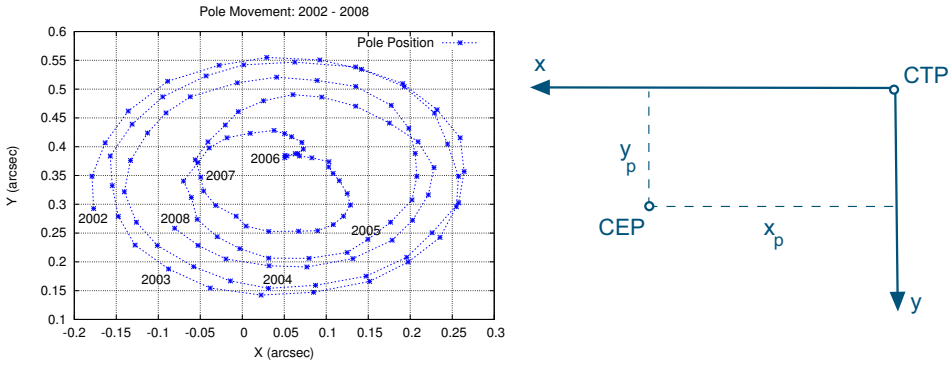


Figure A.7: Polar motion measurements in arcseconds from 2002 to 2008 (*left*) and polar coordinates (*right*).

α_E is the difference between the hour angle of the true equinox (γ_T) of the date and the mean equinox (γ_M) of the date,[21] a difference which is due to the nutation of Earth and can be computed from

$$\alpha_E = \tan^{-1}\left(\frac{\mathbf{N}_{12}}{\mathbf{N}_{11}}\right) \tag{A.37}$$

where $\mathbf{N}_{ij}$ is the component of row i and column j of the nutation matrix (see expression (A.24)). Note that the sidereal time (GAST or GMST) is always measured from the *true Greenwich meridian*.

Polar Motion Matrix $\mathbf{R}_M$

This matrix is given by two elemental rotations, $\mathbf{R}_1(-y_p)$ and $\mathbf{R}_2(-x_p)$, around the x- and y-axis, respectively. With these rotations the CEP is matched to the CTP pole:

$$\mathbf{R}_M = \mathbf{R}_2(-x_p)\,\mathbf{R}_1(-y_p) \tag{A.38}$$

where x_p and y_p define the position of the CEP with respect to the CTP (see Fig. A.7, on the right side).

The x_p and y_p pole displacements between 2002 and 2008 are shown in Fig. A.7, on the left side.

Because the polar movement is quite unpredictable, the x_p and y_p displacements are estimated from observations and provided by the IERS, as mentioned above. These values can be obtained from files for Earth's rotation parameters provided by the IERS or JPL.

[21] That is, between the GAST and GMST. This difference is called the *equation of equinoxes*, see equation (A.1).

Table A.1: Variations in Earth's rotation due to zonal tides [Yoder et al., 1981] (see IERS 96 Conventions). Components with period under 35 days. See also *http://hpiers.obspm.fr/eop-pc/models/UT1/UT1R_tab.html*.

Index	Argument coefficient					Period	A_i
i	k_{i1}	k_{i2}	k_{i3}	k_{i4}	k_{i5}	(days)	($\times 10^{-4}$ s)
1	1	0	2	2	2	5.64	−0.024
2	2	0	2	0	1	6.85	−0.040
3	2	0	2	0	2	6.86	−0.099
4	0	0	2	2	1	7.09	−0.051
5	0	0	2	2	2	7.10	−0.123
6	1	0	2	0	0	9.11	−0.039
7	1	0	2	0	1	9.12	−0.411
8	1	0	2	0	2	9.13	−0.993
9	3	0	0	0	0	9.18	−0.018
10	−1	0	2	2	1	9.54	−0.082
11	−1	0	2	2	2	9.56	−0.197
12	1	0	0	2	0	9.61	−0.076
13	2	0	2	−2	2	12.81	0.022
14	0	1	2	0	2	13.17	0.025
15	0	0	2	0	0	13.61	−0.299
16	0	0	2	0	1	13.63	−3.208
17	0	0	2	0	2	13.66	−7.757
18	2	0	0	0	−1	13.75	0.022
19	2	0	0	0	0	13.78	−0.338
20	2	0	0	0	1	13.81	0.018
21	0	−1	2	0	2	14.19	−0.024
22	0	0	0	2	−1	14.73	0.047
23	0	0	0	2	0	14.77	−0.734
24	0	0	0	2	1	14.80	−0.053
25	0	−1	0	2	0	15.39	−0.051
26	1	0	2	−2	1	23.86	0.050
27	1	0	2	−2	2	23.94	0.101
28	1	1	0	0	0	25.62	0.039
29	−1	0	2	0	0	26.88	0.047
30	−1	0	2	0	1	26.98	0.177
31	−1	0	2	0	2	27.09	0.435
32	1	0	0	0	−1	27.44	0.534
33	1	0	0	0	0	27.56	−8.261
34	1	0	0	0	1	27.67	0.544
35	0	0	0	1	0	29.53	0.047
36	1	−1	0	0	0	29.80	−0.055
37	−1	0	0	2	−1	31.66	0.118
38	−1	0	0	2	0	31.81	−1.824
39	−1	0	0	2	1	31.96	0.132
40	1	0	−2	2	−1	32.61	0.018
41	−1	−1	0	2	0	34.85	−0.086

Table A.2: IAU 1980 Theory of nutation model (source *http://hpiers.obspm.fr/eop-pc/models/nutations/nut_IAU1980.dat*).(.../contd.)

Index	Argument coefficient					Period	A_{0j}	A_{1j}	B_{0j}	B_{1j}
i	k_{i1}	k_{i2}	k_{i3}	k_{i4}	k_{i5}	(days)	($\times 10^{-4}$ ″)		($\times 10^{-4}$ ″)	
1	0	0	0	0	1	−6798.4	−171 996.0	−174.2	92 025.0	8.9
2	0	0	2	−2	2	182.6	−13 187.0	−1.6	5 736.0	−3.1
3	0	0	2	0	2	13.7	−2 274.0	−0.2	977.0	−0.5
4	0	0	0	0	2	−3399.2	2 062.0	0.2	−895.0	0.5
5	0	−1	0	0	0	−365.3	−1 426.0	3.4	54.0	−0.1
6	1	0	0	0	0	27.6	712.0	0.1	−7.0	0.0
7	0	1	2	−2	2	121.7	−517.0	1.2	224.0	−0.6
8	0	0	2	0	1	13.6	−386.0	−0.4	200.0	0.0
9	1	0	2	0	2	9.1	−301.0	0.0	129.0	−0.1
10	0	−1	2	−2	2	365.2	217.0	−0.5	−95.0	0.3
11	−1	0	0	2	0	31.8	158.0	0.0	−1.0	0.0
12	0	0	2	−2	1	177.8	129.0	0.1	−70.0	0.0
13	−1	0	2	0	2	27.1	123.0	0.0	−53.0	0.0
14	1	0	0	0	1	27.7	63.0	0.1	−33.0	0.0
15	0	0	0	2	0	14.8	63.0	0.0	−2.0	0.0
16	−1	0	2	2	2	9.6	−59.0	0.0	26.0	0.0
17	−1	0	0	0	1	−27.4	−58.0	−0.1	32.0	0.0
18	1	0	2	0	1	9.1	−51.0	0.0	27.0	0.0
19	−2	0	0	2	0	−205.9	−48.0	0.0	1.0	0.0
20	−2	0	2	0	1	1305.5	46.0	0.0	−24.0	0.0
21	0	0	2	2	2	7.1	−38.0	0.0	16.0	0.0
22	2	0	2	0	2	6.9	−31.0	0.0	13.0	0.0
23	2	0	0	0	0	13.8	29.0	0.0	−1.0	0.0
24	1	0	2	−2	2	23.9	29.0	0.0	−12.0	0.0
25	0	0	2	0	0	13.6	26.0	0.0	−1.0	0.0
26	0	0	2	−2	0	173.3	−22.0	0.0	0.0	0.0
27	−1	0	2	0	1	27.0	21.0	0.0	−10.0	0.0
28	0	2	0	0	0	182.6	17.0	−0.1	0.0	0.0
29	0	2	2	−2	2	91.3	−16.0	0.1	7.0	0.0
30	−1	0	0	2	1	32.0	16.0	0.0	−8.0	0.0
31	0	1	0	0	1	386.0	−15.0	0.0	9.0	0.0
32	1	0	0	−2	1	−31.7	−13.0	0.0	7.0	0.0
33	0	−1	0	0	1	−346.6	−12.0	0.0	6.0	0.0
34	2	0	−2	0	0	−1095.2	11.0	0.0	0.0	0.0
35	−1	0	2	2	1	9.5	−10.0	0.0	5.0	0.0
36	1	0	2	2	2	5.6	−8.0	0.0	3.0	0.0
37	0	−1	2	0	2	14.2	−7.0	0.0	3.0	0.0
38	0	0	2	2	1	7.1	−7.0	0.0	3.0	0.0
39	1	1	0	−2	0	−34.8	−7.0	0.0	0.0	0.0
40	0	1	2	0	2	13.2	7.0	0.0	−3.0	0.0

Table A.2: (contd.)

Index	Argument coefficient					Period	A_{0j}	A_{1j}	B_{0j}	B_{1j}
i	k_{i1}	k_{i2}	k_{i3}	k_{i4}	k_{i5}	(days)	($\times 10^{-4}$ ″)		($\times 10^{-4}$ ″)	
41	−2	0	0	2	1	−199.8	−6.0	0.0	3.0	0.0
42	0	0	0	2	1	14.8	−6.0	0.0	3.0	0.0
43	2	0	2	−2	2	12.8	6.0	0.0	−3.0	0.0
44	1	0	0	2	0	9.6	6.0	0.0	0.0	0.0
45	1	0	2	−2	1	23.9	6.0	0.0	−3.0	0.0
46	0	0	0	−2	1	−14.7	−5.0	0.0	3.0	0.0
47	0	−1	2	−2	1	346.6	−5.0	0.0	3.0	0.0
48	2	0	2	0	1	6.9	−5.0	0.0	3.0	0.0
49	1	−1	0	0	0	29.8	5.0	0.0	0.0	0.0
50	1	0	0	−1	0	411.8	−4.0	0.0	0.0	0.0
51	0	0	0	1	0	29.5	−4.0	0.0	0.0	0.0
52	0	1	0	−2	0	−15.4	−4.0	0.0	0.0	0.0
53	1	0	−2	0	0	−26.9	4.0	0.0	0.0	0.0
54	2	0	0	−2	1	212.3	4.0	0.0	−2.0	0.0
55	0	1	2	−2	1	119.6	4.0	0.0	−2.0	0.0
56	1	1	0	0	0	25.6	−3.0	0.0	0.0	0.0
57	1	−1	0	−1	0	−3232.9	−3.0	0.0	0.0	0.0
58	−1	−1	2	2	2	9.8	−3.0	0.0	1.0	0.0
59	0	−1	2	2	2	7.2	−3.0	0.0	1.0	0.0
60	1	−1	2	0	2	9.4	−3.0	0.0	1.0	0.0
61	3	0	2	0	2	5.5	−3.0	0.0	1.0	0.0
62	−2	0	2	0	2	1615.7	−3.0	0.0	1.0	0.0
63	1	0	2	0	0	9.1	3.0	0.0	0.0	0.0
64	−1	0	2	4	2	5.8	−2.0	0.0	1.0	0.0
65	1	0	0	0	2	27.8	−2.0	0.0	1.0	0.0
66	−1	0	2	−2	1	−32.6	−2.0	0.0	1.0	0.0
67	0	−2	2	−2	1	6786.3	−2.0	0.0	1.0	0.0
68	−2	0	0	0	1	−13.7	−2.0	0.0	1.0	0.0
69	2	0	0	0	1	13.8	2.0	0.0	−1.0	0.0
70	3	0	0	0	0	9.2	2.0	0.0	0.0	0.0
71	1	1	2	0	2	8.9	2.0	0.0	−1.0	0.0
72	0	0	2	1	2	9.3	2.0	0.0	−1.0	0.0
73	1	0	0	2	1	9.6	−1.0	0.0	0.0	0.0
74	1	0	2	2	1	5.6	−1.0	0.0	1.0	0.0
75	1	1	0	−2	1	−34.7	−1.0	0.0	0.0	0.0
76	0	1	0	2	0	14.2	−1.0	0.0	0.0	0.0
77	0	1	2	−2	0	117.5	−1.0	0.0	0.0	0.0
78	0	1	−2	2	0	−329.8	−1.0	0.0	0.0	0.0
79	1	0	−2	2	0	23.8	−1.0	0.0	0.0	0.0
80	1	0	−2	−2	0	−9.5	−1.0	0.0	0.0	0.0

Table A.2: (contd.)

Index	Argument coefficient					Period	A_{0j}	A_{1j}	B_{0j}	B_{1j}
i	k_{i1}	k_{i2}	k_{i3}	k_{i4}	k_{i5}	(days)	($\times 10^{-4}$ ″)		($\times 10^{-4}$ ″)	
81	1	0	2	−2	0	32.8	−1.0	0.0	0.0	0.0
82	1	0	0	−4	0	−10.1	−1.0	0.0	0.0	0.0
83	2	0	0	−4	0	−15.9	−1.0	0.0	0.0	0.0
84	0	0	2	4	2	4.8	−1.0	0.0	0.0	0.0
85	0	0	2	−1	2	25.4	−1.0	0.0	0.0	0.0
86	−2	0	2	4	2	7.3	−1.0	0.0	1.0	0.0
87	2	0	2	2	2	4.7	−1.0	0.0	0.0	0.0
88	0	−1	2	0	1	14.2	−1.0	0.0	0.0	0.0
89	0	0	−2	0	1	−13.6	−1.0	0.0	0.0	0.0
90	0	0	4	−2	2	12.7	1.0	0.0	0.0	0.0
91	0	1	0	0	2	409.2	1.0	0.0	0.0	0.0
92	1	1	2	−2	2	22.5	1.0	0.0	−1.0	0.0
93	3	0	2	−2	2	8.7	1.0	0.0	0.0	0.0
94	−2	0	2	2	2	14.6	1.0	0.0	−1.0	0.0
95	−1	0	0	0	2	−27.3	1.0	0.0	−1.0	0.0
96	0	0	−2	2	1	−169.0	1.0	0.0	0.0	0.0
97	0	1	2	0	1	13.1	1.0	0.0	0.0	0.0
98	−1	0	4	0	2	9.1	1.0	0.0	0.0	0.0
99	2	1	0	−2	0	131.7	1.0	0.0	0.0	0.0
100	2	0	0	2	0	7.1	1.0	0.0	0.0	0.0
101	2	0	2	−2	1	12.8	1.0	0.0	−1.0	0.0
102	2	0	−2	0	1	−943.2	1.0	0.0	0.0	0.0
103	1	−1	0	−2	0	−29.3	1.0	0.0	0.0	0.0
104	−1	0	0	1	1	−388.3	1.0	0.0	0.0	0.0
105	−1	−1	0	2	1	35.0	1.0	0.0	0.0	0.0
106	0	1	0	1	0	27.3	1.0	0.0	0.0	0.0

B. Coordinate Conversions

B.1 Cartesian and Ellipsoidal Coordinate Conversions

The *global* (x, y, z) ECEF Cartesian coordinates can be expressed in ellipsoidal (geodetic) coordinates (φ, λ, h), where φ and λ are, respectively, the ellipsoidal latitude and longitude, and h is the height above the ellipsoid. The centre of the ellipsoid is at the origin of the Cartesian system and its minor axis b (axis of revolution) is coincident with the z–axis, see Fig. B.1.

B.1.1 From Ellipsoidal to Cartesian Coordinates

The Cartesian coordinates of a point (x, y, z) can be obtained from the ellipsoidal coordinates (φ, λ, h) by the expressions

$$\begin{aligned} x &= (N + h)\cos\varphi\,\cos\lambda \\ y &= (N + h)\cos\varphi\,\sin\lambda \\ z &= \big((1 - e^2)N + h\big)\sin\varphi \end{aligned} \tag{B.1}$$

where N is the radius of curvature in the prime vertical

$$N = \frac{a}{\sqrt{1 - e^2\sin^2\varphi}} \tag{B.2}$$

and where the eccentricity e is related to the semi-major axis a, the semi-minor axis b and the flattening factor $f = 1 - b/a$ by

$$e^2 = \frac{a^2 - b^2}{a^2} = 2f - f^2 \tag{B.3}$$

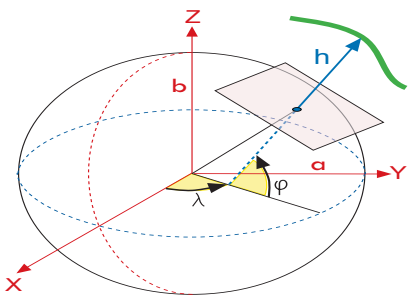


Figure B.1: Cartesian (x, y, z) and ellipsoidal (φ, λ, h) coordinates.

B.1.2 From Cartesian to Ellipsoidal Coordinates

The ellipsoidal coordinates of a point (φ, λ, h) can be obtained from the Cartesian coordinates (x, y, z) as follows.

The longitude λ is given by

$$\lambda = \arctan \frac{y}{x} \tag{B.4}$$

The latitude is computed by an iterative procedure.

1. The initial value is given by

$$\varphi_{(0)} = \arctan \left[\frac{z/p}{1 - e^2} \right] \tag{B.5}$$

   with $p = \sqrt{x^2 + y^2}$.

2. Improved values of φ, as well as the height h, are computed by iterating the equations

$$\begin{aligned} N_{(i)} &= \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi_{(i-1)}}} \\ h_{(i)} &= \frac{p}{\cos \varphi_{(i-1)}} - N_{(i)} \\ \varphi_{(i)} &= \arctan \left[\frac{z/p}{1 - \frac{N_{(i)}}{N_{(i)} + h_{(i)}} e^2} \right] \end{aligned} \tag{B.6}$$

   The iterations are repeated until the change between two successive values of $\varphi_{(i)}$ is smaller than the precision required.

These coordinate transformations are implemented in programs `sub_car2geo.f` and `sub_geo2car.f` in Volume II.

B.2 Transformations between ECEF and ENU Coordinates

Let (x, y, z) be the ECEF Cartesian coordinates of a given point P and (φ, λ) its associated latitude and longitude. Let $\Delta\mathbf{r} = (\Delta x, \Delta y, \Delta z)$ be a displacement vector from that point. These vector coordinates can be transformed from the ECEF to the *local system* East-North-Up $(\Delta\mathrm{e}, \Delta\mathrm{n}, \Delta\mathrm{u})$ coordinates, and vice versa, by two rotations as illustrated in Fig. B.2.

B.2.1 From ENU to ECEF Coordinates

1. A clockwise rotation over the East-axis by an angle $90 - \varphi$ to align the up-axis with the z-axis. That is, $\mathbf{R}_1[-(\pi/2 - \varphi)]$.

2. A clockwise rotation over the z-axis by an angle $90 + \lambda$ to align the East-axis with the x-axis. That is, $\mathbf{R}_3[-(\pi/2 + \lambda)]$.

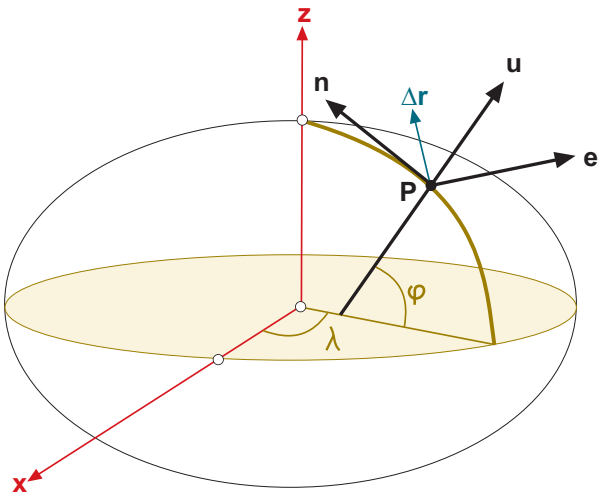


Figure B.2: Transformations between ENU and ECEF coordinates.

Thus

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \mathbf{R}_3[-(\pi/2+\lambda)]\,\mathbf{R}_1[-(\pi/2-\varphi)] \begin{bmatrix} \Delta \mathrm{e} \\ \Delta \mathrm{n} \\ \Delta \mathrm{u} \end{bmatrix} \tag{B.7}$$

which, according to expressions (3.6) of the rotation matrices given in section 3.6, yields

$$\mathbf{R}_3[-(\pi/2+\lambda)]\,\mathbf{R}_1[-(\pi/2-\varphi)] = \begin{bmatrix} -\sin\lambda & -\cos\lambda\sin\varphi & \cos\lambda\cos\varphi \\ \cos\lambda & -\sin\lambda\sin\varphi & \sin\lambda\cos\varphi \\ 0 & \cos\varphi & \sin\varphi \end{bmatrix} \tag{B.8}$$

The unit vectors in the local east, north and up directions as expressed in ECEF Cartesian coordinates are given by the columns of matrix (B.8). That is,

$$\begin{aligned} \hat{\mathbf{e}} &= (-\sin\lambda\,,\ \cos\lambda\,,\ 0) \\ \hat{\mathbf{n}} &= (-\cos\lambda\sin\varphi\,,\ -\sin\lambda\sin\varphi\,,\ \cos\varphi) \\ \hat{\mathbf{u}} &= (\cos\lambda\cos\varphi\,,\ \sin\lambda\ \cos\varphi\,,\ \sin\varphi) \end{aligned} \tag{B.9}$$

Note that, if (φ, λ) are ellipsoidal coordinates, then the vector $\hat{\mathbf{u}}$ is orthogonal to the tangent plane to the ellipsoid, which is defined by $(\hat{\mathbf{e}}, \hat{\mathbf{n}})$. If (φ, λ) are taken as the spherical latitude and longitude, then the vector $\hat{\mathbf{u}}$ is in the radial direction and $(\hat{\mathbf{e}}, \hat{\mathbf{n}})$ defines the tangent plane to the sphere.

B.2.2 From ECEF to ENU Coordinates

Taking into account the properties of the rotation matrices $\mathbf{R}_i(\alpha)$, that is $\mathbf{R}_i^{-1}(\alpha) = \mathbf{R}_i(-\alpha) = \mathbf{R}_i^T(\alpha)$, the inverse transformation of (B.7) is given by

$$\begin{bmatrix} \Delta \mathrm{e} \\ \Delta \mathrm{n} \\ \Delta \mathrm{u} \end{bmatrix} = \mathbf{R}_1[\pi/2-\varphi]\,\mathbf{R}_3[\pi/2+\lambda] \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} \tag{B.10}$$

where the transformation matrix of (B.10) is the transpose of matrix (B.8)

$$\mathbf{R}_1[\pi/2-\varphi]\,\mathbf{R}_3[\pi/2+\lambda] = \begin{bmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\cos\lambda\sin\varphi & -\sin\lambda\sin\varphi & \cos\varphi \\ \cos\lambda\cos\varphi & \sin\lambda\cos\varphi & \sin\varphi \end{bmatrix} \tag{B.11}$$

The unit vectors in the ECEF $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ directions, as expressed in ENU coordinates, are given by the columns of matrix (B.11). That is,

$$\begin{aligned} \hat{\mathbf{x}} &= (-\sin\lambda\,,\ -\cos\lambda\sin\varphi\,,\ \cos\lambda\cos\varphi) \\ \hat{\mathbf{y}} &= (\cos\lambda\,,\ -\sin\lambda\sin\varphi\,,\ \sin\lambda\cos\varphi) \\ \hat{\mathbf{z}} &= (0\,,\ \cos\varphi\,,\ \sin\varphi) \end{aligned} \tag{B.12}$$

B.3 Elevation and Azimuth Computation

Given the line-of-sight unit vector

$$\hat{\boldsymbol{\rho}} = \frac{\mathbf{r}^{sat} - \mathbf{r}_{rcv}}{\|\mathbf{r}^{sat} - \mathbf{r}_{rcv}\|} \tag{B.13}$$

where $\mathbf{r}^{sat}$ and $\mathbf{r}_{rcv}$ are the geocentric positions of the satellite and receiver, respectively, the elevation and azimuth in the local system coordinates (ENU), defined by the unit vectors $\hat{\mathbf{e}}$, $\hat{\mathbf{n}}$ and $\hat{\mathbf{u}}$, can be computed from (see Fig. B.3)

$$\begin{aligned} \hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{e}} &= \cos E\sin A \\ \hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{n}} &= \cos E\cos A \\ \hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{u}} &= \sin E \end{aligned} \tag{B.14}$$

Thus the elevation and azimuth of the satellite in the local coordinate system are given by

$$E = \arcsin(\hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{u}}) \tag{B.15}$$

$$A = \arctan\left(\frac{\hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{e}}}{\hat{\boldsymbol{\rho}}\cdot\hat{\mathbf{n}}}\right) \tag{B.16}$$

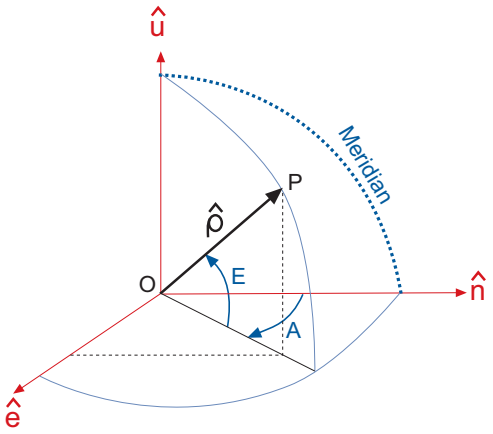


Figure B.3: Local coordinate frame showing the elevation (E) and azimuth (A).

C. Computation of Osculating Orbital Elements

A scheme with the necessary calculations to obtain the osculating orbital elements starting from the position and velocity of a satellite (in a geocentric inertial system), and vice versa, is provided as follows (see Fig. C.1 below). The International System of Units (SI) is used.

C.1 Calculation of the Orbital Elements of a Satellite from its Position and Velocity

$\boxed{(x,\, y,\, z,\, v_x,\, v_y,\, v_z) \Longrightarrow (a,\, e,\, i,\, \Omega,\, \omega,\, M)}$ (see `rv2osc.f` in Volume II)

$\mathbf{c} = \mathbf{r} \times \mathbf{v} \Longrightarrow p = c^2/\mu \Longrightarrow p$
$v^2 = \mu(2/r - 1/a) \Longrightarrow \boxed{a}$
$p = a(1 - e^2) \Longrightarrow \boxed{e}$
$\mathbf{c} = c\,\mathbf{s} \Longrightarrow \Omega = \arctan(-c_x/c_y);\; i = \arccos(c_z/c) \Longrightarrow \boxed{\Omega},\, \boxed{i}$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{R} \begin{bmatrix} r\cos V \\ r\sin V \\ 0 \end{bmatrix} = r \begin{bmatrix} \cos\Omega\cos(\omega+V) - \sin\Omega\sin(\omega+V)\cos i \\ \sin\Omega\cos(\omega+V) + \cos\Omega\sin(\omega+V)\cos i \\ \sin(\omega+V)\sin i \end{bmatrix} \Rightarrow \omega + V$$

$r = p/(1 + e\cos V) \Longrightarrow \boxed{\omega},\, V$

$\tan(E/2) = ((1-e)/(1+e))^{1/2}\tan(V/2) \Longrightarrow E$

$M = E - e\sin E \Longrightarrow \boxed{M}$

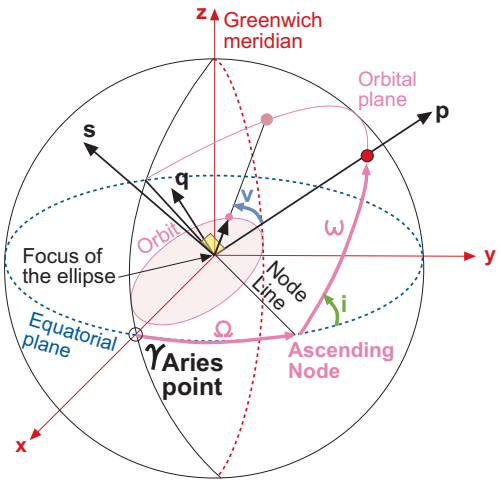


Figure C.1: Orbit in space.

C.2 Calculation of the Position and Velocity of a Satellite from Its Orbital Elements

$$(a,\, e,\, i,\, \Omega,\, \omega,\quad \underbrace{T\,;\,t}_{M \text{ or } E \text{ or } V}\quad) \Longrightarrow (x,\, y,\, z,\, v_x,\, v_y,\, v_z)$$

(program `osc2rv.f`)

$$\begin{array}{ccccc} t & \Longrightarrow & M & \Longrightarrow & E \\ M = n(t-T) & & M = E - e\sin E & & r = a(1-e\cos E) \\ & & & & \tan(V/2) = \sqrt{\dfrac{1+e}{1-e}}\tan(E/2) \end{array} \quad \Longrightarrow \quad (r, V)$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{R} \begin{bmatrix} r\cos V \\ r\sin V \\ 0 \end{bmatrix}; \quad \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = (na^2/r)\{\mathbf{q}(1-e^2)^{1/2}\cos E - \mathbf{p}\sin E\}$$

where:

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_3(-\Omega)\,\mathbf{R}_1(-i)\,\mathbf{R}_3(-\omega) \\ &= \begin{bmatrix} \cos\Omega & -\sin\Omega & 0 \\ \sin\Omega & \cos\Omega & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos i & -\sin i \\ 0 & \sin i & \cos i \end{bmatrix} \begin{bmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} p_x & q_x & s_x \\ p_y & q_y & s_y \\ p_z & q_z & s_z \end{bmatrix} = [\mathbf{p}\ \mathbf{q}\ \mathbf{s}] \end{aligned}$$

$$n^2a^3 = \mu; \quad \mu = G\,M_E = 3\,986\,004.418 \cdot 10^8\ \mathrm{m}^3/\mathrm{s}^2$$

$$c = \sqrt{a^2 - b^2}$$

$$n = 2\pi/P = 1.46 \cdot 10^{-4}\ \mathrm{rad/s}$$

$$e = c/a$$

D. Bancroft Method

The Bancroft method allows a direct solution of the receiver's position and the clock offset to be obtained, without requiring any 'a priori' knowledge of the receiver's location.

D.1 Raising and Resolution

Let PR^j be the prefit residual of satellite j, computed from equation (6.1) after removing all model terms that do not need a priori knowledge of the receiver's position,[1] that is

$$PR^j \equiv R^j + c\,\delta t^j - \mathrm{TGD}^j \tag{D.1}$$

Then, neglecting the tropospheric and ionospheric terms, as well as the multipath and receiver noise, equation (6.2) can be written as

$$PR^j = \sqrt{(x^j - x)^2 + (y^j - y)^2 + (z^j - z)^2} + c\,\delta t \tag{D.2}$$

Developing equation (D.2) further, it follows that

$$\begin{aligned}\left[x^{j\,2} + y^{j\,2} + z^{j\,2} - PR^{j\,2}\right] - 2\left[x^j x + y^j y + z^j z - PR^j c\delta t\,\right] \\ + \left[x^2 + y^2 + z^2 - (c\,\delta t)^2\right] \quad = 0\end{aligned} \tag{D.3}$$

Then, calling $\mathbf{r} = [x, y, z]^T$ and considering the inner product of Lorentz[2] the previous equation can be expressed in a more compact way as

$$\frac{1}{2}\left\langle \begin{bmatrix} \mathbf{r}^j \\ PR^j \end{bmatrix}, \begin{bmatrix} \mathbf{r}^j \\ PR^j \end{bmatrix} \right\rangle - \left\langle \begin{bmatrix} \mathbf{r}^j \\ PR^j \end{bmatrix}, \begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix} \right\rangle + \frac{1}{2}\left\langle \begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix}, \begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix} \right\rangle = 0$$

This equation can be raised for every satellite (or prefit residual PR^j).

[1] The tropospheric and ionospheric terms T^j and $\hat{\alpha}\,I^j$ cannot be included, because of the need to consider the satellite–receiver ray. Of course, after initial computation of the receiver coordinates, the method could be iterated using the ionospheric and tropospheric corrections to improve the solution.

[2] $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T\,\mathbf{M}\,\mathbf{b} = \begin{bmatrix} a_1, a_2, a_3, a_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$

If four measurements are available, then the following matrix can be written, containing all the available information on satellite coordinates and pseudoranges (every row corresponds to a satellite):

$$\mathbf{B} = \begin{bmatrix} x^1 & y^1 & z^1 & PR^1 \\ x^2 & y^2 & z^2 & PR^2 \\ x^3 & y^3 & z^3 & PR^3 \\ x^4 & y^4 & z^4 & PR^4 \end{bmatrix} \tag{D.4}$$

Then, calling

$$\Lambda = \frac{1}{2}\left\langle \begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix}, \begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix} \right\rangle, \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}$$

with

$$a_j = \frac{1}{2}\left\langle \begin{bmatrix} \mathbf{r}^j \\ PR^j \end{bmatrix}, \begin{bmatrix} \mathbf{r}^j \\ PR^j \end{bmatrix} \right\rangle$$

the four equations for pseudorange can be expressed as

$$\mathbf{a} - \mathbf{B}\,\mathbf{M}\begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix} + \Lambda\,\mathbf{1} = 0 \ , \quad \text{being} \quad \mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

whence

$$\begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix} = \mathbf{M}\mathbf{B}^{-1}(\Lambda\,\mathbf{1} + \mathbf{a}) \tag{D.5}$$

Then, taking into account the equality $\langle \mathbf{Mg}, \mathbf{Mh} \rangle = \langle \mathbf{g}, \mathbf{h} \rangle$ and that

$$\Lambda = \frac{1}{2}\left\langle \begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix}, \begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix} \right\rangle$$

from the former expression one obtains

$$\left\langle \mathbf{B}^{-1}\mathbf{1}, \mathbf{B}^{-1}\mathbf{1} \right\rangle \Lambda^2 + 2\left[\left\langle \mathbf{B}^{-1}\mathbf{1}, \mathbf{B}^{-1}\mathbf{a} \right\rangle - 1\right]\Lambda + \left\langle \mathbf{B}^{-1}\mathbf{a}, \mathbf{B}^{-1}\mathbf{a} \right\rangle = 0 \tag{D.6}$$

The previous expression is a quadratic equation in Λ (note that matrix $\mathbf{B}$ and the vector $\mathbf{a}$ are also known) and provides two solutions, one of them with the searched for solution

$$\begin{bmatrix} \mathbf{r} \\ c\,\delta t \end{bmatrix}$$

The other solution is far from Earth's surface.

D.1.1 Generalisation to the Case of *n* Measurements

If more than four observations are available, the matrix $\mathbf{B}$ is not square. However, by multiplying by $\mathbf{B}^T$, one obtains (the least squares solution)

$$\mathbf{B}^T\mathbf{a} - \mathbf{B}^T\mathbf{B}\,\mathbf{M}\begin{bmatrix}\mathbf{r}\\ c\,\delta t\end{bmatrix} + \Lambda\,\mathbf{B}^T\mathbf{1} = 0 \tag{D.7}$$

where:

$$\begin{bmatrix}\mathbf{r}\\ c\,\delta t\end{bmatrix} = \mathbf{M}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T(\Lambda\,\mathbf{1} + \mathbf{a}) \tag{D.8}$$

and then

$$\begin{aligned}&\left\langle(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{1}, (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{1}\right\rangle\Lambda^2\\ &+2\left[\left\langle(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{1}, (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{a}\right\rangle - 1\right]\Lambda\\ &\qquad+\left\langle(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{a}, (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{a}\right\rangle = 0\end{aligned} \tag{D.9}$$

E. Jacobian Matrix: Computation of Partial Derivatives

Let ρ be the geometric range between the antenna phase centre coordinates of satellite $\mathbf{r}^{sat}$ at emission time and receiver $\mathbf{r}_{rcv}$ at reception time, where $\mathbf{r}$ is given in a geocentric inertial system. Both epochs ($t^{emission}$ and $t_{reception}$) are referred to the GNSS time scale, as set by control segment clocks.

The geometric range $\rho = \|\mathbf{r}^{sat} - \mathbf{r}_{rcv}\|$ can be approximated by the first-order Taylor expansion

$$\rho = \rho_0 + \left.\frac{\partial \rho}{\partial x}\right|_{(\mathbf{r}_0,t_0)} \Delta x + \left.\frac{\partial \rho}{\partial y}\right|_{(\mathbf{r}_0,t_0)} \Delta y + \left.\frac{\partial \rho}{\partial z}\right|_{(\mathbf{r}_0,t_0)} \Delta z + \left.\frac{\partial \rho}{\partial t}\right|_{(\mathbf{r}_0,t_0)} \Delta t$$

where $\Delta x = x - x_0$, $\Delta y = y - y_0$, $\Delta z = z - z_0$ are the corrections to be applied to nominal value $\mathbf{r}_{0_{rcv}}$ to get the precise position of receiver $\mathbf{r}_{rcv}$, and Δt is a clock offset.

Calculation of the partial derivatives above depends on how ρ_0 is computed (i.e. the signal transmission time and, then, the satellite coordinates). In Chapter 5, two algorithms were developed to calculate signal emission epoch and, thus, to calculate distance ρ and the satellite coordinates at emission epoch.

The following expressions are developed to determine the partial derivatives for each of these algorithms:

1. Computation of derivative $\partial\rho/\partial t$

As mentioned in Chapter 5, pages 98 and 99, these algorithms relate emission epoch $t^{emission}$ with either the 'satellite clock' ($\iota^{emission}$) or the 'receiver clock' ($\tau_{reception}$). That is,

Pseudorange algorithm: $t^{emission} = \iota^{emission} - d\iota$

Geometric algorithm: $t^{emission} = f(\tau_{reception} - d\tau)$

where $f(\tau_{reception} - d\tau)$ gives the emission epoch computed from the geometric algorithm, as a function of reception epoch τ (according to the receiver clock).

1A. Case of pseudorange-based algorithm

If only the variation of ρ with the emission time is considered, then

$$\rho(t) \simeq \rho(\iota) + \frac{\partial \rho}{\partial t^{emission}}(t - \iota) = \rho(\iota) - \dot{\rho}(\iota) d\iota$$

where $d\iota = \iota - t$.

In this lineal approximation the error incurred when geometric range ρ is calculated using the emission epoch (according to satellite clock ι), instead of GNSS time scale t, is proportional to the rate of variation in the receiver–satellite distance and to the synchronisation error $d\iota$ of satellite clock.

In practice, offset $d\iota$ may be calculated from the navigation message. For instance, with the GPS legacy message, its accuracy is about 10 to 100 ns, depending on S/A off or on, respectively. In this case, and taking into account that $\dot{\rho} < 1\,\text{km/s}$, the error in the calculation of ρ is less than 1 mm and the term $\dot{\rho}(\iota) d\iota$ may be ignored (i.e. $\rho(t) \simeq \rho(\iota)$).

Thence, (considering only the geometric range and clock offsets)

$$\begin{aligned} R = c(\tau_{reception} - \iota^{emission}) = & \quad c(t_{reception} - t^{emission}) + c(d\tau - d\iota) \\ = & \quad \rho(t) + c(d\tau - d\iota) \\ \simeq & \quad \rho(\iota) + c(d\tau - d\iota) \end{aligned}$$

1B. Case of purely geometric algorithm

As in the previous case, linearising ρ around τ, and considering only variations in the reception time, results in

$$\rho(t) \simeq \rho(\tau) + \frac{\partial \rho}{\partial t_{reception}}(t - \tau) = \rho(\tau) - \dot{\rho}(\tau) d\tau$$

where $d\tau = \tau - t$.

In this case, the receiver clock offset $d\tau$ is unknown, but it will be estimated with the receiver coordinates in the navigation solution.[1]

Although some receivers apply clock steering to update their clocks epoch by epoch, so the offset $d\tau$ is kept within a few tens of nanoseconds, many receivers do not update their clocks until the offset reaches 1 ms. In this case, and taking into account that $\dot{\rho} < 1\,\text{km/s}$, the error in the calculation of ρ may be over several decimetres.

If, as usual, the $d\tau$ offset is determined in the navigation solution, then correction $-\dot{\rho}(\tau) d\tau$ must be taken into account in the receiver clock coefficient when the navigation equations are developed (i.e. in the design matrix or Jacobian, see page 140). Therefore, the element 1 associated to $c\delta t$ in the design matrix of equation (6.6) must be substituted by $1 - \dot{\rho}/c$. Thus

[1] It may also be extrapolated from previous estimates, although this is not necessary.

$$\begin{aligned} R = c(\tau_{reception} - \iota^{emission}) = & \quad c(t_{reception} - t^{emission}) + c(d\tau - d\iota) \\ = & \quad \rho(t) + c(d\tau - d\iota) \\ \simeq & \quad \rho(\tau) - \dot{\rho}(\tau)d\tau + c(d\tau - d\iota) \\ = & \quad \rho(\tau) + [1 - \dot{\rho}(\tau)/c]\, cd\tau - cd\iota \end{aligned}$$

Comment: Effect of receiver clock synchronisation error ($d\tau$)
Given a reception epoch, the geometric algorithm calculates the emission epoch (using an iterative procedure) considering only the receiver–satellite geometry. Indeed, it computes the signal propagation time by supposing that it has been received in a given epoch.

Then, if the reception epoch is expressed in the GNSS time scale, it will be also the same for the emission epoch obtained. If, on the other hand, it is given by the receiver clock, the synchronisation error between the receiver clock and GNSS time will introduce an error in the satellite coordinates (because they are not calculated exactly 'at GNSS emission epoch') and, as a result, an error in geometric range ρ.

Computation of range derivative $\dot{\rho} = \partial\rho/\partial t_{reception}$

Calculating the partial derivative $\partial/\partial t_{reception}$ of $\rho = c\left(t_{reception} - t^{emission}\right)$ results in (see [Webb and Zumberge, 1993])

$$\dot{\rho} = c\left(1 - \frac{\partial t^{emission}}{\partial t_{reception}}\right)$$

Equally, from the equation

$$\rho^2 = c^2\left(t_{reception} - t^{emission}\right)^2 = \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)$$

one gets

$$2c^2\left(t_{reception} - t^{emission}\right)\left(1 - \frac{\partial t^{emission}}{\partial t_{reception}}\right) = 2\left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \left(\dot{\mathbf{r}}^{sat}\frac{\partial t^{emission}}{\partial t_{reception}} - \dot{\mathbf{r}}_{rcv}\right)$$

Solving this equation for $\partial t^{emission}/\partial t_{reception}$ results in

$$\frac{\partial t^{emission}}{\partial t_{reception}} = \frac{c\rho + \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \dot{\mathbf{r}}_{rcv}}{c\rho + \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \dot{\mathbf{r}}^{sat}}$$

and therefore

$$\dot{\rho} = c\left(1 - \frac{\partial t^{emission}}{\partial t_{reception}}\right) = \frac{\left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \left(\dot{\mathbf{r}}^{sat} - \dot{\mathbf{r}}_{rcv}\right)}{\rho\left(1 + \frac{(\mathbf{r}^{sat} - \mathbf{r}_{rcv})^T}{\rho} \cdot \frac{\dot{\mathbf{r}}^{sat}}{c}\right)} = \frac{\frac{\boldsymbol{\rho}^T}{\rho} \cdot \left(\dot{\mathbf{r}}^{sat} - \dot{\mathbf{r}}_{rcv}\right)}{1 + \frac{\boldsymbol{\rho}^T}{\rho} \cdot \frac{\dot{\mathbf{r}}^{sat}}{c}}$$

Taking into account that $\dot{r}^{sat} \ll c$, the previous expression becomes:

$$\dot{\rho} = \frac{\partial\rho}{\partial t_{reception}} \simeq \frac{\boldsymbol{\rho}^T}{\rho} \cdot \left(\dot{\mathbf{r}}^{sat} - \dot{\mathbf{r}}_{rcv}\right)$$

where $\boldsymbol{\rho}^T \cdot \left(\dot{\mathbf{r}}^{sat} - \dot{\mathbf{r}}_{rcv}\right)$ can be evaluated either in a CRS or ECEF system, as Earth's rotation effect (i.e., $\boldsymbol{\omega}_E \times \boldsymbol{r}^{sat} - \boldsymbol{\omega}_E \times \boldsymbol{r}_{rcv} = \boldsymbol{\omega}_E \times \boldsymbol{\rho}$) cancels out in the scalar product.

2. Computation of derivatives $[\partial\rho/\partial x, \partial\rho/\partial y, \partial\rho/\partial z]$

As in the computation of partial derivative $\frac{\partial\rho}{\partial t}$, one must distinguish the cases where the emission epoch is calculated using the pseudorange algorithm and where the purely geometric algorithm is used, because implicit relationships among the variables involved are different in each case.

Note that in the case of the purely geometric algorithm, a nominal value of receiver coordinates $\mathbf{r}_{rcv} = (x, y, z)$ is used to compute the transmission time. Consequently, any error in these coordinates will also affect the results, and hence the geometric range ρ, through this transmission time estimate. Nevertheless, as is shown in section 2B, its effect on the spatial partial derivatives is negligible.

2A. Case of pseudorange-based algorithm

In this case, the nominal value chosen for the receiver's position does not affect, in any way, computation of the signal emission epoch, or the satellite coordinates at that instant. In other words, receiver coordinates $\mathbf{r}_{rcv} = (x, y, z)$ and satellite coordinates $\mathbf{r}^{sat} = (x^{sat}, y^{sat}, z^{sat})$ are 'independent variables'.

Consequently, as stated on page 140, when building the navigation equations

$$\frac{\partial\rho}{\partial x} = \frac{x - x^{sat}}{\rho}\ , \quad \frac{\partial\rho}{\partial y} = \frac{y - y^{sat}}{\rho}\ , \quad \frac{\partial\rho}{\partial z} = \frac{z - z^{sat}}{\rho}$$

or equivalently

$$\left[\frac{\partial\rho}{\partial x}, \frac{\partial\rho}{\partial y}, \frac{\partial\rho}{\partial z}\right] = -\frac{\boldsymbol{\rho}^T}{\rho}$$

2B. Case of purely geometric algorithm

Unlike the previous case 2A, receiver and satellite coordinates are not independent variables, but they are linked when the geometric algorithm is used to compute the emission time. Indeed, the satellite coordinates at emission epoch $\mathbf{r}^{sat}$ depend on the emission epoch $t^{emission}$ obtained by the geometric algorithm, which further depends on the receiver coordinates $\mathbf{r}_{rcv} = (x, y, z)$ used to calculate the receiver-satellite geometric distance.

Taking into account the previous considerations, computation of partial derivatives requires the reiterative use of the chain rule.

Given the geometric distance

$$\rho = c\left(t_{reception} - t^{emission}\right) = \sqrt{\boldsymbol{\rho}^T \cdot \boldsymbol{\rho}} = \left\|\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right\|$$

this results in

$$\frac{\partial\rho}{\partial x} = \frac{1}{\rho}\,\boldsymbol{\rho}^T \cdot \frac{\partial\boldsymbol{\rho}}{\partial x}$$

On the other hand,

$$\frac{\partial\boldsymbol{\rho}}{\partial x} = \frac{\partial\left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)}{\partial x} = \frac{\partial\mathbf{r}^{sat}}{\partial\mathbf{r}_{rcv}} \cdot \frac{\partial\mathbf{r}_{rcv}}{\partial x} - \frac{\partial\mathbf{r}_{rcv}}{\partial x}$$

Taking into account that the satellite coordinates at emission epoch $\mathbf{r}^{sat}$ depend on emission epoch $t^{emission}$ obtained by the geometric algorithm, which further depends on the receiver coordinates $\mathbf{r}_{rcv}$ used to calculate the receiver–satellite geometric distance, then

$$\frac{\partial \mathbf{r}^{sat}}{\partial \mathbf{r}_{rcv}} = \frac{\partial \mathbf{r}^{sat}}{\partial t^{emission}} \cdot \frac{\partial t^{emission}}{\partial \mathbf{r}_{rcv}} = \dot{\mathbf{r}}^{sat} \cdot \frac{\partial t^{emission}}{\partial \mathbf{r}_{rcv}}$$

Derivative $\partial t^{emission}/\partial \mathbf{r}_{rcv}$ may be obtained implicitly by differentiating the equation

$$\rho^2 = c^2 \left(t_{reception} - t^{emission}\right)^2 = \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)$$

leading to

$$c^2 \left(t_{reception} - t^{emission}\right) \left(-\frac{\partial t^{emission}}{\partial \mathbf{r}_{rcv}}\right) = \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \left(\frac{\partial \mathbf{r}^{sat}}{\partial \mathbf{r}_{rcv}} - \frac{\partial \mathbf{r}_{rcv}}{\partial \mathbf{r}_{rcv}}\right)$$

Taking into account in this expression that

$$\partial \mathbf{r}^{sat}/\partial \mathbf{r}_{rcv} = \dot{\mathbf{r}}^{sat} \cdot \partial t^{emission}/\partial \mathbf{r}_{rcv} \quad \text{and} \quad \partial \mathbf{r}_{rcv}/\partial \mathbf{r}_{rcv} = I_3$$

gives

$$\begin{aligned} \frac{\partial t^{emission}}{\partial \mathbf{r}_{rcv}} &= \frac{\left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T}{c^2 \left(t_{reception} - t^{emission}\right) + \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \dot{\mathbf{r}}^{sat}} \\ &= \frac{\left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T}{c\rho \left(1 + \frac{\left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T}{\rho} \cdot \frac{\dot{\mathbf{r}}^{sat}}{c}\right)} \end{aligned}$$

Finally, substituting into the equation for $\partial \rho/\partial x$,

$$\frac{\partial \rho}{\partial x} = \frac{1}{\rho} \boldsymbol{\rho}^T \cdot \frac{\partial \boldsymbol{\rho}}{\partial x} = -\frac{1}{\rho} \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T \cdot \left[I_3 - \frac{\dot{\mathbf{r}}^{sat} \cdot \left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T}{c\rho \left(1 + \frac{\left(\mathbf{r}^{sat} - \mathbf{r}_{rcv}\right)^T}{\rho} \cdot \frac{\dot{\mathbf{r}}^{sat}}{c}\right)}\right] \frac{\partial \mathbf{r}_{rcv}}{\partial x}$$

where $\partial \mathbf{r}_{rcv}/\partial x = (1, 0, 0)^T$.

In general, considering that $\boldsymbol{\rho} = \mathbf{r}^{sat} - \mathbf{r}_{rcv}$, then

$$\left[\frac{\partial \rho}{\partial x}, \frac{\partial \rho}{\partial y}, \frac{\partial \rho}{\partial z}\right] = -\frac{\boldsymbol{\rho}^T}{\rho} \cdot \left[I_3 - \frac{\frac{\dot{\mathbf{r}}^{sat}}{c} \cdot \frac{\boldsymbol{\rho}^T}{\rho}}{1 + \frac{\boldsymbol{\rho}^T}{\rho} \cdot \frac{\dot{\mathbf{r}}^{sat}}{c}}\right]$$

The term joining the matrix I_3 in the previous expression accounts for the contribution of errors in the estimation of satellite coordinates due to errors in the nominal values of receiver coordinates (being the $t^{emission}$ computed using the geometric algorithm). This term, however, can be neglected taking into account that $\dot{r}^{sat} \ll c$.

Thence

$$\left[\frac{\partial \rho}{\partial x}, \frac{\partial \rho}{\partial y}, \frac{\partial \rho}{\partial z}\right] \simeq -\frac{\boldsymbol{\rho}^T}{\rho}$$

which leads to the same result as in the previous case 2A.

3. Jacobian matrix

Taking into account the above results, the matrix $\mathbf{G}$ associated with the navigation equation system $\mathbf{y} = \mathbf{G}\,\mathbf{x}$, see page 140, is given by the following.

3A. Case of pseudorange-based algorithm

$$\mathbf{G} = \begin{bmatrix} \frac{x_0 - x^1}{\rho_0^1} & \frac{y_0 - y^1}{\rho_0^1} & \frac{z_0 - z^1}{\rho_0^1} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{x_0 - x^n}{\rho_0^n} & \frac{y_0 - y^n}{\rho_0^n} & \frac{z_0 - z^n}{\rho_0^n} & 1 \end{bmatrix}$$

3B. Case of purely geometric algorithm

$$\mathbf{G} = \begin{bmatrix} \frac{x_0 - x^1}{\rho_0^1} & \frac{y_0 - y^1}{\rho_0^1} & \frac{z_0 - z^1}{\rho_0^1} & 1 - \frac{\dot{\rho}_0^1}{c} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{x_0 - x^n}{\rho_0^n} & \frac{y_0 - y^n}{\rho_0^n} & \frac{z_0 - z^n}{\rho_0^n} & 1 - \frac{\dot{\rho}_0^1}{c} \end{bmatrix}$$

where:

$$\dot{\rho} = \frac{\partial \rho}{\partial t_{reception}} \simeq \frac{\boldsymbol{\rho}^T}{\rho} \cdot \left(\dot{\mathbf{r}}^{sat} - \dot{\mathbf{r}}_{rcv} \right)$$

can be evaluated either in a CRS or ECEF system.

Comment: The emission time computed using the geometric algorithm will also be corrupted by other delays (from the ionosphere, troposphere…) but together these are at most a few hundreds of metres and their effect is negligible in this context.

F. Antenna Phase Centres for GPS Broadcast Orbits and Clocks

The following table is from the National Geospatial-Intelligence Agency (NGA) *http://earth-info.nga.mil/GandG/sathtml/gpsdoc2009_10a.html*.

```
Satellite antenna offset (satellite body centred coordinates, meters)
  Block II  PRN's  - Delta x=  0.2794, Delta y=  0.0000, Delta z=  0.9519

  Block IIA PRN's  - Delta x=  0.2794, Delta y=  0.0000, Delta z=  0.9519

  Block IIR PRN 02 - Delta x= -0.0099, Delta y=  0.0061, Delta z= -0.0820
  Block IIR PRN 11 - Delta x=  0.0019, Delta y=  0.0011, Delta z=  1.5141
  Block IIR PRN 13 - Delta x=  0.0024, Delta y=  0.0025, Delta z=  1.6140
  Block IIR PRN 14 - Delta x=  0.0018, Delta y=  0.0002, Delta z=  1.6137
  Block IIR PRN 16 - Delta x= -0.0098, Delta y=  0.0060, Delta z=  1.6630
  Block IIR PRN 18 - Delta x= -0.0098, Delta y=  0.0060, Delta z=  1.5923
  Block IIR PRN 19 - Delta x= -0.0079, Delta y=  0.0046, Delta z= -0.0180
  Block IIR PRN 20 - Delta x=  0.0022, Delta y=  0.0014, Delta z=  1.6140
  Block IIR PRN 21 - Delta x=  0.0023, Delta y= -0.0006, Delta z=  1.5840
  Block IIR PRN 22 - Delta x=  0.0018, Delta y= -0.0009, Delta z=  0.0598
  Block IIR PRN 23 - Delta x= -0.0088, Delta y=  0.0035, Delta z=  0.0004
  Block IIR PRN 28 - Delta x=  0.0019, Delta y=  0.0007, Delta z=  1.5131

  Block IIR-M PRN 01 - Delta x=  0.01245, Delta y= -0.00038, Delta z= -0.02283
  Block IIR-M PRN 05 - Delta x=  0.00292, Delta y= -0.00005, Delta z= -0.01671
  Block IIR-M PRN 07 - Delta x=  0.00127, Delta y=  0.00025, Delta z=  0.00056
  Block IIR-M PRN 12 - Delta x= -0.01016, Delta y=  0.00587, Delta z= -0.09355
  Block IIR-M PRN 15 - Delta x= -0.00996, Delta y=  0.00579, Delta z= -0.01227
  Block IIR-M PRN 17 - Delta x= -0.00996, Delta y=  0.00599, Delta z= -0.10060
  Block IIR-M PRN 29 - Delta x= -0.01012, Delta y=  0.00591, Delta z= -0.01512
  Block IIR-M PRN 31 - Delta x=  0.00160, Delta y=  0.00033, Delta z= -0.05750
```

G. List of Acronyms

AltBOC Alternate Binary Offset Carrier

AMCS Alternate Master Control Station

ANTEX ANTenna EXchange format

AOAD/M_T Allen Osborne Associates Dorne Margolin Model T

APC Antenna Phase Centre

ARNS Aeronautical Radio Navigation Service

ARP Antenna Reference Point

ASCII American Standard Code for Information Interchange

A/S Anti-Spoofing

BDT BeiDou Time

BIPM Bureau International des Poids et Mesures

BIE Bureau International de l'Heure

BLUE Best Linear Unbiased Minimum Variance Estimator

BM Benchmark

BOC Binary Offset Carrier

BPSK Binary Phase Shift Keying

BRS Barycentric Reference System

C/A Coarse/Acquisition

CASC China Aerospace Science and Technology Corporation

CBOC Composite Binary Offset Carrier

CC-M Central Synchroniser

CDDIS Crustal Dynamics Data Information System

CDMA Code Division Multiple Access

CEP Celestial Ephemeris Pole

CGCS2000 China Geodetic Coordinate System 2000

CIO Conventional International Origin

CIP Celestial Intermediate Pole

CIS Conventional Inertial System

CODE Centre for Orbit Determination in Europe

C/NAV Commercial Navigation Message

CORS Continuously Operating Reference Stations

CRC Cyclic Redundancy Check

CRF Celestial Reference Frame

CRS Conventional Celestial Reference System

CS Commercial Service

CTP Conventional Terrestrial Pole

CTS Conventional Terrestrial System

DCB Differential Code Bias

DoD US Department of Defense

DOP Dilution Of Precision

ECEF Earth-Centred, Earth-Fixed

ECI Earth-Centred Inertial

EGNOS European Geostationary Navigation Overlay System

EMR Energy Mines and Resources

ENU East North Up

EOP Earth Orientation Parameters

ERP Earth Rotation Parameters

ESA European Space Agency

ET Ephemeris Time

EUV Extreme UltraViolet

FDMA Frequency Division Multiple Access

FEC Forward Error Correction

FOC Full Operational Capability

F/NAV Freely accessible Navigation Message

F-PPP Fast Precise Point Positioning

GA Ground Antennas

gAGE Research group of Astronomy and Geomatics

GAST Greenwich Apparent Sidereal Time

GCC Ground Control Centre

GCS Ground Control Segment

GEO Geostationary Orbit

GGSP Galileo Geodetic Service Provider

GIOVE Galileo In-Orbit Validation

GIM Global Ionospheric Maps

Glonass GLObal NAvigation Satellite System

GLNT Glonass Time

GMT Greenwich Mean Time

GMS Ground Mission Segment

GMST Greenwich Mean Sidereal Time

G/NAV Governmental Navigation Message

GNSS Global Navigation Satellite System

GPS Global Positioning System

GPST GPS Time

GRAPHIC Group and Phase Ionospheric Calibration

GRS Geocentric Reference System

GSS Galileo Sensor Stations

GST Galileo System Time

GTRF Galileo Terrestrial Reference Frame

GUI Graphical User Interface

HPPS High-Precision Positioning Service

HTML HyperText Markup Language

IAU International Astronomical Union

ICAO International Civil Aviation Organization

ICD Interface Control Document

ICRF International Celestial Reference Frame

ICRS International Celestial Reference System

IDF Ionospheric Disturbance Flag

IERS International Earth Rotation and Reference Systems Service

IGEX International Glonass EXperiment

IGS International GNSS Service

IGSO Inclined Geosynchronous Satellite Orbit

I/NAV Integrity Navigation Message

IOV In-Orbit Validation

IPP Ionospheric Pierce Point

ITRF International Terrestrial Reference Frame

ITRS International Terrestrial Reference System

ITU International Telecommunications Union

ITU-R International Telecommunications Union Radiocommunication Sector

JD Julian Day

JPL Jet Propulsion Laboratory

LAST Local Apparent Sidereal Time

LC Carrier phase ionosphere-free combination

LEO Low Earth Orbit

LMST Local Mean Sidereal Time

LS Laser Station

L2CL L2 Civil Long

L2CM L2 Civil Moderate

MBOC Multiplexed Binary Offset Carrier

MC Satellite Mass Centre

MCS Master Control Station

MEO Medium Earth Orbit

MJD Modified Julian Day

MM Monument Marker

MODIP MOdified DIP latitude

MS Monitoring Stations

MSE Mean-Square Error

MW Melbourne–Wübbena

NASA National Aeronautics and Space Administration

NAV Navigation Message

NEU North East Up

NGA National Geospatial-Intelligence Agency

NIST National Institute of Standards and Technology

OS Open Service

PCV Phase Centre Variation

PPP Precise Point Positioning

PPS Precise Positioning Service

PRN Pseudo-Random Noise

PRS Public Regulated Service

PZ-90 Parametry Zemli 1990 (Parameters of the Earth 1990)

QPSK Quadrature Phase-Shifted Keying

QZSS Quasi-Zenith Satellite System

RINEX Receiver INdependent EXchange format

RK4 Fourth-order Runge–Kutta

RMS Root Mean Square

RMSE Root Mean Square Error

RNSS Radionavigation Satellite Service

RTK Real-Time Kinematics

S/A Selective Availability

SAR Search and Rescue

SBAS Satellite-Based Augmentation System

SCC System Control Centre

SEP Spherical Error Probable

SI International System of Units

SINEX Solution (Software/technique) INdependent EXchange format

SIS Signal In Space

SoL Safety-of-Life

SP3 Standard Product #3

SPP Standard Point Positioning

SPS Standard Positioning Service

SSTL Surrey Satellite Technology Ltd.

STEC Slant Total Electron Content

SU Soviet Union

SVN Space Vehicle Number

SW Synchronisation Word

TAI International Atomic Time

TCAR Three Carrier Ambiguity Resolution

TCB Barycentric Coordinate Time, from the French Temps Coordonné Barycentrique

TCG Geocentric Coordinate Time, from the French Temps Coordonné Géocentrique

TDB Barycentric Dynamic Time, from the French Temps Dinamique Barycentrique

TDT Terrestrial Dynamic Time

TEC Total Electron Content

TECU Total Electron Content Unit

TLM TeLeMetry

TMBOC Time Multiplexed BOC

TGD Total Group Delay

TRF Terrestrial Reference Frame

TRS Conventional Terrestrial Reference System

TT Terrestrial Time

UERE User Equivalent Range Error

ULS Mission Uplink Stations

UNE Up North East

UPC Technical University of Catalonia

USAF United States Air Force

USNO United States Naval Observatory

UT Universal Time

UTC Coordinated Universal Time

UT0 Universal Time 0

UT1 Universal Time 1

UT2 Universal Time 2

VLBI Very Long Baseline Interferometry

WAAS Wide Area Augmentation System

WARTK Wide-Area Real-Time Kinematics

WGS World Geodetic System

WGS-84 World Geodetic System 84

WLS Weighted Least Squares

ZTD Zenith Tropospheric Delay

Bibliography

[Arbesser-Rastburg, B., 2006] Arbesser-Rastburg, B., 2006. The Galileo Single Frequency Ionospheric Correction Algorithm. *http://sidc.oma.be/esww3/presentations/Session4/Arbesser.pdf*.

[Ashby, N., 2003] Ashby, N., 2003. Relativity in the Global Positioning System. *http://relativity.livingreviews.org/Articles/lrr-2003-1/download/lrr-2003-1Color.pdf*.

[Avila-Rodriguez, 2008] Avila-Rodriguez, J., 2008. On Generalized Signal Waveforms for Satellite Navigation. PhD thesis. University of Munich, Munich, Germany.

[Avila-Rodriguez et al., 2007] Avila-Rodriguez, J., Hein, G., Wallner, S., Issler, J., Ries, L., Lestarquit, L., De Latour, A., Godet, J., Bastide, F., Pratt, T. and Owen, J., 2007. The MBOC Modulation. A Final Touch for the Galileo Frequency and Signal Plan. *InsideGNSS*. September/October 2007.

[Avila-Rodriguez et al., 2006] Avila-Rodriguez, J., Wallner, S., Hein, G., Rebeyrol, E., Julien, O., Macabiau, C., Ries, L. DeLatour, A., Lestarquit, L. and Issler, J., 2006. CBOC- An Implementation of MBOC. In: *Proceedings of 1rst CNES-ESA Workshop on Galileo Siganals and Signal Processing*, Toulouse, France.

[Banville et al., 2008] Banville, S., Santerre1, R., Cocard, M. and R.B., L., 2008. Satellite and Receiver Phase Bias Calibration for Undifferenced Ambiguity Resolution. *Proceedings Institute of Navigation, National Technical Meeting. San Diego, CA, USA.*

[Bar-Sever, Y.E., 1994] Bar-Sever, Y.E., 1994. Improvement to the GPS Attitude Control Subsystem Enables Predictable Attitude during Eclipse Seasons. *http://igscb.jpl.nasa.gov/mail/igsmail/1994/msg00166.html*.

[BeiDou-SIS-ICD, 2012] BeiDou-SIS-ICD, 2012. Technical report. BeiDou Navigation Satellite System Signal in Space. Interface Control Document. Open Service Signal B1I (Version 1.0), December 2012.

[BeiDou-SIS-ICD-Test, 2011] BeiDou-SIS-ICD-Test, 2011. Technical report. BeiDou Navigation Satellite System Signal in Space. Interface Control Document (Test version), December 2011.

[Bierman, 1976] Bierman, G., 1976. *Factorization Methods for Discrete Sequential Estimation*. Academic Press, New York, USA.

[Black and Eisner, 1984] Black, H. and Eisner, A., 1984. Correcting Satellite Doppler Data for Tropospheric Effects. *Journal of Geophysical Research* **89**, pp. 2616–2626.

[Blewitt, 1990] Blewitt, G., 1990. An Automatic Editing Algorithms for GPS Data. *Geophysical Research Letters* **17**(3), pp. 199–202.

[Boucher and Altamimi, 2001] Boucher, C. and Altamimi, Z., 2001. ITRS, PZ-90 and WGS 84: Current Realizations and the Related Transformation Parameters. *Journal of Geodesy* **75**, pp. 613–619.

[Cao et al., 2008] Cao, C., Jing, G. and Luo, M., 2008. COMPASS Satellite Navigation System Development. *PNT Challenges and Opportunities Symposium. Stanford, California, USA, 5–6 November 2008.*

[Cheng and Dang, 2011] Cheng, P. and Dang, Y., 2011. China National Report on Geodesy (2007-2010). Chinese National Committee for IAG, Chinese National Committee for The International Union of Geodesy and Geophysics, Beijing, China.

[Chengqi, R., 2011] Chengqi, R., 2011. The Statement at the Press Conference of the State Council Information Office, 27 December 2011. *http://www.china.com.cn/zhibo/zhuanti/ch-xinwen/2011-12/27/content_24259420.htm.*

[Collins, 1999] Collins, J., 1999. Assessment and Development of a Tropospheric Delay Model for Aircraft Users of the Global Positioning System. MScE thesis, University of New Brunswick, Fredericton, New Brunswick, Canada.

[Crawford, 1968] Crawford, F., 1968. *Berkeley Physics Course, Vol. 3, Waves*. McGraw-Hill,, New York, USA.

[Davies, 1989] Davies, K., 1989. *Ionosphere Radio.* IEE Electromagnetic Waves Series 31. Peter Peregrinus Ltd., Stevenage, UK.

[Di Giovanni and Radicella, 1990] Di Giovanni, G. and Radicella, S., 1990. An Analytical Model of the Electron Density Profile in the Ionosphere. *Advances in Space Research* **10**(11), pp. 27–30.

[Diggelen, 2007] Diggelen, F. v., 2007. GNSS Accuracy: Lies, Damn Lies, and Statistics. *GPS World* **18**(1), pp. 26–32.

[Enge and Misra, 1999] Enge, P. and Misra, P., 1999. Special Issue on Global Positioning System. *Proceedings of the IEEE* **87**(1), pp. 3–15.

[Fantino et al., 2011] Fantino, M., Lo Presti, L. and Pini, M., 2011. Digital Signal Processing in GNSS Receivers, in *Handbook of Position Location: Theory, Practice, and Advances.* Eds. S.A. Zekavat and R. M. Buehrer, John Wiley & Sons, Inc., NJ, USA.

[Galileo SIS ICD, EU, 2010] Galileo SIS ICD, EU, 2010. Galileo Open Service Signal In Space Control Document (OS SIS IDC), issue 1.1. *http://ec.europa.eu/enterprise/policies/satnav/galileo/open-service/index_en.htm.*

[Ge et al., 2008] Ge, M., Gendt, M., Rothacher, M., Shi, C. and Liu, J., 2008. Resolution of GPS Carrier-Phase Ambiguities in Precise Point Positioning (PPP) with Daily Observations. *Journal of Geodesy* **82**, pp. 389–399.

[Gendt et al., 2011] Gendt, G., Altamimi, Z., Dach, R., Springer, T. and The GNSS Prototype Team., 2011. GGSP: Realisation and Maintenance of the Galileo Terrestrial Reference Frame. *Advances in Space Research* **47**(2), pp. 174–185.

[Gendt, G. and Schmid, R., 2005] Gendt, G. and Schmid, R., 2005. IGSMAIL-5189: Planned Changes to IGS Antenna Calibrations. *http://igscb.jpl.nasa.gov/mail/igsmail/2005/msg00111.html*.

[GLONASS ICD, 1998] GLONASS ICD, 1998. Technical report. v.4.0.

[GLONASS ICD, 2008] GLONASS ICD, 2008. Technical report. v.5.1.

[GPS/PPS-PS, 2007] GPS/PPS-PS, 2007. Global Positioning System Precise Positioning Service. Performance Standard. *http://www.gps.gov/technical/ps/2007-PPS-performance-standard.pdf*.

[GPS/SPS-PS, 2008] GPS/SPS-PS, 2008. Global Positioning System Standard Positioning Service. Performance Standard. *http://www.gps.gov/technical/ps/2008-SPS-performance-standard.pdf*.

[GPS/SPS-SS, 1995] GPS/SPS-SS, 1995. Global Positioning System Standard Positioning Service. Signal Specification. *http://www.navcen.uscg.gov/pubs/gps/sigspec/gpssps1.pdf*.

[Hajj et al., 2002] Hajj, G., Kursinsky, E., Romans, L., Bertier, W. and Leroy, S., 2002. A technical description of atmospheric sounding by GPS occultation. *Journal of Atmospheric and Solar-Terrestrial Physics* **64**, pp. 451–469.

[Hein, G.W., 2010] Hein, G.W., 2010. Galileo and GNSS under Development 2 – Compass/Beidou, IRNSS and QZSS. ESA International Summer School on GNSS, Slettestrand, Denmark, September 2010.

[Hein, G.W., 2011] Hein, G.W., 2011. Towards Multi-Frequency Multi-Constellation GNSS. ESA International Summer School on GNSS, Berchtesgaden, Germany, 20–30 July 2011.

[Hernández-Pajares, 2004] Hernández-Pajares, M., 2004. IGS Ionosphere WG Status Report: Performance of IGS Ionosphere TEC Maps - Position Paper-. In: IGS Workshop and Symposium 2004 "Celebrating a Decade of the International GPS Service, Bern, Switzerland.

[HJS, 2007] Hernández-Pajares, M., Juan, M. and Sanz, J., 2007. Second-Order Ionospheric Term in GPS: Implementation and Impact on Geodetic Estimates. *Journal of Geophysical Research* **112**, pp. 1–16.

[HJS, 2010] Hernández-Pajares, M., Juan, M. and Sanz, J., 2010. Enhanced Precise Point Positioning using Galileo and GPS (EPPP). TN8: Recommendations for a PPP Processing Facility. ESA report (ESA ITT AO/1-5823/08/NL/AT), ESOC/ESA, Noordwijk, The Netherlands.

[HJS and Colombo, 1999] Hernández-Pajares, M., Juan, M., Sanz, J. and Colombo, O., 1999. New Approaches in Global Ionospheric Determination Using Ground GPS Data. *Journal of Atmospheric and Solar Terrestrial Physics* **61**, pp. 1237–1247.

[HJS et al., 2010] Hernández-Pajares, M., Juan, M., Sanz, J., Aragon-Angel, A., Samson, J., Tossaint, M., Odijk, D., Teunissen, P., De Bakker, P., Verhagen, S. and Van der Marel, H., 2010. Wide-Area RTK. *InsideGNSS* **5**(2), pp. 36–46.

[HJS et al., 2011] Hernández-Pajares, M., Juan, M. and Sanz, J., Samson, J. and Tossaint, M., 2011. Method, Apparatus and System for Determining a Position of an Object Having a Global Navigation Satellite System Receiver by Processing Undifferenced Data Like Carrier Phase Measurements and External Products Like Ionosphere Data. international patent application PCT/EP2011/001512 (ESA ref.: ESA/PAT/566).

[Hofmann-Wellenhof et al., 2003] Hofmann-Wellenhof, B., Legat, K. and Wieser, M., 2003. *Navigation: Principles of Positioning and Guidance*. Springer, Vienna, Austria.

[Hofmann-Wellenhof et al., 2008] Hofmann-Wellenhof, B., Lichtenegger, H. and Wasle, E., 2008. *GNSS – Global Navigation Satellite Systems*. Springer, Vienna, Austria.

[Hoque and Jakowsky, 2008] Hoque, M. and Jakowsky, N., 2008. Estimate of Higher Order Ionospheric Errors in GNSS Positioning. *Radio Science* **43**, pp. 1–15.

[IS-GPS-200, 2010] IS-GPS-200, 2010. GPS Interface Specification IS-GPS-200. Revision E., *http://www.gps.gov/technical/icwg/IS-GPS-200E.pdf*.

[Januszewski, 2010] Januszewski, J., 2010. Time, Its Scales and Part in Satellite Navigation Systems. *Scientific Journals, Maritime University of Szczecin* **20**(92), pp. 52–59.

[Juan et al., 2010] Juan, J., Hernández-Pajares, M. and Sanz, J., 2010. Precise Real Time Orbit Determination and Time Synchronization (PRTODTS). TN2 & TN3: GPS PRTODTS Design and Validation Documents, v0.0. ESA report (ESA ITT AO/1-5823/08/NL/AT), ESOC/ESA, Noordwijk, The Netherlands.

[Juan et al., 2012a] Juan, M., Hernández-Pajares, M., Sanz, J., Ramos-Bosch, P., Aragon-Angel, A., Orus, R., Ochieng, W., Feng, S., Coutinho, P., Samson, J. and Tossaint, M., 2012a. Enhanced Precise Point Positioning for GNSS Users. *IEEE Transactions on Geoscience and Remote Sensing* PP, issue 99, pp. 1–10.

[Juan et al., 2012b] Juan, M., Hernández-Pajares, M., Sanz, J., Samson, J., Tossaint, M., Aragon-Angel, A. and Salazar, D., 2012b. Wide Area RTK: A Satellite Navigation System Based on Precise Real-Time Ionospheric Modelling. *Radio Science* **47**, pp. 1–14.

[Kaplan, 1996] Kaplan, E., 1996. *Understanding GPS: Principles and Applications*. Artech House, Boston, MA, USA.

[Kaplan, 1981] Kaplan, G., 1981. Nutation Data from Radio Interferometer Observations. *Proceedings of the Ninth International Symposium on Earth Tides. E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart* pp. 519–526.

[Kim and Langley, 2000] Kim, D. and Langley, R., 2000. GPS Ambiguity Resolution and Validation: Methodologies, Trends and Issues. Proceedings of 7th GNSS Workshop - International Symposium on GPS/GNSS, Seoul, Korea.

[Klobuchar, 1987] Klobuchar, J., 1987. Ionospheric Time-Delay Algorithms for Single-Frequency GPS Users. *IEEE Transactions on Aerospace and Electronic Systems* **AES-23**(3), pp. 325–331.

[Konno, 2007] Konno, H., 2007. Design of an aircraft landing system using dual-frequency gnss. PhD dissertation. Stanford University, Stanford, CA, USA.

[Kouba and Héroux, 2000] Kouba, J. and Héroux, P., 2000. GPS Precise Point Positioning using IGS Orbit Products. Geodetic Survey Division, Natural Resources Canada.

[Laurichesse and Mercier, 2007] Laurichesse, D. and Mercier, F., 2007. Integer Ambiguity Resolution on Undifferenced GPS Phase Measurements and Its Application to PPP. *Proceedings of the 20th International Technical Meeting of the Satellite Division, Institute of Navigation (GNSS 2007). Fort Worth, TX, USA*.

[Leick, 1994] Leick, A., 1994. *GPS Satellite Surveying*. Wiley-Interscience, New York, USA.

[Lewandowski et al., 2006] Lewandowski, W., Matsakis, D., Panfilo, G. and Tavella, P., 2006. The Evaluation of Uncertainties in [UTC-UTC(k)]. *Metrologia* **43**, pp. 278–286.

[Mader, 1999] Mader, G., 1999. GPS Antenna Calibration at the National Geodetic Survey. *GPS Solutions* **3**(1), pp. 50–58.

[Marini, 1972] Marini, J., 1972. Correction of Satellite Tracking Data for an Arbitrary Tropospheric Profile. *Radio Science* **7**(2), pp. 223–231.

[Martin-Neira et al., 1995] Martin-Neira, M., Toledo, M. and Pelaez, A., 1995. The Null Space Method for GPS Integer Ambiguity Resolution. *Proceedings of DSNS'95, Bergen, Norway*, Paper No. 31, 8pp.

[McCarthy, 1996] McCarthy, D., 1996. IERS Conventions (1996). IERS Technical Note 21. IERS Convention Center, US Naval Observatory., Washington, DC, USA.

[McCarthy and Petit, 2004] McCarthy, D. and Petit, G., 2004. IERS Conventions (2003). IERS Technical Note 32. IERS Convention Center., Frankfurt am Main.

[McCarthy, D. and Petit, G., 2009] McCarthy, D. and Petit, G., 2009. IERS Conventions (2009) (draft), Chapter 9. *ftp://tai.bipm.org/iers/convupdt/chapter9/icc9.pdf*.

[McDonald and Hegarty, 2000] McDonald, K. and Hegarty, C., 2000. Modernization GPS Performance Capabilities. In: *Proceedings of ION 56th Annual Meeting, San Diego, California, USA*, pp. 242–249.

[Mervat, 1995] Mervat, L., 1995. Ambiguity Resolution Techniques in Geodetic and Geodynamic Applications of the Global Positioning System. PhD thesis. University of Bern, Bern, Switzerland.

[Misra and Enge, 2001] Misra, P. and Enge, P., 2001. *Global Positioning System: Signals, Measurements and Performance*. Ganga-Jamuna Press, Lincoln, MA, USA.

[Montenbruck and Eberhard, 2005] Montenbruck, O. and Eberhard, G., 2005. *Satellite Orbits: Models, Methods, Applications*. Springer, Berlin, Germany.

[Montenbruck and Ramos-Bosch, 2007] Montenbruck, O. and Ramos-Bosch, P., 2007. Precision Real-Time Navigation of LEO Satellites Using Global Positioning System Measurements. *GPS Solutions* **12**(3), pp. 187–198.

[Niell, 1996] Niell, A., 1996. Global Mapping Functions for the Atmosphere Delay at Radio Wavelengths. *Journal of Geophysical Research* **101**, pp. 3227–3246.

[Parkinson et al., 1996] Parkinson, B., Spilker, J. and Enge, P., 1996. *Global Positioning System, Vols I and II, Theory and Applications*. American Institute of Aeronautics, Reston, VA, USA.

[Powe, M., 2006] Powe, M., 2006. Introduction to Galileo. PowerPoint Presentation, Progeny.

[Rawer, 1963] Rawer, K., 1963. *Propagation of Decameter Waves (HF-band) in Meteorological and Astronomical Influences on Radio Wave Propagation*. Ed. Landmark, B. Pergamon Press, New York.

[Revnivykh, 2007] Revnivykh, S., 2007. Glonass Status and Progress. In: Minutes of the 47th CGSIC Meeting, Forth Worth, TX, USA.

[RTCA-MOPS, 2006] RTCA-MOPS, 2006. Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment.rtca document 229-C.

[Sanz et al., 2010] Sanz, J., Juan, J. and Hernández-Pajares, M., 2010. *GNSS Data Processing laboratory exercises*. 4th ESA GNSS Summer School., Slettestrand (Aalborg), Denmark. *http://www.gage.upc.edu/tutorials*.

[Sanz et al., 2012] Sanz, J., Juan, J. and Hernández-Pajares, M., 2012. *Analysis of propagation effects from GNSS observables based on laboratory exercises.* Course on Propagation effects, channel models and related error sources on GNSS, ESA/ESAC., Madrid, Spain. *http://www.gage.upc.edu/tutorials*.

[Schaer, S. and Steingenberger, P., 2006] Schaer, S. and Steingenberger, P., 2006. Determination and Use of GPS Differential Code Bias Values. *http://acc.igs.org/biases/signal-biases_igsws10.pdf*.

[Seeber, 1993] Seeber, G., 1993. *Satellite Geodesy: Foundations, Methods, and Applications*. Walter de Gruyter,, Berlin, Germany.

[Seidelmann, 1982] Seidelmann, P., 1982. 1980 IAU Theory of Nutation: The Final Report of the IAU Working Group on Nutation. *Celestial Mechanics* **27**, pp. 79–106.

[Strang and Borre, 1997] Strang, G. and Borre, K., 1997. Linear Algebra, Geodesy, and GPS. Wellesley-Cambridge Press, Wellesley MA.

[Tapley et al., 2004] Tapley, B., Schutz, B. and Born, G., 2004. *Statistical Orbit Determination.* Academic Press, Amsterdam, The Netherlands.

[Teunissen, 1996] Teunissen, P., 1996. GPS Carrier Phase Ambiguity Fixing Concepts, in GPS for Geodesy. *Lecture Notes in Earth Sciences* **60**, pp. 263–335.

[Teunissen et al., 1997] Teunissen, P., De Jonge, P. and Tiberius, J., 1997. Performance of LAMBDA Method for Fast GPS Ambiguity Resolution. *Navigation* **44**(3), pp. 373–383.

[United Nations, 2004] United Nations, 2004. Report of the Action Team on Global Navigation Satellite Systems (GNSS) – Follow-up to the Third United Nations Conference on the Exploration and Peaceful Uses of Outer Space (UNISPACE III). *http://www.unoosa.org/pdf/publications/st_space_24E.pdf*

[United Nations, 2010] United Nations, 2010. *Current and Planned Global and Regional Navigation Satellite Systems and Satellite-Based Augmentation Systems of the International Committee on Global Navigation Satellite Systems Providers' Forum*. United Nations, New York, USA.

[Urlichich et al., 2011] Urlichich, Y., Subbotin, V., Stupak, G., Dvorkin, V., Povalyaev, A., Karutin, S. and Bakitko, R., 2011. Glonass Modernization. *GPS World* **22**(11), pp. 34–39.

[Ventura-Traveset, J. and Flament, 2006] Ventura-Traveset, J. and Flament, D., 2006. *EGNOS: The European Geostationary Navigation Overlay System: A Cornerstone of Galileo*. ESA Publications Division, ESTEC. Series ESA SP 1303., Noordwijk, the Netherlands.

[Webb and Zumberge, 1993] Webb, F. and Zumberge, J., 1993. *An Introduction to GIPSY/OASIS-II*. Jet Propulsion Laboratory, Pasadena, CA, USA.

[Wu et al., 1993] Wu, J., Wu, S., Hajj, G., Bertiguer, W. and Lichten, S., 1993. Effects of Antenna Orientation on GPS Carrier Phase Measurements. *Manuscripta Geodaetica* **18**, pp. 91–98.

[Xu, 2007] Xu, G., 2007. *GPS: Theory, Algorithms, and Applications*. Springer, Berlin, Germany.

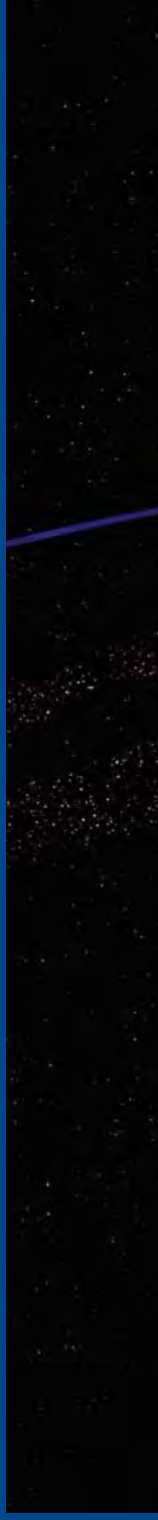
[Yoder et al., 1981] Yoder, C., Williams, J. and Parke, M., 1981. Tidal Variations of Earth Rotation. *Journal of Geophysical Research* **86**, pp. 881–891.

[Yunck, 1993] Yunck, T., 1993. Coping with the atmosphere and ionosphere in precise satellite and ground positioning. *Environmental Effects on Spacecraft Trajectories and Positioning, Geophysical Monograph Series*, Edited by A. Valance Jones. AGU. Washington, DC, USA.

Index

About the authors

Jaume Sanz Subirana, PhD, José Miguel Juan Zornoza, PhD, and Manuel Hernández-Pajares, PhD, are Professors at the Technical University of Catalonia (UPC). They created the Research Group of Astronomy and Geomatics (gAGE) at the UPC in 1987. Their current research interests are in the areas of GNSS data processing algorithms, GNSS ionospheric sounding, satellite and ground-based augmentation systems (SBAS and GBAS), and high-accuracy GNSS navigation. This last is one of the main research topics of the gAGE/UPC group, where new algorithms for wide-area high-accuracy navigation are being developed and tested, leading to the Wide Area RTK (WARTK) and the Fast-Precise Point Positioning (Fast-PPP) techniques. The authors share several international patents on these techniques, founded by ESA. As a result of the gAGE/UPC research group's activities in GNSS, in 2009 they founded the spin-off company gAGE-NAV S.L.



ESA Member States

Austria
Belgium
Czech Republic
Denmark
Finland
France
Germany
Greece
Ireland
Italy
Luxembourg
Netherlands
Norway
Poland
Portugal
Romania
Spain
Sweden
Switzerland
United Kingdom

An ESA Communications Production